ÁLGEBRA LINEAR

por

KENNETH HOFFMAN
Associate Professor of Mathematics
Massachusetts Institute of Technology

e

RAY KUNZE
Associate Professor of Mathematics
Washington University
St. Louis, Mo.

Tradução de

ADALBERTO PANOBIANCO BERGAMASCO

EDITÔRA DA UNIVERSIDADE DE SÃO PAULO
EDITÔRA POLÍGONO

Título do original:
Linear Algebra

Capa de
Studio Ito

512.897 Hoffman, Kenneth

Álgebra linear, por Kenneth Hoffman e Ray Kunze; traduzido por Adalberto P. Bergamasco. S. Paulo, Ed. Univ. de S. Paulo e Polígono, 1970.
356p. ilus.

Álgebra linear

PREFÁCIO

Nosso propósito original ao escrever êste livro foi o de fornecer um texto para o curso de graduação de álgebra linear no Massachussetts Institute of Technology. Êste curso era destinado ao terceiro ano dos optantes de matemática. Atualmente, cêrca de três quartos dos alunos especializam-se em ciências ou engenharia e variam de calouros a estudantes de pós-graduação.

Concessão alguma se fêz ao fato de a maioria dos alunos não estar interessada primordialmente em matemática. Isso porque acreditamos que um curso de matemática não deveria fornecer a estudantes de ciências ou engenharia um amontoado de métodos, e sim proporcionar a êles uma compreensão dos conceitos matemáticos fundamentais.

Por outro lado, estivemos profundamente conscientes da grande variação de conhecimentos que os estudantes poderiam possuir e, em particular, do fato de terem os estudantes tido muito pouca experiência com o raciocínio matemático abstrato. Por essa razão, evitamos a introdução de muitas idéias abstratas logo no início do livro. Como complemento, incluímos um Apêndice, onde são apresentadas idéias básicas tais como conjunto, função e relação de equivalência. Achamos mais proveitoso não insistir nessas idéias independentemente, e sim aconselhar os estudantes a lerem o Apêndice à medida que surjam tais idéias.

Em todo o livro incluímos uma grande diversidade de exemplos dos conceitos importantes que ocorrem. O estudo de tais exemplos é de fundamental importância e tende a minimizar o número de estudantes que conseguem repetir definições, teoremas e demonstrações em ordem lógica, sem apreender o significado dos conceitos abstratos. O livro contém também uma ampla variedade de exercícios graduados (em tôrno de quinhentos), variando desde aplicações rotineiras aos que solicitarão até os melhores alunos. Pretende-se que êsses exercícios sejam parte importante do texto.

O Capítulo 1 trata de sistemas de equações lineares e sua resolução por meio de operações elementares sôbre linhas de matrizes. Tem sido nosso costume despender seis aulas nessa matéria. Isso proporciona ao estudante um esbôço das origens da álgebra linear e das técnicas de cálculo computacionais necessárias ao entendimento de exemplos das idéias mais abstratas ocorrentes nos capítulos posteriores. O Capítulo 2 discorre sôbre espaços vetoriais, subespaços, bases e dimensão. O Capítulo 3 trata das transformações lineares, sua álgebra, sua representação por matrizes, bem como isomorfismo, funcionais lineares e espaços duais. O Capítulo 4 define a álgebra dos polinômios sôbre um corpo, os ideais naquela álgebra e a decomposição de um polinômio em fatôres primos. O Capítulo 5 desenvolve determinantes de matrizes quadradas, sendo o determinante encarado como uma função n-linear alternada das linhas de uma matriz. Os Capítulos 6 e 7 contêm uma discussão dos conceitos básicos para a análise de uma transformação linear isolada sôbre um espaço vetorial de dimensão finita, a análise de transformações diagonalizáveis, o conceito das partes diagonalizável e nilpotente de uma transformação mais geral e as formas canônicas racional e de Jordan. O Capítulo 8 considera com algum detalhe espaços de dimensão finita com produto interno. Êle cobre, em particular, a geometria básica e o estudo dos operadores auto-adjuntos, positivos, unitários e normais. O Capítulo 9 discute formas bilineares, enfatizando as formas canônicas para formas simétricas e anti-simétricas, assim com o grupo que conserva uma forma não-degenerada.

A interdependência dos capítulos é como segue. Os Capítulos 1 e 2 e a maior parte do Capítulo 3 são básicos para o livro todo. Os Capítulos 4 e 5 também são fundamentais; entretanto, podem ser tratados de uma forma mais abreviada se o professor deseja passar aos capítulos subseqüentes mais ràpidamente. Os Capítulos 6 e 7 são uma unidade. Os Capítulos 8 e 9 são independentes entre si e não necessitam dos Capítulos 6 e 7 (exceto talvez das primeiras páginas do Capítulo 6). O Capítulo 9 não depende do Capítulo 4 nem do 5 tampouco.

Somos gratos a nossos colegas, em particular, aos Professôres Louis Howard e Daniel Kan e Doutores Harry Furstenberg e Edward Thorp, por suas tantas e tão profícuas sugestões. Pela preparação do manuscrito, agradecemos às Srtas. Betty Ann Sargent e Phyllis Ruby, que datilografaram as notas originais; Srta. Judith Bowers, que datilografou o manuscrito final; e à equipe da Prentice-Hall, Inc.

Cambridge, Massachussetts KENNETH HOFFMAN

Waltham, Massachussetts RAY KUNZE

SUMÁRIO

Capítulo 1. EQUAÇÕES LINEARES 1
1.1. Corpos comutativos 1
1.2. Sistemas de equações lineares 3
1.3. Matrizes e operações elementares sôbre linhas 6
1.4. Matrizes linha-reduzidas à forma em escada 12
1.5. Multiplicação de matrizes 18
1.6. Matrizes inversíveis 24

Capítulo 2. ESPAÇOS VETORIAIS 30
2.1. Espaços vetoriais 30
2.2. Subespaços 37
2.3. Bases e dimensão 43
2.4. Coordenadas 50
2.5. Resumo de linha-equivalência 57
2.6. Cálculos concernentes a subespaços 61

Capítulo 3. TRANSFORMAÇÕES LINEARES 67
3.1. Transformações lineares 67
3.2. A álgebra das transformações lineares 73
3.3. Isomorfismo 83
3.4. Representação de transformações por matrizes 85
3.5. Funcionais lineares 97
3.6. Anuladores 104
3.7. A transposta de uma transformação 109

Capítulo 4. POLINÔMIOS 115
4.1. Álgebras 115
4.2. A álgebra dos polinômios 117
4.3. Interpolação de Lagrange 122
4.4. Ideais de polinômios 126
4.5. A decomposição de um polinômio em fatôres primos . . 132

Capítulo 5. DETERMINANTES 138
5.1. Anéis comutativos 138
5.2. Funções determinantes 139
5.3. Permutações e a unicidade dos determinantes 149
5.4. Propriedades adicionais dos determinantes 156

Capítulo 6. DECOMPOSIÇÕES EM SOMAS DIRETAS INVARIANTES 166
6.1. Decomposições em somas diretas 166

6.2. Valores característicos e vetores característicos 177
6.3. Operadores diagonalizáveis 184
6.4. O teorema da decomposição primária 193

CAPÍTULO 7. AS FORMAS RACIONAL E DE JORDAN . . . 201
7.1. Subespaços cíclicos e anuladores 201
7.2. O teorema da decomposição racional 205
7.3. A forma de Jordan 219
7.4. Resumo: operadores semi-simples 226

CAPÍTULO 8. ESPAÇOS COM PRODUTO INTERNO. 235
8.1. Produtos internos 235
8.2. Espaços com produto interno 242
8.3. Funcionais lineares e adjuntos 251
8.4. Operadores positivos. 260
8.5. Operadores unitários. 271
8.6. Operadores normais 282
8.7. O teorema espectral 288
8.8. Diagonalização simultânea de operadores normais 302

CAPÍTULO 9. FORMAS BILINEARES 305
9.1. Formas bilineares 305
9.2. Formas bilineares simétricas 314
9.3. Formas bilineares anti-simétricas 323
9.4. Grupos que conservam formas bilineares 327

APÊNDICE 335
A.1. Conjuntos 335
A.2. Funções 337
A.3. Relações de equivalência 341
A.4. Espaços quocientes 344
A.5. Relações de equivalência em álgebra linear 348

BIBLIOGRAFIA 350
ÍNDICE 351

CAPÍTULO 1

EQUAÇÕES LINEARES

1.1 Corpos Comutativos

Supomos que o leitor tenha familiaridade com a álgebra elementar dos números reais e complexos. Para uma grande parte dêste livro as propriedades algébricas dos números que usaremos podem ser fàcilmente deduzidas da pequena lista abaixo de propriedades da adição e da multiplicação. Indicamos por F o conjunto dos números reais ou o conjunto dos números complexos.

(1) A adição é comutativa,

$$x + y = y + x$$

para todos x e y em F.

(2) A adição é associativa,

$$x + (y + z) = (x + y) + z$$

para todos x, y e z em F.

(3) Existe um único elemento 0 (zero) em F tal que $x + 0 = x$, para todo x em F.

(4) A cada x em F corresponde um único elemento $(-x)$ em F tal que $x + (-x) = 0$.

(5) A multiplicação é comutativa,

$$xy = yx$$

para todos x e y em F.

(6) A multiplicação é associativa,

$$x(yz) = (xy)z$$

para todos x, y e z em F.

(7) Existe um único elemento não-nulo 1 (um) em F tal que $x1 = x$, para todo x em F.

(8) A cada x não-nulo em F corresponde um único x^{-1} (ou $1/x$) em F tal que $xx^{-1} = 1$.

(9) A multiplicação é distributiva em relação à adição; isto é, $x(y + z) = xy + xz$, para todos x, y e z em F.

Suponhamos que se tenha um conjunto F de objetos $x, y, z, \ldots$ e duas operações sôbre os elementos de F como segue. A primeira operação, denominada adição, associa a cada par de elementos x, y em F um elemento $(x + y)$ em F; a segunda operação, denominada multiplicação, associa a cada par x, y um elemento xy em F; e estas duas operações satisfazem as condições (1)-(9) acima. O conjunto F, munido destas duas operações, é então denominado um **corpo comutativo***. A grosso modo, um corpo é um conjunto munido de algumas operações sôbre seus objetos, as quais se comportam como a adição, subtração, multiplicação e divisão usuais de números no sentido de que elas obedecem às nove regras de álgebra acima relacionadas. Com as propriedades usuais da adição e multiplicação, o conjunto C dos números complexos é um corpo, como o é o conjunto R dos números reais.

Na maior parte dêste livro, os "números" que usamos podem ser os elementos de qualquer corpo F. Para permitir esta generalização, usaremos a palavra "escalar" ao invés de "número". O leitor não perderá muito se supuser sempre que o corpo de escalares seja um subcorpo do corpo dos números complexos. Um **subcorpo** do corpo C é um conjunto F de números complexos que é um corpo em relação às operações usuais de adição e multiplicação de números complexos. Isto significa que 0 e 1 estão no conjunto F e que se x e y são elementos de F então $(x + y)$, $-x$, xy e x^{-1} (se $x \neq 0$) também o são. Um exemplo de um subcorpo desta natureza é o corpo R dos números reais; de fato, se identificarmos os números reais com os números complexos $(a + ib)$ para os quais $b = 0$, o 0 e o 1 do corpo complexo são números reais e, se x e y são reais, $(x + y)$, $-x$, xy, e x^{-1} (se $x \neq 0$) também o são. Daremos outros exemplos abaixo. O objetivo de nossa discussão sôbre subcorpos é essencialmente o seguinte: quando trabalhamos com escalares de um certo subcorpo de C, a realização das operações de adição, subtração, multiplicação ou divisão sôbre êstes escalares não nos tira daquele subcorpo.

(*) Neste livro, sempre teremos corpos comutativos, portanto abreviaremos a denominação escrevendo simplesmente corpos. (N. do T.)

Exemplo 1. O conjunto dos **inteiros positivos**: 1, 2, 3, ..., não é um subcorpo de C, por diversas razões. Por exemplo, 0 não é um inteiro positivo; para qualquer inteiro positivo n, $-n$ não é um inteiro positivo; para qualquer inteiro n, exceto 1, $1/n$ não é um inteiro positivo.

Exemplo 2. O conjunto dos **inteiros**: ..., -2, -1, 0, 1, 2, ..., não é um subcorpo de C, pois para um inteiro n, $1/n$ não é um inteiro a menos que n seja 1 ou -1. Com as operações usuais de adição e multiplicação, o conjunto dos inteiros satisfaz tôdas as condições (1)—(9) com exceção da condição (8).

Exemplo 3. O conjunto dos **números racionais**, isto é, números da forma p/q, onde p e q são inteiros e $q \neq 0$, é um subcorpo do corpo dos números complexos. A divisão, que não é possível dentro do conjunto dos inteiros, pode ser feita dentro do conjunto dos números racionais. O leitor interessado deve verificar que qualquer subcorpo de C contém todos os números racionais.

Exemplo 4. O conjunto de todos os números complexos da forma $x + y\sqrt{2}$, onde x e y são racionais, é um subcorpo de C. Deixamos a cargo do leitor a verificação dêste fato.

1.2 Sistemas de Equações Lineares

Suponhamos que F seja um corpo. Consideremos o problema da determinação de n escalares (elementos de F) $x_1, \ldots, x_n$ que satisfaçam as condições

$$
\begin{array}{c}
A_{11}x_1 + A_{12}x_2 + \ldots + A_{1n}x_n = y_1 \\
A_{21}x_1 + A_{22}x_2 + \ldots + A_{2n}x_n = y_2 \\
\vdots \\
A_{m1}x_1 + A_{m2}x_2 + \ldots + A_{mn}x_n = y_m
\end{array} \tag{1-1}
$$

onde $y_1, \ldots, y_m$ e A_{ij}, $1 \leq i \leq m$, $1 \leq j \leq n$, são elementos dados de F. Denominamos (1—1) um **sistema de m equações lineares a n incógnitas.** Toda n-upla $(x_1, \ldots, x_n)$ de elementos de F que satisfaz a cada uma das equações em (1—1) é dita uma **solução** do sistema. Se $y_1 = y_2 = \ldots = y_m = 0$, dizemos que o sistema é **homogêneo**, ou que cada uma das equações é homogênea.

O método mais importante para determinar as soluções de um sistema de equações lineares é talvez o método de eliminação. Podemos ilustrar êste método com o sistema homogêneo

$$\begin{aligned} 2x_1 - x_2 + x_3 &= 0 \\ x_1 + 3x_2 + 4x_3 &= 0. \end{aligned}$$

Somando (—2) vêzes a segunda equação à primeira equação obtemos

$$-7x_2 - 7x_3 = 0$$

ou $x_2 = -x_3$. Somando 3 vêzes a primeira equação à segunda equação obtemos

$$7x_1 + 7x_3 = 0$$

ou $x_1 = -x_3$. Assim, concluímos que se (x_1, x_2, x_3) é uma solução então $x_1 = x_2 = -x_3$. Reciprocamente, pode-se verificar prontamente que tôda terna dêste tipo é uma solução. Assim, o conjunto de soluções consiste de tôdas as ternas $(-a, -a, a)$.

Determinamos as soluções dêste sistema de equações "eliminando incógnitas", isto é, multiplicando equações por escalares e daí somando-as para obter equações em que alguns dos x_j não estavam presentes. Queremos formalizar ligeiramente êste processo para que possamos compreender por que êle funciona e para que possamos efetuar os cálculos necessários para resolvermos um sistema de uma maneira organizada.

Para o sistema arbitrário (1—1), suponhamos selecionar m escalares, multiplicar a j-ésima equação por c_j e daí somar. Obtemos a equação

$$\begin{aligned} (c_1A_{11} + \ldots + c_mA_{m1})x_1 + \ldots + (c_1A_{1n} + \ldots + c_mA_{mn})x_n \\ = c_1y_1 + \ldots + c_my_m. \end{aligned}$$

Tal equação será por nós denominada uma **combinação linear** das equações em (1—1). Evidentemente, tôda solução do sistema de equações (1—1) também será uma solução desta nova equação. Esta é a idéia fundamental do processo de eliminação. Se temos outro sistema de equações lineares

$$\begin{array}{ll} & B_{11}x_1 + \ldots + B_{1n}x_n = z_1 \\ (1\text{—}2) & \quad\vdots \qquad\qquad\qquad \vdots \quad\;\; \vdots \\ & B_{k1}x_1 + \ldots + B_{kn}x_n = z_k \end{array}$$

no qual cada uma das k equações é uma combinação linear das equações em (1—1), então tôda solução de (1—1) é uma solução dêste nôvo sistema. É claro que pode acontecer que algumas soluções de (1—2) não sejam soluções de (1—1). Isto òbviamente não acontece se cada equação do sistema original é uma combinação linear das equações do nôvo sistema. Diremos que dois sistemas de equações lineares são **equivalentes** se cada equação de cada sistema fôr uma combinação linear das equações do outro sistema. Podemos então enunciar formalmente nossas observações como segue.

Teorema 1. *Sistemas equivalentes de equações lineares têm exatamente as mesmas soluções.*

Para o processo de eliminação ser eficiente na determinação das soluções de um sistema como (1—1), é necessário que se saiba, formando combinações lineares das equações dadas, como produzir um sistema equivalente de equações que seja mais fácil de resolver. Na próxima seção discutiremos um método para conseguir isto.

Exercícios

1. Verificar que o conjunto dos números complexos descrito no Exemplo 4 é um subcorpo de C.

2. Seja F o corpo dos números complexos. Os dois seguintes sistemas de equações lineares são equivalentes? Em caso afirmativo, exprimir cada equação de cada sistema como uma combinação linear das equações do outro sistema.

$$\begin{aligned} x_1 - x_2 &= 0 \\ 2x_1 + x_2 &= 0 \end{aligned} \qquad \begin{aligned} 3x_1 + x_2 &= 0 \\ x_1 + x_2 &= 0 \end{aligned}$$

3. Repetir o Exercício 2 para os seguintes sistemas de equações:

$$\begin{aligned} -x_1 + x_2 + 4x_3 &= 0 \\ x_1 + 3x_2 + 8x_3 &= 0 \\ \tfrac{1}{2}x_1 + x_2 + \tfrac{5}{2}x_3 &= 0 \end{aligned} \qquad \begin{aligned} x_1 \qquad - x_3 &= 0 \\ x_2 + 3x_3 &= 0 \end{aligned}$$

4. Repetir o Exercício 2 para os sistemas seguintes:

$$\begin{aligned} 2x_1 + (-1 + i)x_2 \qquad + x_4 &= 0 \\ 3x_2 - 2ix_3 + 5x_4 &= 0 \end{aligned} \qquad \begin{aligned} \left(1 + \frac{i}{2}\right) x_1 + 8x_2 - ix_3 - x_4 &= 0 \\ \tfrac{2}{3}x_1 - \tfrac{1}{2}x_2 + x_3 + 7x_4 &= 0 \end{aligned}$$

5. Seja F um conjunto que contém exatamente dois elementos, 0 e 1. Definamos uma adição e uma multiplicação pelas tábuas:

+	0	1
0	0	1
1	1	0

.	0	1
0	0	0
1	0	1

Verificar que o conjunto F, munido destas duas operações, é um corpo.

6. Demonstrar que se dois sistemas homogêneos de equações lineares a duas incógnitas têm as mesmas soluções, então êles são equivalentes.

7. Demonstrar que todo subcorpo do corpo dos números complexos contém todos os números racionais.

1.3 Matrizes e Operações Elementares sôbre Linhas

Não podemos deixar de observar que, ao formarmos combinações lineares de equações lineares, não há necessidade de continuarmos escrevendo as "incógnitas" $x_1, \ldots, x_n$, uma vez que, na realidade, fazemos cálculos apenas com os coeficientes A_{ij} e os escalares y_i. Abreviaremos o sistema (1—1) por

$$AX = Y$$

onde

$$A = \begin{bmatrix} A_{11} & \ldots & A_{1n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ A_{m1} & \ldots & A_{mn} \end{bmatrix}$$

$$X = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \quad \text{e} \quad Y = \begin{bmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_m \end{bmatrix}.$$

Denominamos A a **matriz dos coeficientes** do sistema. Rigorosamente falando, a tabela retangular acima exibida não é uma matriz, mas sim uma representação de um matriz. Uma $m \times n$ **matriz sôbre o corpo** F é uma função A do conjunto dos pares de inteiros (i, j), $1 \leq i \leq m$, $1 \leq j \leq n$, no corpo F. Os **elementos** da matriz A são os escalares $A(i, j) = A_{ij}$ e, com bastante freqüência, o mais conveniente é descrever a matriz exibindo seus elementos numa tabela retangular com m linhas e n colunas, como acima. Assim X (acima) é, ou define uma $n \times 1$ matriz e Y é uma $m \times 1$ matriz. Por ora, $AX = Y$ nada mais é que uma notação taquigráfica para o nosso sistema de equações lineares. Posteriormente, quando houvermos definido uma multiplicação de matrizes, aquilo significará que Y é o produto de A por X.

Queremos agora considerar operações sôbre linhas da matriz A que correspondam a formar combinações lineares das equações do sistema $AX = Y$. Restringiremos nossa atenção a três **operações elementares sôbre as linhas** de uma $m \times n$ matriz A sôbre o corpo F:

(1) multiplicação de uma linha de A por um escalar c não-nulo;

(2) substituição da r-ésima linha de A pela linha r mais c vêzes a linha s, sendo c um escalar arbitrário e $r \neq s$;

(3) transposição de duas linhas de A.

Uma operação elementar sôbre linhas é assim um tipo particular de função (regra) e que associa a cada $m \times n$ matriz A uma $m \times n$ matriz $e(A)$. Pode-se descrever e com precisão nos três casos acima como segue:

(1) $e(A)_{ij} = A_{ij}$ se $i \neq r$, $e(A)_{rj} = cA_{rj}$.

(2) $e(A)_{ij} = A_{ij}$ se $i \neq r$, $e(A)_{rj} = A_{rj} + cA_{sj}$.

(3) $e(A)_{ij} = A_{ij}$ se i é diferente de r e de s, $e(A)_{rj} = A_{sj}$, $e(A)_{sj} = A_{rj}$.

Ao definirmos $e(A)$ não importa muito o número de colunas de A, mas o número de linhas de A é crucial. Por exemplo, deve-se tomar cuidado ao decidir o que significa trocar as linhas 5 e 6 de uma 5×5 matriz. Para evitar tais complicações, convencionaremos que uma operação elementar e sôbre as linhas é definida sôbre a classe das $m \times n$ matrizes sôbre F, para um certo m fixo mas para n arbitrário. Em outras palavras, um e particular é definido sôbre a classe das matrizes com m linhas sôbre F.

Uma razão para nos restringirmos a êstes três tipos simples de operações sôbre linhas é que, tendo efetuado uma tal operação e sôbre uma matriz A, podemos voltar a A efetuando uma operação do mesmo tipo sôbre $e(A)$.

Teorema 2. *A cada operação elementar sôbre linhas* e *corresponde uma operação elementar sôbre linhas* e_2, *do mesmo tipo que* e, $e_1(e(A)) = e(e_1(A)) = A$ *para qualquer* A. *Em outras palavras, a operação (função) inversa de uma operação elementar sôbre linhas existe e é uma operação elementar sôbre linhas do mesmo tipo.*

Demonstração. (1) Suponhamos que e seja a operação que multiplica a r-ésima linha de uma matriz pelo escalar não-nulo c. Seja e_1 a operação que multiplica a linha r por c^{-1}. (2) Suponhamos que e seja a operação que substitui a linha r pela linha r mais c vêzes a linha s, $r \neq s$. Seja e_1 a operação que substitui a linha r pela linha r mais $(-c)$ vêzes a linha s. (3) Se e transpõe as linhas r e s, seja $e_1 = e$. Em cada um dêstes três casos temos evidentemente $e_1(e(A)) = e(e_1(A)) = A$ para cada A.

Definição. *Se* A *e* B *são* m $\times$ n *matrizes sôbre o corpo* F, *dizemos que* B **é linha-equivalente a** A *se* B *pode ser obtida de* A *por uma seqüência finita de operações elementares sôbre linhas.*

Usando o Teorema 2, o leitor deverá achar fácil verificar o que segue. Tôda matriz é linha-equivalente a si mesma; se B é linha-equivalente a A, então, A é linha-equivalente a B; se B é linha-equivalente a A e C é linha-equivalente a B, então C é linha-equivalente a A. Em outras palavras, a linha-equivalência é uma relação de equivalência (ver Apêndice).

Teorema 3. *Se* A *e* B *são* m $\times$ n *matrizes linha-equivalentes, os sistemas homogêneos de equações lineares* AX = 0 *e* BX = 0 *têm exatamente as mesmas soluções.*

Demonstração. Suponhamos passar de A para B por meio de uma seqüência finita de operações elementares sôbre linhas:

$$A = A_0 \rightarrow A_2 \rightarrow \ldots \rightarrow A_k = B.$$

Basta demonstrar que os sistemas $A_jX = 0$ e $A_{j+1}X = 0$ têm as mesmas soluções, isto é, que uma operação elementar sôbre linhas não altera o conjunto das soluções.

Assim, suponhamos que B seja obtida de A por uma única operação elementar sôbre linhas. Qualquer que seja o tipo da operação, (1), (2) ou (3), cada equação dos sistema $BX = 0$ será uma combinação linear das equações do sistema $AX = 0$. Como a inversa de uma operação elementar sôbre linhas é uma operação elementar sôbre linhas, cada equação em $AX = 0$ também será uma combinação linear das equações em $BX = 0$. Logo êstes dois sistemas são equivalentes e, pelo Teorema 1, têm as mesmas soluções.

Exemplo 5. Suponhamos que F seja o corpo dos números racionais e que

$$A = \begin{bmatrix} 2 & -1 & 3 & 2 \\ 1 & 4 & 0 & -1 \\ 2 & 6 & -1 & 5 \end{bmatrix}.$$

Efetuaremos uma seqüência finita de operações elementares sôbre as linhas de A, indicando por números entre parênteses o tipo de operação efetuada.

$$\begin{bmatrix} 2 & -1 & 3 & 2 \\ 1 & 4 & 0 & -1 \\ 2 & 6 & -1 & 5 \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & -9 & 3 & 4 \\ 1 & 4 & 0 & -1 \\ 2 & 6 & -1 & 5 \end{bmatrix} \xrightarrow{(2)}$$

$$\begin{bmatrix} 0 & -9 & 3 & 4 \\ 1 & 4 & 0 & -1 \\ 0 & -2 & -1 & 7 \end{bmatrix} \xrightarrow{(1)} \begin{bmatrix} 0 & -9 & 3 & 4 \\ 1 & 4 & 0 & -1 \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \end{bmatrix} \xrightarrow{(2)}$$

$$\begin{bmatrix} 0 & -9 & 3 & 4 \\ 1 & 0 & -2 & 13 \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & 0 & \frac{15}{2} & -\frac{55}{2} \\ 1 & 0 & -2 & 13 \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \end{bmatrix} \xrightarrow{(1)}$$

$$\begin{bmatrix} 0 & 0 & 1 & -\frac{11}{3} \\ 1 & 0 & -2 & 13 \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & 0 & 1 & -\frac{11}{3} \\ 1 & 0 & 0 & \frac{17}{3} \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \end{bmatrix} \xrightarrow{(2)}$$

$$\begin{bmatrix} 0 & 0 & 1 & -\frac{11}{3} \\ 1 & 0 & 0 & \frac{17}{3} \\ 0 & 1 & 0 & -\frac{5}{3} \end{bmatrix}$$

A linha-equivalência de A com a matriz final na seqüência acima nos diz em particular que as soluções de

$$\begin{aligned} 2x_1 - x_2 + 3x_3 + 2x_4 &= 0 \\ x_1 + 4x_2 - x_4 &= 0 \\ 2x_1 + 6x_2 - x_3 + 5x_4 &= 0 \end{aligned}$$

e

$$\begin{aligned} x_3 - \tfrac{11}{3} x_4 &= 0 \\ x_1 + \tfrac{17}{3} x_4 &= 0 \\ x_2 - \tfrac{5}{3} x_4 &= 0 \end{aligned}$$

são exatamente as mesmas. No segundo sistema é evidente que atribuindo um valor racional arbitrário c a x_4, obtemos uma solução $(-\frac{17}{3}c, \frac{5}{3}c, \frac{11}{3}c, c)$, e também que tôda solução é desta forma.

Exemplo 6. Suponhamos que F seja o corpo dos números complexos e que

$$A = \begin{bmatrix} -1 & i \\ -i & 3 \\ 1 & 2 \end{bmatrix}.$$

Ao efetuarmos operações sôbre linhas freqüentemente convém combinar várias operações do tipo (2). Com isto em mente

$$\begin{bmatrix} -1 & i \\ -i & 3 \\ 1 & 2 \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & 2+i \\ 0 & 3+2i \\ 1 & 2 \end{bmatrix} \xrightarrow{(1)} \begin{bmatrix} 0 & 1 \\ 0 & 3+2i \\ 1 & 2 \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Assim o sistema de equações

$$\begin{aligned} -x_1 + ix_2 &= 0 \\ -ix_1 + 3x_2 &= 0 \\ ix_1 + 2x_2 &= 0 \end{aligned}$$

possui apenas a solução trivial $x_1 = x_2 = 0$.

Nos Exemplos 5 e 6 é óbvio que não efetuamos operações sôbre linhas ao acaso. Nossa escolha de operações sôbre linhas foi motivada por um desejo de simplificar a matriz dos coeficientes de uma maneira análoga à "eliminação de incógnitas" no sistema de equações lineares. Coloquemos agora uma definição formal do tipo da matriz à qual estávamos tentando chegar.

Definição. *Uma* m $\times$ n *matriz* R *é dita* **linha-reduzida** *se*:

(a) *o primeiro elemento não-nulo em cada linha não-nula de* R *é igual a* 1;

(b) *cada coluna de* R *que contém o primeiro elemento não-nulo de alguma linha tem todos os seus outros elementos nulos.*

Exemplo 7. Um exemplo de uma matriz linha-reduzida é a $n \times n$ **matriz** (quadrada) **unidade** I. Esta é a $n \times n$ matriz definida por

$$I_{ij} = \delta_{ij} = \begin{cases} 1, & \text{se } i = j \\ 0, & \text{se } i \neq j. \end{cases}$$

Esta é a primeira de muitas ocasiões em que usaremos o **símbolo de Kronecker** (δ).

Nos Exemplos 5 e 6, as matrizes finais nas seqüências apresentadas são matrizes linha-reduzidas. Dois exemplos de matrizes que *não* são linha-reduzidas são:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 2 & 1 \\ 1 & 0 & -3 \\ 0 & 0 & 0 \end{bmatrix}.$$

A segunda matriz não satisfaz a condição (a) porque o primeiro elemento não-nulo da primeira linha não é 1. A primeira matriz satisfaz a condição (a) mas não satisfaz a condição (b) na coluna 3.

Demonstraremos agora que podemos passar de uma matriz arbitrária a uma matriz linha-reduzida, por meio de um número finito de operações elementares sôbre linhas. Combinado com o Teorema 3, isto nos fornecerá um instrumento eficiente para a resolução de sistemas de equações lineares.

Teorema 4. *Tôda* m $\times$ n *matriz sôbre o corpo* F *é linha-equivalente a uma matriz linha-reduzida.*

Demonstração. Seja A uma $m \times n$ matriz sôbre F. Se todo elemento na primeira linha de A é 0, então a condição (a) está satisfeita no que diz respeito à linha 1. Se a linha 1 tem um elemento não-nulo, seja k o menor inteiro positivo j para o qual $A_{ij} \neq 0$. Multipliquemos a linha 1 por A_{1k}^{-1} e então a condição (a) está satisfeita em relação à linha 1. Agora, para cada $i \geq 2$, somemos $(-A_{ik})$ vêzes a linha 1 à linha i. Agora o primeiro elemento não-nulo da linha 1 ocorre na coluna k, êste elemento é 1, e todos os outros elementos na coluna k são nulos.

Consideremos agora a matriz que resultou das operações acima. Se todo elemento na linha 2 é nulo, nada fazemos à linha 2. Se algum elemento na linha 2 é diferente de 0, multiplicamos a linha 2 por um escalar de modo que o primeiro elemento não-nulo seja 1. No caso em que a linha 1 tenha um primeiro elemento não-nulo na coluna k, êste primeiro elemento não-nulo na linha 2 não pode ocorrer na coluna k; digamos que êle aparece na coluna $k' \neq k$. Somando múltiplos adequados da linha 2 às diversas linhas, podemos fazer com que todos os elementos na coluna k' sejam nulos, com exceção do 1 na linha 2. O fato importante a ser notado é êste: ao efetuarmos estas últimas operações, não alteramos os elementos da linha 1 nas colunas $1, \ldots, k$; além disso, não alteramos nenhum elemento da coluna k. É claro que, se a linha 1 fôsse idênticamente nula, as operações com a linha 2 não afetariam a linha 1.

Trabalhando com uma linha de cada vez da maneira acima, é evidente que, com um número finito de passos, chegaremos a uma matriz linha-reduzida.

Exercícios

1. Determinar tôdas as soluções do sistema de equações

$$\begin{aligned} (1 - i)x_1 - ix_2 &= 0 \\ 2x_1 + (1 - i)x_2 &= 0. \end{aligned}$$

2. Se

$$A = \begin{bmatrix} 3 & -1 & 2 \\ 2 & 1 & 1 \\ 1 & -3 & 0 \end{bmatrix}$$

determinar tôdas as soluções de $AX = 0$, tornando A linha-reduzida.

3. Se

$$A = \begin{bmatrix} 6 & -4 & 0 \\ 4 & -2 & 0 \\ -1 & 0 & 3 \end{bmatrix}$$

determinar tôdas as soluções de $AX = 2X$ e tôdas as soluções de $AX = 3X$. (O símbolo cX indica a matriz cujos elementos são c vêzes os elementos correspondentes de X.)

4. Encontrar uma matriz linha-reduzida que seja linha-equivalente a

$$A = \begin{bmatrix} i & -(1+i) & 0 \\ 1 & -2 & 1 \\ 1 & 2i & -1 \end{bmatrix}.$$

5. Demonstrar que as duas matrizes seguintes *não* são linha-equivalentes:

$$\begin{bmatrix} 2 & 0 & 0 \\ a & -1 & 0 \\ b & c & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 2 \\ -2 & 0 & -1 \\ 1 & 3 & 5 \end{bmatrix}.$$

6. Seja

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

uma 2×2 matriz com elementos complexos. Suponhamos que A seja linha-reduzida e também que $a + b + c + d = 0$. Demonstrar que existem exatamente três destas matrizes.

7. Demonstrar que a transposição de duas linhas de uma matriz pode ser conseguida por uma seqüência finita de operações elementares sôbre linhas dos outros dois tipos.

8. Consideremos o sistema de equações $AX = 0$ onde

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

é uma 2×2 matriz sôbre o corpo F. Demonstrar o que segue.

(i) Se todo elemento de A é nulo, então todo par (x_1, x_2) é uma solução de $AX = 0$.

(ii) Se $ad - bc \neq 0$, o sistema $AX = 0$ possui apenas a solução trivial $x_1 = x_2 = 0$.

(iii) Se $ad - bc = 0$ e algum elemento de A é diferente de 0, então existe uma solução (x_1^0, x_2^0) tal que (x_1, x_2) é uma solução se e sòmente se existe um escalar y tal que $x_1 = yx_1^0$, $x_2 = yx_2^0$.

1.4 Matrizes Linha-reduzidas à Forma em Escada

Até agora, nosso trabalho com sistemas de equações lineares foi motivado por uma tentativa de determinar as soluções de um tal sistema. Na Seção 1.3 estabelecemos um método padronizado para determinar estas soluções. Desejamos agora obter algum conhecimento que seja um pouco mais teórico, e para tal propósito é conveniente ir um pouco além de matrizes linha-reduzidas.

Definição. *Uma* m $\times$ n *matriz* R *é dita uma* **matriz linha-reduzida à forma em escada** *se*

(a) R *é linha-reduzida*;

(b) *tôda linha de* R *cujos elementos são todos nulos ocorre abaixo de tôdas as linhas que possuem um elemento não-nulo*;

(c) *se as linhas* 1, . . . , r *são as linhas não-nulas de* R *e se o primeiro elemento não-nulo da linha* i *ocorre na coluna* k_i, i = 1, . . . , r, *então* $k_1 < k_2 < \ldots < k_r$.

Pode-se também descrever uma $m \times n$ matriz R linha-reduzida à forma em escada como segue. Todo elemento em R é nulo ou então existe um inteiro positivo r, $1 \leq r \leq m$, e r inteiros positivos k_2, . . . , k_r com $1 \leq k_i \leq n$ e

(a) $R_{ij} = 0$ para $i > r$, e $R_{ij} = 0$ se $j < k_i$.

(b) $R_{ik_j} = \delta_{ij}$, $1 \leq i \leq r$, $1 \leq j \leq r$.

(c) $k_1 < \ldots < k_r$.

Exemplo 8. Dois exemplos de matrizes linha-reduzidas à forma em escada são $n \times n$ matriz unidade e a $m \times n$ **matriz nula** $0^{m,n}$, na qual todos os elementos são nulos. O leitor não deverá encontrar nenhuma dificuldade para encontrar outros exemplos, mas gostaríamos de dar mais um exemplo não-trivial:

$$\begin{bmatrix} 0 & 1 & -3 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Teorema 5. *Tôda* m $\times$ n *matriz* A *é linha-equivalente a uma matriz linha-reduzida à forma em escada.*

Demonstração. Sabemos que A é linha-equivalente a uma matriz linha-reduzida. Portanto, basta observar que, efetuando um número finito de permutações das linhas de uma matriz linha-reduzida, podemos transformá-la numa matriz linha-reduzida à forma em escada.

Nos Exemplos 5 e 6, vimos a importância de matrizes linha-reduzidas na solução de sistemas homogêneos de equações lineares. Discutamos ràpidamente o sistema $RX = 0$, no caso em que R é uma matriz linha-reduzida à forma em escada. Sejam as linhas 1, . . . , r as linhas não-nulas de R e suponhamos que o primeiro elemento não-nulo da linha i ocorra na coluna k_i. O sistema $RX = 0$ consiste então de r equações não-triviais. Além disso, a incógnita x_{k_i} aparecerá (com coeficiente não-nulo) apenas na i-ésima equação. Se indicarmos por $u_1, \ldots, u_{n-r}$ as $(n - r)$ incógnitas que são diferentes de $x_{k_1}, \ldots, x_{k_r}$, então as r equações não-triviais em $RX = 0$ são da forma

$$
(1\text{—}3) \qquad
\begin{aligned}
x_{k_1} + \sum_{j=1}^{n-r} C_{1j}u_j &= 0 \\
\vdots \qquad\quad & \ \vdots \\
x_{k_r} + \sum_{j=1}^{n-r} C_{rj}u_j &= 0\,.
\end{aligned}
$$

Tôdas as soluções dos sistemas de equações $RX = 0$ são obtidas atribuindo-se valores arbitrários a $u_1, \ldots, u_{n-r}$ e calculando os valores correspondentes de $x_{k_1}, \ldots, x_{k_r}$ por meio de (1—3). Por exemplo, se R é a matriz do exemplo 8 acima, então $r = 2$, $k_1 = 2$, $k_2 = 4$, e as duas equações não-triviais do sistema $RX = 0$ são

$$
\begin{aligned}
x_2 - 3x_3 \qquad + \tfrac{1}{2}x_5 &= 0 \text{ ou } x_2 = 3x_3 - \tfrac{1}{2}x_5 \\
x_4 + 2x_5 &= 0 \text{ ou } x_4 = 2x_5.
\end{aligned}
$$

Assim, podemos atribuir valores arbitrários a x_1, x_3 e x_5, digamos $x_1 = a$, $x_3 = b$, $x_5 = c$, e obter a solução $(a, 3b - \frac{1}{2}c, b, -2c, c)$.

Observemos mais um fato sôbre o sistema de equações $RX = 0$. Se o número r de linhas não-nulas de R é menor que n, então o sistema $RX = 0$ admite uma solução não-trivial, isto é, uma solução $(x_1, \ldots, x_n)$ em que nem todo x_j é nulo. De fato, como $r < n$, podemos tomar algum x_j que não esteja entre as r incógnitas $x_{k_1}, \ldots, x_{k_r}$ e daí construir uma solução como acima na qual êste x_j é 1. Esta observação nos leva a um dos conceitos mais fundamentais relativos a sistemas de equações lineares homogêneas.

Teorema 6. *Se* A *é uma* m $\times$ n *matriz e* m $<$ n, *então o sistema homogêneo de equações lineares* AX $= 0$ *admite uma solução não-trivial.*

Demonstração. Seja R uma matriz linha-reduzida à forma em escada que seja linha-equivalente a A. Então os sistemas $AX = 0$ e $RX = 0$ possuem, pelo Teorema 3, as mesmas soluções. Se r é o número de linhas não-nulas em R, então certamente $r \leqslant$ m e como $m < n$, temos $r < n$. Decorre imediatamente de nossas observações acima que $AX = 0$ admite uma solução não-trivial.

Teorema 7. *Se* A *é uma* n $\times$ n *matriz* (*quadrada*) *e se o sistema de equações* AX $= 0$ *não possui solução não-trivial, então* A *é linha-equivalente a* n $\times$ n *matriz unidade.*

Demonstração. Seja R uma $n \times n$ matriz linha-reduzida à forma em escada que seja linha-equivalente a A, e seja r o número de elementos não-nulos de R. Como $AX = 0$ não admite solução não-trivial, $RX = 0$ não admite solução não-trivial. Assim, $r \geqslant n$. Mas como R possui n linhas, certamente $r \leqslant n$ e temos $r = n$. Como isto significa que R possui na verdade um primeiro elemento não-nulo igual a 1 em cada uma de suas n linhas e como êstes 1 ocorrem cada um numa das n colunas, R é, necessàriamente, a $n \times n$ matriz unidade.

Perguntemos agora que operações elementares sôbre linhas efetuar para resolver um sistema de equações lineares $AX = Y$ que não seja homogêneo. De início, devemos observar uma diferença básica entre êste caso e o caso homogêneo, a saber, que enquanto o sistema homogêneo sempre admite a solução trivial $x_1 = \ldots = x_n = 0$, um sistema não homogêneo pode não ter nenhuma solução.

Formemos a **matriz completa** A' do sistema $AX = Y$. Esta é a $m \times (n + 1)$ matriz cujas n primeiras colunas são as colunas de A e cuja última coluna é Y. Mais precisamente

$$\begin{aligned} A'_{ij} &= A_{ij}, \text{ se } j \leqslant n \\ A'_{i(n+1)} &= y_i \end{aligned}$$

Suponhamos que efetuemos uma seqüência de operações elementares sôbre as linhas de A, obtendo uma matriz R linha-reduzida à forma em escada. Se efetuarmos esta mesma seqüência de operações sôbre a matriz completa A', obteremos uma matriz R' cujas n primeiras colunas são as colunas de R e cuja coluna contém certos escalares $z_1, \ldots, z_m$. Os escalares z_i são os elementos da $m \times 1$ matriz

$$Z = \begin{bmatrix} z_1 \\ \cdot \\ \cdot \\ \cdot \\ z_m \end{bmatrix}$$

que resulta de se aplicar a seqüência de operações sôbre as linhas da matriz Y. Deve ser evidente ao leitor que, como na demonstração do Teorema 3, os sistemas $AX = Y$ e $RX = Z$ são equivalentes e portanto admitem as mesmas soluções. É bem fácil saber se o sistema $RX = Z$ possui soluções e em caso afirmativo determinar tôdas as soluções. De fato, se R possuir r linhas não-nulas, com o primeiro elemento não-nulo da linha i ocorrendo na coluna k_i, $i = 1, \ldots, r$, então as r primeiras equações de $RX = Z$ exprimem realmente $x_{k_1}, \ldots, x_{k_r}$ em têrmos dos $(n - r)$ x_j restantes e dos escalares $z_1, \ldots, z_r$. As $(m - r)$ últimas equações são

$$\begin{aligned} 0 &= z_{r+1} \\ &\vdots \\ 0 &= z_m \end{aligned}$$

portanto a condição para o sistema ter uma solução é que $z_i = 0$ para $i > r$. Se esta condição é satisfeita, tôdas as soluções dêste sistema podem ser determinadas, como no caso homogêneo, atribuindo-se valores arbitrários a $(n - r)$ dos x_j e daí calculando x_{k_i} por meio da i-ésima equação

Exemplo 9. Seja F o corpo dos números racionais e

$$A = \begin{bmatrix} 1 & -2 & 1 \\ 2 & 1 & 1 \\ 0 & 5 & -1 \end{bmatrix}$$

e suponhamos que se deseje resolver o sistema $AX = Y$ para certos y_1, y_2 e y_3. Efetuemos uma seqüência de operações sôbre as linhas da matriz completa A' que torne A linha-reduzida:

$$\begin{bmatrix} 1 & -2 & 1 & y_1 \\ 2 & 1 & 1 & y_2 \\ 0 & 5 & -1 & y_3 \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 1 & -2 & 1 & y_1 \\ 0 & 5 & -1 & (y_2 - 2y_1) \\ 0 & 5 & -1 & y_3 \end{bmatrix} \xrightarrow{(2)}$$

$$\begin{bmatrix} 1 & -2 & 1 & y_1 \\ 0 & 5 & -1 & (y_2 - 2y_1) \\ 0 & 0 & 0 & (y_3 - y_2 + 2y_1) \end{bmatrix} \xrightarrow{(1)}$$

$$\begin{bmatrix} 1 & -2 & 1 & y_1 \\ 0 & 1 & -\frac{1}{5} & \frac{1}{5}(y_2 - 2y_1) \\ 0 & 0 & 0 & (y_3 - y_2 + 2y_1) \end{bmatrix} \xrightarrow{(2)}$$

$$\begin{bmatrix} 1 & 0 & \frac{3}{5} & \frac{1}{5}(y_1 + 2y_2) \\ 0 & 1 & -\frac{1}{5} & \frac{1}{5}(y_2 - 2y_1) \\ 0 & 0 & 0 & (y_3 - y_2 + 2y_1) \end{bmatrix}.$$

A condição para que o sistema $AX = Y$ tenha uma solução é portanto

$$2y_1 - y_2 + y_3 = 0$$

e se os escalares y_i dados satisfazem esta condição, tôdas as soluções são obtidas atribuindo-se um valor c a x_3 e depois calculando

$$\begin{aligned} x_1 &= -\tfrac{3}{5}c + \tfrac{1}{5}(y_1 + 2y_2) \\ x_2 &= \tfrac{1}{5}c + \tfrac{1}{5}(y_2 - 2y_1). \end{aligned}$$

Façamos uma observação final sôbre o sistema $AX = Y$. Suponhamos que os elementos da matriz A e os escalares $y_1, \ldots, y_m$ estejam num subcorpo F_1 do corpo F. Se o sistema de equações $AX = Y$ admite uma solução com $x_1, \ldots, x_n$ em F, êle admite uma solução com $x_1, \ldots, x_n$ em F_1. De fato, sôbre qualquer um dos dois corpos, a condição para o sistema admitir uma solução é que valham certas relações entre $y_i, \ldots, y_m$ em F_1 (a saber, as relações $z_i = 0$ para $i > r$, acima). Por exemplo, se $AX = Y$ é um sistema de equações lineares no qual os escalares y_k e A_{ij} são números reais e, se existe uma solução na qual $x_1, \ldots, x_n$ são números complexos, então existe uma solução com $x_1, \ldots, x_n$ números reais.

Exercícios

1. Determinar tôdas as soluções do seguinte sistema de equações linha-reduzindo a matriz dos coeficientes:

$$\begin{aligned} \tfrac{1}{3}x_1 + 2x_2 - 6x_3 &= 0 \\ -4x_1 + 5x_3 &= 0 \\ -3x_1 + 6x_2 - 13x_3 &= 0 \\ -\tfrac{7}{3}x_1 + 2x_2 - \tfrac{8}{3}x_3 &= 0 \end{aligned}$$

2. Determinar uma matriz linha-reduzida à forma em escada que seja equivalente a

$$A = \begin{bmatrix} 1 & -i \\ 2 & 2 \\ i & 1 + i \end{bmatrix}.$$

Quais são as soluções de $AX = 0$?

3. Descrever explìcitamente tôdas as 2×2 matrizes linha-reduzidas à forma em escada.

4. Consideremos o sistema de equações

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 1 \\ 2x_1 + 2x_3 &= 1 \\ x_1 - 3x_2 + 4x_3 &= 2. \end{aligned}$$

Êste sistema admite solução? Em caso afirmativo, descrever explìcitamente tôdas as soluções.

5. Dar um exemplo de um sistema de duas equações lineares a duas incógnitas que não admite solução.

6. Seja

$$A = \begin{bmatrix} 3 & -1 & 2 \\ 2 & 1 & 1 \\ 1 & -3 & 0 \end{bmatrix}.$$

Para que ternas (y_1,y_2,y_3) o sistema $AX = Y$ admite solução?

7. Seja

$$A = \begin{bmatrix} 3 & -6 & 2 & -1 \\ -2 & 4 & 1 & 3 \\ 0 & 0 & 1 & 1 \\ 1 & -2 & 1 & 0 \end{bmatrix}.$$

Para que (y_1,y_2,y_3,y_4) o sistema de equações $AX = Y$ admite solução?

8. Suponhamos que R e R', sejam 2×3 matrizes linha-reduzidas à forma em escada e que os sistemas $RX = 0$ e $R'X = 0$ admitam as mesmas soluções. Demonstrar que $R = R'$.

1.5 Multiplicação de Matrizes

É evidente (ou, de qualquer modo, deveria ser) que o processo de formar combinações lineares das linhas de uma matriz é um processo fundamental. Por esta razão é vantajoso introduzir um esquema sistemático para indicar exatamente que operações devem ser efetuadas. Mais especìficamente, suponhamos que B seja uma $n \times p$ matriz sôbre um corpo F com linhas $\beta_1, \ldots, \beta_n$ e que a partir de B construamos uma matriz C com linhas $\gamma_1, \ldots, \gamma_m$ formando certas combinações lineares

$$\gamma_i = A_{i1}\beta_1 + A_{i2}\beta_2 + \ldots + A_{in}\beta_m. \tag{1-4}$$

As linhas de C são determinadas pelos mn escalares A_{ij} que são os elementos de uma $m \times n$ matriz A. Se (1-4) é desenvolvido como

$$(C_{i1} \ldots C_{ip}) = \sum_{r=1}^{n} (A_{ir}B_{r1} \ldots A_{ir}B_{rp})$$

vemos que os elementos de C são dados por

$$C_{ij} = \sum_{r=1}^{n} A_{ir}B_{rj}.$$

Definição. *Seja* A *uma* m $\times$ n *matriz sôbre o corpo* F *e seja* B *uma* n $\times$ p *matriz sôbre* F. *O* **produto** AB *é a* m $\times$ p *matriz* C *cujo elemento* i,j *e*

$$C_{ij} = \sum_{r=1}^{n} A_{ir}B_{rj}.$$

Exemplo 10. Eis alguns produtos de matrizes com elementos racionais.

(a) $$\begin{bmatrix} 5 & -1 & 2 \\ 0 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 5 & -1 & 2 \\ 15 & 4 & 8 \end{bmatrix}$$

Neste caso

$$\gamma_1 = (5\ \ -1\ \ 2) = 1\ .\ (5\ \ -1\ \ 2) + 0\ .\ (15\ \ 4\ \ 8)$$
$$\gamma_2 = (0\ \ \ 7\ \ 2) = -3\ (5\ \ -1\ \ 2) + 1\ .\ (15\ \ 4\ \ 8)$$

(b) $$\begin{bmatrix} 0 & 6 & 1 \\ 9 & 12 & -8 \\ 12 & 62 & -3 \\ 3 & 8 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -2 & 3 \\ 5 & 4 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 6 & 1 \\ 3 & 8 & -2 \end{bmatrix}$$

Neste caso

$$\gamma_2 = (\ 9\ \ 12\ \ -8) = -2(0\ \ 6\ \ 1) + 3(3\ \ 8\ \ -2)$$
$$\gamma_3 = (12\ \ 62\ \ -3) = 5(0\ \ 6\ \ 1) + 4(3\ \ 8\ \ -2)$$

(c) $$\begin{bmatrix} 8 \\ 29 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 5 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 6 \end{bmatrix}$$

(d) $$\begin{bmatrix} -2 & -4 \\ 6 & 12 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \end{bmatrix} [2\ \ 4]$$

Neste caso

$$\gamma_2 = (6\ \ 12) = 3(2\ \ 4)$$

(e) $$[2\ \ 4] \begin{bmatrix} -1 \\ 3 \end{bmatrix} = [10]$$

(f) $$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -5 & 2 \\ 2 & 3 & 4 \\ 9 & -1 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

(g) $$\begin{bmatrix} 1 & -5 & 2 \\ 2 & 3 & 4 \\ 9 & -1 & 3 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 9 & 0 \end{bmatrix}$$

É importante observar que o produto de duas matrizes pode não estar definido; o produto é definido se, e sòmente se, o número de coluna da primeira matriz coincide com o número de linhas da segunda matriz. Assim, não faz sentido trocar a ordem dos fatôres em (a), (b) e (c) acima. Freqüentemente escreveremos produtos como AB sem mencionar explìcitamente as dimensões dos fatôres e, em tais casos, estará subentendido que o produto está definido. De (d), (e), (f), (g) vemos que mesmo quando ambos os produtos AB e BA estão definidos não é necessàriamente verdade que $AB = BA$; em outras palavras a multiplicação de matrizes *não é comutativa*.

Exemplo 11.

(a) Se I é a $m \times m$ matriz unidade e A é uma $m \times n$ matriz, $IA = A$.

(b) Se I é a $n \times n$ matriz unidade e A é uma $m \times n$ matriz, $AI = A$.

(c) Se $0^{k,m}$ é a $k \times m$ matriz nula, $0^{k,n} = 0^{k,m}A$. Anàlogamente, $A0^{n,p} = 0^{m,p}$.

Exemplo 12. Seja A uma $m \times n$ matriz sôbre F. Nossa notação taquigráfica anterior, $AX = Y$, para sistemas de equações lineares, é coerente com nossa definição de produtos de matrizes. De fato, se

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

com x_i em F, então AX é a $m \times 1$ matriz

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_m \end{bmatrix}$$

tal que $y_i = A_{i1}x_1 + A_{i2}x_2 + \ldots + A_{in}x_n$.

A despeito do fato de que um produto de matrizes depende da ordem em que os fatôres são escritos, êle é independente da maneira pela qual elas são associadas, como o próximo teorema mostra.

Teorema 8. *Se* A, B, C *são matrizes sôbre o corpo F tais que os produtos* BC *e* A(BC) *são definidos, então estão definidos os produtos* AB, (AB)C *e*

$$A(BC) = (AB)C.$$

Demonstração. Suponhamos que B seja uma $n \times p$ matriz. Como BC está definida, C é uma matriz com p linhas e BC tem n linhas. Como $A(BC)$ está definido podemos supor que A é uma $m \times n$ matriz. Assim, o produto AB existe e é uma $m \times p$ matriz, do que segue que o produto $(AB)C$ existe. Mostrar que $A(BC) = (AB)C$ significa mostrar que

$$[A(BC)]_{ij} = [(AB)C]_{ij}$$

para cada i, j. Por definição

$$\begin{aligned}
[A(BC)]_{ij} &= \sum_r A_{ir}(BC)_{rj} \\
&= \sum_r A_{ir} \sum_s B_{rs}C_{sj} \\
&= \sum_r \sum_s A_{ir}B_{rs}C_{sj} \\
&= \sum_s \sum_r A_{ir}B_{rs}C_{sj} \\
&= \sum_s (\sum_r A_{ir}B_{rs})C_{sj} \\
&= \sum_s (AB)_{is}C_{sj} \\
&= [(AB)C]_{ij}.
\end{aligned}$$

Quando A é uma $n \times m$ matriz (quadrada), o produto AA está definido. Indicaremos esta matriz por A^2. Pelo Teorema 8, $(AA)A = A(AA)$ ou $A^2A = AA^2$, de modo que o produto AAA está definido sem ambigüidade. Indicaremos êste produto por A^3. Em geral, o produto $AA \ldots A$ (k vêzes) está definido sem ambigüidade e indicaremos êste produto por A^k.

Notemos que a relação $A(BC) = (AB)C$ implica, entre outras coisas, que combinações lineares de combinações lineares das linhas de C são novamente combinações lineares das linhas de C.

Se B é uma dada matriz e C é obtida de B por meio de uma operação elementar sôbre linhas, então cada linha de C é uma combinação linear das linhas de B, logo existe uma matriz A tal que $AB = C$. Em geral, existem muitas dessas matrizes A e, dentre elas tôdas, é conveniente e possível escolher uma que tenha um número de propriedades especiais. Antes de passar a isto precisamos introduzir uma classe de matrizes.

Definição. *Uma* $n \times m$ *matriz é dita uma* **matriz elementar** *se ela pode ser obtida da* $m \times m$ *matriz unidade por meio de uma única operação elementar.*

Exemplo 13. Uma 2×2 matriz elementar é necessàriamente uma das seguintes:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix}$$

$$\begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}, \ c \neq 0, \quad \begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix}, \ c \neq 0.$$

Lema. *Seja* e *uma operação elementar sôbre linhas de matrizes com* p *linhas. Seja* A *uma* m $\times$ n *matriz e* B *uma* p $\times$ m *matriz. Então*

$$e(B)A = e(BA). \tag{1—5}$$

Demonstração. Indiquemos as linhas de A por $\alpha_1, \ldots, \alpha_m$. As linhas $\gamma_1, \ldots, \gamma_p$ de $C = BA$ são então dadas por

$$\gamma_i = \sum_j B_{ij}\alpha_j. \tag{1—6}$$

Se a operação e é a multiplicação de r-ésima linha por $c \neq 0$, então a r-ésima linha de $e(C)$ é dada por

$$\gamma'_r = \sum_j cB_{rj}\alpha_j \tag{1—7}$$

enquanto $\gamma'_i = \gamma_i$ para $i \neq r$. Por outro lado, se e é uma operação que substitui a linha r pela linha r mais c vêzes a linha s, $r \neq s$, então

$$\gamma'_r = \sum_j (B_{rj} + cB_{sj})\alpha_j \tag{1—8}$$

e $\gamma'_i = \gamma_i$ para $i \neq r$. No caso restante, quando e transpõe as linhas r e s, temos

$$\begin{aligned} \gamma'_r &= \sum_j B_{sj}\alpha_j \\ \gamma'_s &= \sum_j B_{rj}\alpha_j \end{aligned} \tag{1—9}$$

e $\gamma'_i = \gamma_i$ se i é diferente de r e de s. Considerando (1—7), (1—8) e (1—9) é evidente que em cada caso

$$\gamma'_i = \sum_j e(B)_{ij}\alpha_j$$

para $i = 1, \ldots, p$.

Tomando B como sendo a $m \times m$ matriz unidade em (1—5) obtemos

$$e(I)A = e(A). \tag{1—10}$$

Por ser êste resultado de importância fundamental reenunciamo-lo como segue:

Teorema 9. *Seja* A *uma* m × n *matriz sôbre o corpo* F *e seja* C *uma matriz obtida efetuando-se uma única operação elementar sôbre as linhas de* A. *Seja* E *a matriz elementar obtida efetuando-se a mesma operação elementar sôbre a* m × m *matriz unidade. Então* C = EA.

Corolário. *Sejam* A *e* B m × n *matrizes sôbre o corpo* F. *Então* B *é linha-equivalente a* A *se e sòmente se* B = PA, *onde* P *é um produto de* m × m *matrizes elementares.*

Demonstração. Suponhamos que $B = PA$ onde $P = E_s \ldots E_2E_1$ e as E_i são $m \times m$ matrizes elementares. Então E_1A é linha-equivalente a A e $E_2(E_1A)$ é linha-equivalente a E_2A. Assim E_2E_1A é linha-equivalente a A; continuando desta maneira, vemos que $(E_s \ldots E_1)A$ é linha-equivalente a A.

Suponhamos agora que B seja linha-equivalente a A. Sejam $E_1, E_2, \ldots, E_s$ as matrizes elementares correspondentes a alguma seqüência de operações elementares sôbre linhas que levam A em B. Então $B = (E_s \ldots E_1)A$.

Exercícios

1. Sejam

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 2 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix} \qquad C = [1 \quad -1].$$

Calcular ABC e CAB.

2. Sejam

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \\ 3 & 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & -2 \\ 1 & 3 \\ 4 & 4 \end{bmatrix}.$$

Verificar diretamente que $A(AB) = A^2B$.

3. Determinar duas 2×2 matrizes A distintas tais que $A^2 = 0$ mas $A \neq 0$.

4. Para a matriz A do Exercício 2, determinar matrizes elementares $E_1, E_2, \ldots, E_k$ tais que

$$E_k \ldots E_2E_1A = I.$$

5. Sejam

$$A = \begin{bmatrix} 1 & -1 \\ 2 & 2 \\ 1 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 3 & 1 \\ -1 & 4 \end{bmatrix}.$$

Existe alguma matriz C tal que $CA = B$?

6. Seja A uma $m \times n$ matriz e B uma $n \times k$ matriz. Mostrar que as colunas de $C = AB$ são combinações lineares das colunas de A. Se $\alpha_1, \ldots, \alpha_n$ são as colunas de A e $\gamma_1, \ldots, \gamma_k$ são as colunas de C, então

$$\gamma_j = \sum_{r=1}^{n} B_{rj}\alpha_r.$$

7. Sejam A e B duas 2×2 matrizes tais que $AB = I$. Demonstrar que $BA = I$.

8. Seja

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

uma 2×2 matriz. Perguntamos quando é possível encontrar 2×2 matrizes A e B tais que $C = AB - BA$. Demonstrar que tais matrizes podem ser encontradas se e sòmente se $C_{11} + C_{22} = 0$.

1.6 Matrizes Inversíveis

Suponhamos que P seja uma $m \times m$ matriz que seja um produto de matrizes elementares. Para cada $m \times n$ matriz A, a matriz $B = PA$ é linha-equivalente a A; logo A é linha-equivalente a B e existe um produto Q de matrizes elementares tal que $A = QB$. Em particular, isto é válido quando A é a $m \times m$ matriz unidade. Em outras palavras, existe uma $m \times m$ matriz Q, que é um produto de matrizes elementares, tal que $QP = I$. Como logo veremos, a existência de uma Q tal que $QP = I$ é equivalente ao fato de P ser um produto de matrizes elementares.

Definição. *Seja* A *uma* n $\times$ n *matriz (quadrada) sôbre o corpo* F. *Uma* n $\times$ n *matriz* B *tal que* BA = I *é dita uma* **inversa à esquerda** *de* A; *uma* n $\times$ n *matriz* B *tal que* AB = I *é dita uma* **inversa à direita** *de* A. *Se* AB = BA = I, *então* B *é dita* **inversível.**

Lema. Se A possui uma inversa à esquerda B e uma inversa à direita C, então $B = C$.

Demonstração. Suponhamos que $BA = I$ e $AC = I$. Então

$$B = BI = B(AC) = (BA)C = IC = C.$$

Assim, se A possui uma inversa à esquerda e uma à direita, A é inversível e possui uma única inversa bilateral, que indicaremos por A^{-1} e denominaremos simplesmente a *inversa* de A.

Teorema 10. *Sejam* A *e* B n $\times$ n *matrizes sôbre* F.

(a) *Se* A *é inversível,* A^{-1} *tambem o é e* $(A^{-1})^{-1} = A$.

(b) *Se* A *e* B *são inversíveis,* AB *também o é e* $(AB)^{-1} = B^{-1}A^{-1}$.

Demonstração. A primeira afirmação é evidente pela simetria da definição. A segunda decorre da verificação das relações

$$(AB)(B^{-1}A^{-1}) = (B^{-1}A^{-1})(AB) = I.$$

Corolário. *Um produto de matrizes inversíveis é inversível.*

Teorema 11. *Uma matriz elementar é inversível.*

Demonstração. Seja E uma matriz elementar correspondente à operação elementar sôbre linhas e. Se e_1 é a operação inversa de e (Teorema 2) e $E_1 = e_1(I)$, então

$$EE_1 = e(E_1) = e(e_1(I)) = I$$

$$E_1E = e_1(E) = e_1(e(I)) = I$$

de modo que E é inversível e $E_1 = E^{-1}$.

Exemplo 14.

(a) $$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

(b) $$\begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -c \\ 0 & 1 \end{bmatrix}$$

(c) $$\begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ -c & 1 \end{bmatrix}$$

(d) Quando $c \neq 0$,

$$\begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} c^{-1} & 0 \\ 0 & 1 \end{bmatrix} \text{ e } \begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & c^{-1} \end{bmatrix}.$$

Teorema 12. *Se* A *e uma* n $\times$ n *matriz sôbre* F, *as seguintes afirmações são equivalentes (isto é, tôdas verdadeiras ou tôdas falsas).*

(i) A *e inversível.*

(ii) A *possui uma inversa à esquerda.*

(iii) AX = 0 *é um sistema de equações sem solução além da trivial.*

(iv) A *é um produto de matrizes elementares.*

Demonstração. Demonstraremos as implicações (i) $\rightarrow$ (ii) $\rightarrow$ (iii) $\rightarrow$ (iv) $\rightarrow$ (i). (i) $\rightarrow$ (ii). Se A é inversível, A^{-1} é uma inversa à esquerda de A.

(ii) $\rightarrow$ (iii). Suponhamos que P seja uma inversa à esquerda de A e que $AX = 0$. Então $X = IX = (PA)X = P(AX) = P \cdot 0 = 0$.

(iii) $\rightarrow$ (iv). Suponhamos que o sistema de equaçoes lineares homogêneas $AX = 0$ não possua solução não-trivial. Seja R uma matriz-reduzida à forma em escada e linha-equivalente a A. Então R é uma $n \times n$ matriz quadrada e $RX = 0$ não possui soluções $X \neq 0$. Assim, R é a $n \times n$ matriz unidade e, pelo Corolário do Torema 9, $A = P$ onde P é um produto de matrizes elementares.

(iv) $\rightarrow$ (i). Suponhamos que $E_1, E_2, \ldots, E_S$ sejam $n \times n$ matrizes elementares tais que $A = E_1E_2 \ldots E_S$. Pelo Teorema 11, cada E_i é inversível e é evidente que

$$A^{-1} = E_s^{-1} \ldots E_2^{-1}E_1^{-1}.$$

Corolário 1. *Se* A *e uma* n $\times$ n *matriz inversível e se uma seqüência de operações elementares sôbre linhas reduz* A *a unidade, então aquela mesma seqüência de operações sôbre linhas quando aplicada a* I *produz* A^{-1}.

Corolário 2. *Uma matriz quadrada com inversa à esquerda ou à direita é inversível.*

Demonstração. Se A e B são $n \times n$ matrizes tais que $AB = I$, então A é uma inversa à esquerda de B, logo B é inversível, o que implica $B^{-1} = A$ e $A^{-1} = (B^{-1})^{-1} = B$.

Corolário 3. *Uma* n $\times$ n *matriz* A *e inversível se e somente se o sistema de equações* AX = X *possui uma solução* X *para cada* n $\times$ 1 *matriz* Y.

Demonstração. Suponhamos que A seja inversível. Então $X = A^{-1}Y$ é uma solução da equação $AX = Y$.

Suponhamos que $AX = Y$ possua uma solução para cada Y. Seja Y_i a i-ésima coluna da $n \times n$ matriz unidade. Tomemos X_i de modo que $AX_i = Y_i$. Se B é a $n \times n$ matriz com colunas $X_1, X_2, \ldots, X_n$ então $AB = I$; agora o corolário anterior se aplica e mostra que $B = A^{-1}$.

Deve-se observar que o Corolário 3 mostra que se A é $n \times n$ e $AX = Y$ possui uma solução para cada Y, então na verdade $AX = Y$ possui uma *única* solução para cada Y.

Corolário 4. *Sejam* A *e* B m $\times$ n *matrizes. Então* B *é linha-equivalente a* A *se e sòmente se* B = PA *onde* P *é uma* m $\times$ m *matriz inversível.*

Tomando $m = n$ e fazendo B igual à $n \times n$ matriz unidade obtemos o resultado que segue.

Corolário 5. *Uma* n $\times$ n *matriz* A *é linha-equivalente à matriz unidade se e sòmente se* A *é inversível.*

Corolário 6. *Seja* $A = A_1A_2, \ldots A_k$, *onde* $A_1, \ldots, A_k$ *são* $n \times n$ *matrizes (quadradas). Então* A *é inversível se e sòmente se cada* A_j *é inversível.*

Demonstração. Já demonstramos que o produto de duas matrizes inversíveis é inversível. A partir disto vê-se fàcilmente que se cada A_j é inversível então A é inversível.

Suponhamos agora que A seja inversível. Demonstremos primeiro que A_k é inversível. Suponhamos que X seja uma $n \times 1$ matriz e $A_kX = 0$. Então $AX = (A_1 \ldots A_{k-1})A_kX = 0$. Como A é inversível temos $X = 0$. Desta maneira, o sistema de equações $A_kX = 0$ não possui soluções não-triviais, portanto, A_k é inversível. Mas então $A_1 \ldots A_{k-1} = AA_k^{-1}$ é inversível. Pela razão anterior A_{k-1} é inversível. Prosseguindo desta forma, concluímos que cada A_j é inversível.

Gostaríamos de fazer um comentário final sôbre a resolução de equações lineares. Suponhamos que A seja uma $m \times n$ matriz e que desejamos resolver o sistema de equações $AX = Y$. Se R é uma matriz linha-reduzida à forma em escada que é equivalente a A, então $R = PA$, onde P é uma $m \times m$ matriz inversível. As soluções do sistema $AX = Y$ são exatamente as soluções do sistema $RX = PY$ $(=Z)$. Na prática, não é muito mais difícil determinar a matriz P do que linha-reduzir A a R. De fato, suponhamos que formemos a matriz completa A' do sistema $AX = Y$, com escalares arbitrários $y_1, \ldots, y_m$ na última coluna. Se agora efetuarmos sôbre A' uma seqüência de operações elementares sôbre linhas que reduza A a R, tornar-se-á evidente o que é a matriz P. (O leitor deve consultar o Exemplo 9 onde, em essência, aplicamos êste processo.) Em particular, se A é uma matriz quadrada, êste processo mostrará claramente se A é inversível ou não e, se A fôr inversível, qual é a inversa de P. Como já apresentamos o núcleo de um exemplo dêste tipo de cálculo, contentar-nos-emos com um exemplo 2×2.

Exemplo 15. Suponhamos que F seja o corpo dos números racionais e

$$A = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}.$$

Então

$$\begin{bmatrix} 2 & -1 & y_1 \\ 1 & 3 & y_2 \end{bmatrix} \xrightarrow{(3)} \begin{bmatrix} 1 & 3 & y_2 \\ 2 & -1 & y_1 \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 1 & 3 & y_2 \\ 0 & -7 & y_1 - 2y_2 \end{bmatrix} \xrightarrow{(1)}$$

$$\begin{bmatrix} 1 & 3 & y_2 \\ 0 & 1 & \frac{1}{7}(2y_2 - y_1) \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 1 & 0 & \frac{1}{7}(y_2 + 3y_1) \\ 0 & 1 & \frac{1}{7}(2y_2 - y_1) \end{bmatrix}$$

onde se vê claramente que A é inversível e que

$$A^{-1} = \begin{bmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{bmatrix}.$$

Deve ter ocorrido ao leitor que fizemos uma longa discussão sôbre linhas de matrizes e pouco dissemos sôbre colunas. Concentramos nossa atenção sôbre as linhas porque isto pareceu mais natural do ponto de vista de equações lineares. Como não existe evidentemente nada sagrado sôbre linhas, a discussão das últimas seções poderia muito bem ter sido feita usando-se colunas em vez de linhas. Se se define uma operação elementar sôbre colunas e uma coluna equivalência de maneira análoga à operação elementar sôbre linhas e à linha-equivalência é evidente que cada $m \times n$ matriz será coluna equivalente a uma matriz "coluna-reduzida à forma em escada". Além disso, cada operação elementar sôbre colunas será da forma $A \to AE$, onde E é uma $n \times n$ matriz elementar e assim por diante.

Exercícios

1. Seja

$$A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ -1 & 0 & 3 & 5 \\ 1 & -2 & 1 & 1 \end{bmatrix}.$$

Determinar uma matriz R linha-reduzida à forma em escada que seja linha-equivalente a A e uma 3×3 matriz inversível P tal que $R = PA$.

2. Fazer o Exercício 1, com

$$A = \begin{bmatrix} 2 & 0 & i \\ 1 & -3 & -i \\ i & 1 & 1 \end{bmatrix}.$$

3. Para cada uma das matrizes

$$\begin{bmatrix} 2 & 5 & -1 \\ 4 & -1 & 2 \\ 6 & 4 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & -1 & 2 \\ 3 & 2 & 4 \\ 0 & 1 & -2 \end{bmatrix}$$

usar operações elementares sôbre linhas para descobrir se é inversível e, em caso afirmativo, determinar a inversa.

4. Seja

$$A = \begin{bmatrix} 5 & 0 & 0 \\ 1 & 5 & 0 \\ 0 & 1 & 5 \end{bmatrix}.$$

Para que X existe um escalar c tal que $AX = cX$?

5. Suponhamos que A seja uma 2×1 matriz e que B seja uma 1×2 matriz. Demonstrar que $C = AB$ não é inversível.

6. Seja A uma $n \times n$ matriz (quadrada). Demonstrar as duas afirmações seguintes:

(a) Se A é inversível e $AB = 0$ para alguma $n \times n$ matriz, então $B = 0$.

(b) Se A não é inversível, então existe uma $n \times n$ matriz B tal que $AB = 0$ mas $B = 0$.

7. Seja

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Demonstrar, usando operações elementares sôbre linhas, que A é inversível se, e sòmente se, $(ad - bc) \neq 0$.

8. Demonstar a seguinte generalização do Exercício 5. Se A é uma $m \times n$ matriz, B é uma $n \times m$ matriz e $n < m$, então AB não é inversível.

9. Seja A uma $m \times n$ matriz. Mostrar que, por meio de um número finito de operações elementares sôbre linhas e/ou colunas, pode-se passar de A a uma matriz R, "linha-reduzida à forma em escada" e "coluna-reduzida à forma em escada", isto é, $R_{ij} = 0$ se $i \neq j$, $R_{ii} = 1$, $1 \leq i \leq r$, $R_{ii} = 0$ se $i > r$. Mostrar que $R = PAQ$, onde P é uma $n \times m$ matriz inversível e Q é uma $n \times n$ matriz inversível.

CAPÍTULO 2

ESPAÇOS VETORIAIS

2.1 Espaços Vetoriais

Em várias partes da matemática, defrontamo-nos com um conjunto, tal que é, ao mesmo tempo, significativo e interessante lidar com "combinações lineares" dos objetos daquele conjunto. Por exemplo, em nosso estudo de equações lineares, foi bastante natural considerar combinações lineares das linhas de uma matriz. É provável que o leitor tenha estudado cálculo e tenha já lidado com combinações lineares de funções; isto certamente ocorreu se êle estudou equações diferenciais. Talvez o leitor tenha tido alguma experiência com vetores no espaço euclidiano tridimensional e, em particular, com combinações lineares de tais vetores.

A grosso modo, a álgebra linear é o ramo da matemática que trata das propriedades comuns a sistemas algébricos constituídos por um conjunto mais uma noção razoável de uma "combinação linear" de elementos do conjunto. Nesta seção definiremos o objeto matemático que, como a experiência mostrou, é a abstração mais útil dêste tipo de sistema algébrico.

Definição. *Um* **espaço vetorial** *(ou espaço linear) consiste do seguinte:*

(1) *um corpo* F *de escalares;*

(2) *um corpo* V *de objetos, denominados vetores;*

(3) *uma regra (ou operação), dita adição de vetores, que associa a cada par de vetores* α, β *em* V *um vetor* $\alpha + \beta$ *em* V, *denominado a soma de* α *e* β, *de maneira tal que*

(a) *a adição é comunicativa,* $\alpha + \beta = \beta + \alpha$;

(b) *a adição é associativa,* $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$;

(c) *existe um único vetor* 0 *em* V, *denominado o vetor nulo, tal que* $\alpha + 0 = \alpha$ *para todo* α *em* V;

(d) *para cada vetor* α *em* V *existe um único vetor* $-\alpha$ *em* V *tal que* $\alpha + (-\alpha) = 0$;

(4) *uma regra (ou operação), dita multiplicação escalar, que associa a cada escalar* c *em* F *e cada vetor* α *em* V *um vetor* $c\alpha$ *em* V, *denominado o produto de* c *por* α *de maneira tal que*

(a) $1\,\alpha = \alpha$ *para todo* α *em* V;

(b) $(c_1c_2)\alpha = c_1(c_2\alpha)$;

(c) $c(\alpha + \beta) = c\alpha + c\beta$;

(d) $(c_1 + c_2)\alpha = c_1\alpha + c_2\alpha$.

É importante observar, como afirma a definição, que um espaço vetorial é um objeto composto de um corpo, um conjunto de "vetores" e duas operações com certas propriedades especiais. O mesmo conjunto de vetores pode ser parte de diversos espaços vetoriais (ver Exemplo 5 abaixo). Quando não há possibilidade de confusão, podemos simplesmente nos referir ao espaço vetorial por V ou, quando fôr desejável especificar o corpo, dizer que V é um **espaço vetorial sôbre o corpo** F. O nome "vetor" é aplicado aos elementos do conjunto V mais por conveniência. A origem do nome é encontrada no Exemplo 1 abaixo, mas não se deve emprestar muita importância ao nome uma vez que a variedade de objetos que aparecem como sendo os vetores em V podem não apresentar muita semelhança com qualquer conceito de vetor adquirido *a priori* pelo leitor. Tentaremos indicar esta variedade através de uma lista de exemplos; nossa lista será consideràvelmente ampliada assim que iniciarmos o estudo de espaços vetoriais.

Exemplo 1. O espaço das n**-uplas,** F^n. Seja F um corpo arbitrário e seja V o conjunto de tôdas as n-uplas $\alpha = (x_1, x_2, \ldots, x_n)$ de escalares x_i em F. Se $\beta = (y_1, y_2, \ldots, y_n)$ com y_i em F, a soma de α e β é definida por

$$\alpha + \beta = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n). \tag{2-1}$$

O produto de um escalar c por um vetor α é definido por

$$c\alpha = (cx_1, cx_2 \ldots, cx_n). \tag{2-2}$$

O fato de que esta adição de vetores e multiplicação escalar satisfazem as condições (3) e (4) é fácil de verificar, usando as propriedades semelhantes da adição e multiplicação de elementos de F.

Exemplo 2. O espaço das $m \times n$ **matrizes sôbre o corpo** F. Seja F um corpo arbitrário e sejam m e n inteiros positivos. Seja V o conjunto de tôdas as $m \times n$ matrizes sôbre o corpo F. A soma de dois vetores A e B em V é definida por

$$(2\text{-}3) \qquad (A + B)_{ij} = A_{ij} + B_{ij}.$$

O produto de um escalar c por A em V é definido por

$$(2\text{-}4) \qquad (cA)_{ij} = cA_{ij}.$$

Exemplo 3. O espaço das funções de um conjunto em um corpo. Seja F um corpo arbitrário e seja S um conjunto não-vazio arbitrário. Seja V o conjunto das funções do conjunto S em F. A soma de dois vetores f e g em V é o vetor $f + g$, isto é, a função de S em F, definida por

$$(2\text{-}5) \qquad (f + g)(s) = f(s) + g(s).$$

O produto do escalar c pela função f é a função cf definida por

$$(2\text{-}6) \qquad (cf)(s) = cf(s).$$

Os exemplos anteriores são os casos particulares dêste. De fato, uma n-upla de elementos de F pode ser considerada como uma função do conjunto S dos inteiros $1, \ldots, n$ em F. Anàlogamente, uma $m \times n$ matriz sôbre o corpo F é uma função do conjunto S de pares de inteiros (i,j), $1 \leq i \leq m$, $1 \leq j \leq$ n, no corpo F. Para êste terceiro exemplo indicaremos como se faz para verificar que as operações por nós definidas satisfazem as condições (3) e (4). Para a adição de vetores:

(a) Como a adição em F é comutativa,

$$f(s) + g(s) = g(s) + f(s)$$

para cada s em S, portanto as funções $f + g$ e $g +$ f são idênticas.

(b) Como a adição em F é associativa,

$$f(s) + [g(s) + h(s)] = f(s) + g(s) + h(s)$$

para cada s, portanto $f + (g + h)$ e $(f + g) + h$ são a mesma função.

(c) O único vetor nulo é a função nula que associa a cada elemento de S o escalar 0 em F.

(d) Para cada f em V, $(-f)$ é a função dada por

$$(-f)(s) = -f(s).$$

O leitor deverá achar fácil verificar que a multiplicação escalar satisfaz as condições de (4), fazendo como fizemos para a adição de vetores.

Exemplo 4. O espaço das funções polinomiais sôbre um corpo F. Seja F um corpo e seja V o conjunto das funções f de F em F que são da forma

$$f(x) = c_0 + c_1x + \ldots + c_nx^n \tag{2-7}$$

onde $c_0, c_1, \ldots, c_n$ são escalares fixos em F (independentes de x). Uma função dêste tipo é denominada uma **função polinomial sôbre** F. Sejam a adição e multiplicação escalar definidas como no Exemplo 3. Deve-se observar aqui que se f e g são funções polinomiais e c está em F, então $f + g$ e cf são também funções polinomiais.

Exemplo 5. O corpo C dos números complexos pode ser considerado como um espaço vetorial sôbre o corpo R dos números reais. De maneira mais geral, seja F o corpo dos números reais, e seja V o conjunto das n-uplas $\alpha = (x_1, \ldots, x_n)$ onde $x_1, \ldots, x_n$ são números *complexos*. Definamos a adição de vetores e a multiplicação escalar por (2—1) e (2—2), como no Exemplo 1. Desta forma obtemos um espaço vetorial sôbre o corpo R que é bem diferente do espaço C^n e do espaço R^n.

Há alguns fatos simples que decorrem quase imediatamente da definição de um espaço vetorial e que passamos a deduzir. Se c é um escalar e 0 é o vetor nulo, então, por (3) (c) e (4) (c),

$$c0 = c(0 + 0) = c0 + c0.$$

Somando $-(c0)$ e usando 3(d) obtemos

$$c0 = 0. \tag{2-8}$$

Anàlogamente, para o escalar 0 e qualquer vetor α temos que

$$0\alpha = 0. \tag{2-9}$$

Se c é um escalar não-nulo e α é um vetor tal que $c\,\alpha = 0$, então por (2-8), $c^{-1}(c\alpha) = 0$. Mas

$$c^{-1}(c\alpha) = (c^{-1}c)\alpha = 1\,\alpha = \alpha$$

logo, $\alpha = 0$. Assim, vemos que se c é um escalar e α um vetor tal que $c\alpha = 0$, então c é o escalar nulo ou α é o vetor nulo.

Se α é um vetor arbitrário em V, então

$$0 = 0\alpha = (1 - 1)\alpha = 1\,\alpha + (-1)\alpha = \alpha + (-1)\alpha$$

do que segue que

$$(-1)\alpha = -\alpha. \tag{2-10}$$

Finalmente, as propriedades associativa e comutativa da adição de vetores implicam que uma soma envolvendo um certo número de vetores é independente da maneira pela qual êstes vetores são combinados ou associados. Por exemplo, se α_1, α_2, α_3, α_4 são vetores em V, então

$$(\alpha_1 + \alpha_2) + (\alpha_3 + \alpha_4) = [\alpha_2 + (\alpha_1 + \alpha_3)] + \alpha_4$$

e esta pode ser escrita sem confusão como

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4.$$

Definição. *Um vetor β em* V *é dito uma* **combinação linear** *dos vetores $\alpha_1, \ldots, \alpha_n$ em* V *se existem escalares $c_1, \ldots, c_n$ em* F *tais que*

$$\begin{aligned}\beta &= c_1\alpha_1 + \ldots + c_n\alpha_n \\ &= \sum_{i=1}^{n} c_i\alpha_i.\end{aligned}$$

Outras extensões da propriedade associativa da adição de vetores e das propriedades distribuitivas (4) (c) e (4) (d) da multiplicação escalar aplicam-se a combinações lineares:

$$\sum_{i=1}^{n} c_i\alpha_i + \sum_{i=1}^{n} d_i\alpha_i = \sum_{i=1}^{n} (c_i + d_i)\alpha_i$$

$$c \sum_{i=1}^{n} c_i\alpha_i = \sum_{i=1}^{n} (cc_i)\alpha_i.$$

Certas partes da álgebra linear são intimamente relacionadas com a geometria. A própria palavra "espaço" sugere algo geométrico, como o faz a palavra "vetor" à maioria das pessoas. À medida que prossigamos nosso estudo de espaços vetoriais, o leitor observará que grande parte da terminologia possui uma conotação geométrica. Antes de concluirmos esta seção introdutória sôbre espaços vetoriais, seria bom discutirmos a relação dos espaços vetoriais com a geometria até um ponto que indique pelo menos a origem do nome "espaço vetorial". Esta será uma discussão breve e intuitiva.

Consideremos o espaço vetorial R^3. Na geometria analítica, identificamos as ternas (x_1, x_2, x_3) de números reais com os pontos do espaço euclidiano tridimensional. Naquele contexto, um vetor é usualmente definido como sendo um segmento de reta orientado PQ, que vai de um ponto P do espaço a outro ponto Q. Isto significa uma formulação cuidadosa da idéia da "flecha" de P a Q. Da forma como os vetores são usados, pretende-se que êles sejam determinados por seu comprimento, direção e sentido. Assim, é necessário identificar dois segmentos de reta orientados se êles têm o mesmo comprimento, direção e sentido.

O segmento de reta orientado PQ, que vai do ponto $P = (x_1, x_2, x_3)$ ao ponto $Q = (y_1, y_2, y_3)$, tem o mesmo comprimento, direção e sentido que o segmento de reta orientado que vai da origem $O = (0, 0, 0)$ ao ponto $(y_1 - x_1, y_2 - x_2, y_3 - x_3)$. Além disso, êste é o único segmento que emana da origem e tem o mesmo comprimento, direção e sentido que PQ. Assim, se resolvermos estudar apenas os vetores que emanam da origem, existe exatamente um vetor associado a cada comprimento, direção e sentido dados.

O vetor OP, que vai da origem a $P = (x_1, x_2, x_3)$, é completamente determinado por P, portanto é possível identificar êste vetor com o ponto P. Em nossa definição do espaço vetorial R^3, os vetores são definidos como sendo simplesmente as ternas (x_1, x_2, x_3).

Dados pontos $P = (x_1, x_2, x_3)$ e $Q = (y_1, y_2, y_3)$, a definição da soma dos vetores OP e OQ pode ser dada geomètricamente. Se os vetores não são paralelos, então os segmentos OP e OQ determinam um plano e êstes segmentos são dois dos lados de um paralelogramo naquele plano (ver Figura 1). Uma diagonal dêste para-

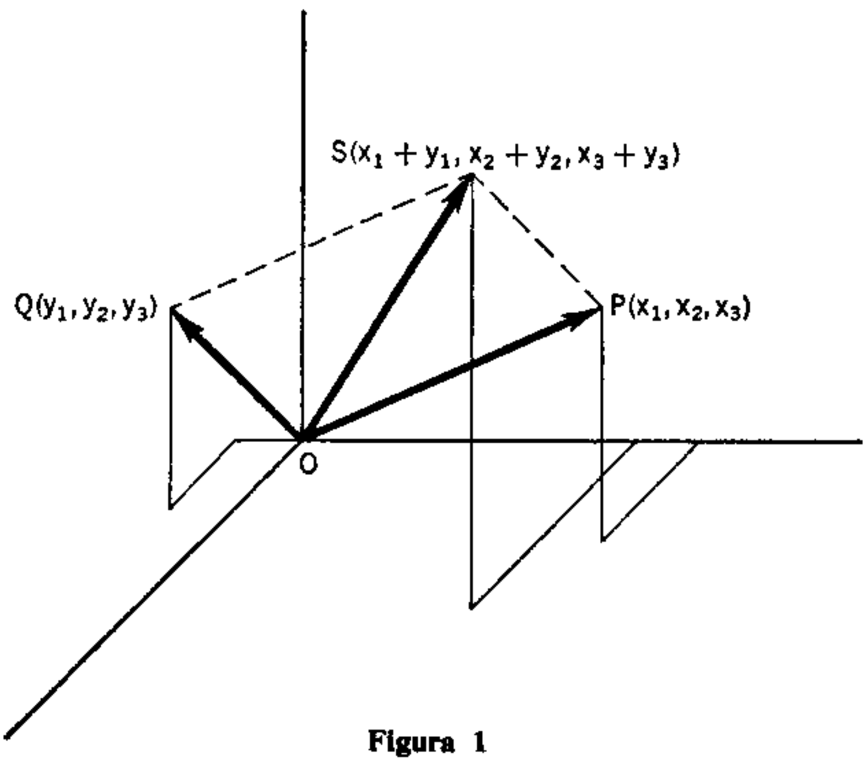

Figura 1

lelogramo estende-se de O a um ponto S e a soma de OP e OQ é definida como sendo o vetor OS. As coordenadas do ponto S são $(x_1 + y_1, x_2 + y_2, x_3 + y_3)$, logo esta definição geométrica da adição de vetores é equivalente à definição algébrica do Exemplo 1.

A multiplicação escalar tem uma interpretação geométrica mais simples. Se c é um número real, então, o produto de c pelo vetor OP é o vetor que parte da origem, tem comprimento $|c|$ vêzes o comprimento de OP, mesma direção que OP e um sentido que concorda com o de OP se $c > 0$ e é oposto ao de OP se $c < 0$. Esta multiplicação escalar produz exatamente o vetor OT onde $T = (cx_1, cx_2, cx_3)$ e é portanto compatível com a definição algébrica dada para R^3.

De vez em quando, o leitor provàvelmente achará útil "pensar geomètricamente" sôbre espaços vetoriais, isto é, desenhar figuras para uso próprio para ilustrar e motivar algumas idéias. Na verdade deve fazer isto. Contudo, ao fazer tais ilustrações, deve ter em mente que, por estarmos tratando de espaços vetoriais como sistemas algébricos, tôdas as demonstrações que fizermos serão de natureza algébrica.

Exercícios

1. Se F é um corpo, verificar que F^n (tal como definido no Exemplo 1) é um espaço vetorial sôbre o corpo F.

2. Se V é um espaço vetorial sôbre o corpo F, verificar que

$$(\alpha_1 + \alpha_2) + (\alpha_3 + \alpha_4) = [\alpha_2 + (\alpha_3 + \alpha_1)] + \alpha_4$$

para todos vetores α_1, α_2, α_3 e α_4 em V.

3. Se C é o corpo dos números complexos, quais vetores em C^3 são combinações lineares de $(1, 0, -1)$, $(0, 1, 1)$ e $(1, 1, 1)$?

4. Seja V o conjunto de todos os pares (x, y) de números reais e seja F o corpo dos números reais. Definamos

$$(x, y) + (x_1, y_1) = (x + x_1, y + y_1)$$
$$c(x, y) = (cx, y).$$

V, com estas operações, é um espaço vetorial sôbre o corpo dos números reais?

5. Seja V o conjunto de todos os pares (x, y) de números reais e seja F o corpo dos números reais. Definamos

$$(x, y) + (x_1, y_1) = (3y + 3y_1, -x - x_1)$$

$$c(x, y) = (3cy, -cx).$$

Verificar que V, com estas operações, não é um espaço vetorial sôbre o corpo dos números reais.

2.2 Subespaços

Nesta seção introduziremos alguns conceitos básicos no estudo dos espaços vetoriais.

Definição. *Seja* V *um espaço vetorial sôbre o corpo* F. *Um* **subespaço** *de* V *é um subconjunto* W *de* V *que é um espaço vetorial sôbre* F *com as operações de adição de vetores e multiplicação escalar de* V.

Uma verificação direta dos axiomas para um espaço vetorial mostra que o subconjunto W de V é um subespaço se para todos α e β em W o vetor $\alpha + \beta$ está ainda em W; o vetor nulo está em W; para todo α em W o vetor $(-\alpha)$ está em W; para todo α em W e todo escalar c o vetor $c\,\alpha$ está em W. A comutatividade e associatividade da adição de vetores e as propriedades (4) (*a*), (*b*), (*c*) e (*d*) da multiplicação escalar não precisam ser verificadas, uma vez que são propriedades das operações em V. Podemos simplificar ainda mais as coisas.

Teorema 1. *Um subconjunto não-vazio* W *de* V *é um subespaço de* V *se, e sòmente se, para cada par de vetores* α, β *em* W *e cada escalar* c *em* F, *o vetor* $c\alpha + \beta$ *está em* W.

Demonstração. Suponhamos que W seja um subconjunto não-vazio de V tal que $c\,\alpha + \beta$ pertença a W para todos os vetores α, β em W e todos escalares c em F. Como W é não-vazio, existe um vetor ρ em W, logo $(-1)\,\rho + \rho = 0$ está em W. Então se α é um vetor arbitrário em W e c um escalar arbitrário, o vetor $c\alpha = c\,\alpha + 0$ está em W. Em particular $(-1)\alpha = -\alpha$ está em W. Finalmente se α e β estão em W, então $\alpha + \beta = 1\,\alpha + \beta$ está em W. Assim, W é um subespaço de V.

Reciprocamente, se W é um subespaço de V, α e β estão em W e c é um escalar, certamente $c\,\alpha + \beta$ está em W.

Exemplo 6.

(a) Se V é um espaço vetorial arbitrário, V é um subespaço de V; o subconjunto constituído sòmente pelo vetor nulo é um subespaço de V, denominado o **subespaço nulo** de V.

(b) Em F^n, o conjunto das n-uplas $(x_1, \ldots, x_n)$ com $x_1 = 0$ é um subespaço: contudo, o conjunto das n-uplas com $x_1 = 1 + x_2$ não é um subespaço ($n \geq 2$).

(c) O espaço das funções polinomiais sôbre o corpo F é um subespaço do espaço de tôdas as funções de F em F.

(d) Uma $n \times n$ matriz (quadrada) A sôbre o corpo F é **simétrica** se $A_{ij} = A_{ji}$ para todos i e j. As matrizes simétricas formam um subespaço do espaço de tôdas as $n \times n$ matrizes sôbre F.

(e) Uma $n \times n$ matriz (quadrada) A sôbre o corpo C dos números complexos é **hermitiana** (ou **auto-adjunta**) se

$$A_{jk} = \overline{A_{kj}}$$

para todos j, k, sendo que a barra indica conjugação complexa. Uma 2×2 matriz é hermetiana se e sòmente se é da forma

$$\begin{bmatrix} z & x + iy \\ x - iy & w \end{bmatrix}$$

onde x, y, z e w são números reais. O conjunto de tôdas as matrizes hermitianas *não* é um subespaço do espaço de tôdas as $n \times n$ matrizes sôbre C. De fato, se A é hermitiana, todos os elementos A_{11}, $A_{22}, \ldots$, de sua diagonal são números reais mas os elementos diagonais de iA em geral não são reais. Por outro lado, verifica-se fàcilmente que o conjunto das $n \times n$ matrizes hermitianas complexas é um espaço vetorial sôbre o corpo R dos números reais (com as operações usuais).

Exemplo 7. O espaço-solução de um sistema de equações lineares homogêneas. Seja A uma $m \times n$ matriz sôbre F. Então o conjunto de tôdas as $n \times 1$ matrizes - (colunas) X sôbre F tais que $AX = 0$ é um subespaço do espaço de tôdas as $n \times 1$ matrizes sôbre F. Para demonstrar isto precisamos mostrar que $A(cX + Y) = 0$ para $AX = 0$, $AY = 0$ e c um escalar arbitrário em F. Isto decorre imediatamente do seguinte fato geral:

Lema. *Se* A *é uma* m $\times$ n *matriz sôbre* F *e* B, C *são* n $\times$ p *matrizes sôbre* F, *então*

$$A(dB + C) = d(AB) + AC \tag{2-11}$$

para todo escalar d *em* F.

Demonstração.

$$\begin{aligned} [A(dB + C)]_{ij} &= \sum_k A_{ik}(dB + C)_{kj} \\ &= \sum_k (dA_{ik}B_{kj} + A_{ik}C_{kj}) \\ &= d\sum_k A_{ik}B_{kj} + \sum_k A_{ik}C_{kj}) \\ &= d(AB)_{ij} + (AC)_{ij} \\ &= [d(AB) + AC]_{ij}. \end{aligned}$$

Anàlogamente, pode-se mostrar que $(dB + C)A = d(BA) + CA$, se as somas e produtos de matrizes estão definidos.

Teorema 2. *Seja* V *um espaço vetorial sôbre o corpo F. A interseção de uma coleção arbitrária de subespaços de* V *é um subespaço de* V.

Demonstração. Seja $\{W_a\}$ uma coleção de subespaços de V e seja $W = \bigcap_a W_a$ a sua interseção. Recordemos que W é definido como sendo o conjunto dos elementos pertencentes simultâneamente a W_a (ver Apêndice). Como cada W_a é um subespaço, todos contêm o vetor nulo. Assim o vetor nulo está na interseção W e W é não vazio. Sejam α e β vetores em W e seja c um escalar. Pela definição de W, tanto α como β pertencem a cada W_a e, como cada W_a é um subespaço, o vetor $(c\alpha + \beta)$ está em todo W_a. Assim, $(c\alpha + \beta)$ está em W. Pelo Teorema 1, W é um subespaço de V.

Do Teorema 2 decorre que se S é uma coleção arbitrária de vetores em V, então existe um menor subespaço de V que contém S, isto é, um subespaço que contém S e que está contido em todos os outros subespaços que contêm S.

Definição. *Seja* S *um conjunto de vetores num espaço vetorial* V. O **subespaço gerado** *por* S *é definido como sendo a interseção* W *de todos os subespaços de* V *que contêm* S. *Quando* S *é um conjunto finito de vetores,* $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, *denominaremos* W *simplesmente o* **subespaço gerado pelos vetores** $\alpha_1, \alpha_2, \ldots, \alpha_n$.

Teorema 3. *O subespaço gerado por um subconjunto não vazio* S *de um espaço vetorial* V *é o conjunto de tôdas as combinações lineares de vetores em* S.

Demonstração. Seja W o subespaço gerado por S. Então, cada combinação linear

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + \ldots + x_m\alpha_m$$

de vetores $\alpha_1, \alpha_2, \ldots, \alpha_m$ em S evidentemente está em W. Assim, W contém o conjunto L de tôdas as combinações lineares de vetores em S. O conjunto L, por outro lado, contém S e é não vazio. Se α, β pertencem a L, então α é uma combinação linear,

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + \ldots + x_m\alpha_m$$

de vetores α_i em S e β é uma combinação linear

$$\beta = y_1\beta_1 + y_2\beta_2 + \ldots + y_n\beta_n$$

de vetores β_j em S. Para cada escalar c,

$$c\alpha + \beta = \sum_{i=1}^{m} (cx_i)\alpha_i + \sum_{j=1}^{n} y_j\beta_j.$$

Logo $c\alpha + \beta$ pertence a L. Assim, L é um subespaço de V.

Mostramos acima que L é um subespaço de V que contém S e também que todo subespaço que contém S contém L. Decorre que L é a interseção de todos os subespaços que contêm S, isto é, que L é o subespaço gerado pelo conjunto S.

Definição. *Se* $S_1, S_2, \ldots, S_k$ *são subconjuntos de um espaço vetorial* V, *o conjunto de tôdas as somas*

$$\alpha_1 + \alpha_2 + \ldots + \alpha_k$$

de vetores x_i *em* S_i *é dito a* **soma** *dos subconjuntos* $S_1, S_2, \ldots, S_k$ *e é indicado por*

$$S_1 + S_2 + \ldots + S_k$$

ou por

$$\sum_{i=1}^{k} S_i.$$

Se $W_1, W_2, \ldots, W_k$ são subespaços de V, então vê-se fàcilmente que a soma

$$W + W_1 + W_2 + \ldots + W_k$$

é um subespaço de V que contém cada um dos subespaços W_i. Disto decorre, como na demonstração do Teorema 3, que W é o subespaço gerado pela reunião de $W_1, W_2, \ldots, W_k$.

Exemplo 8. Seja F um subcorpo do corpo C dos números complexos. Suponhamos que

$$\begin{aligned} \alpha_1 &= (1, 2, 0, 3, 0) \\ \alpha_2 &= (0, 0, 1, 4, 0) \\ \alpha_3 &= (0, 0, 0, 0, 1). \end{aligned}$$

Pelo Teorema 3, um vetor α está no subespaço W de F^5 gerado por $\alpha_1, \alpha_2, \alpha_3$ sè, e sòmente se, existem escalares c_1, c_2, c_3 em F tais que

$$\alpha = c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3.$$

Portanto, W consiste de todos os vetores da forma

$$\alpha = (c_1, 2c_1, c_2, 3c_1 + 4c_2, c_3)$$

onde c_1, c_2, c_3 são escalares arbitrários em F. W pode ser descrito de outra forma como sendo o conjunto de tôdas as quíntuplas

$$\alpha = (x_1, x_2, x_3, x_4, x_5)$$

com x_i em F tais que

$$\begin{aligned} x_2 &= 2x_1 \\ x_4 &= 3x_1 + 4x_3. \end{aligned}$$

Assim, (—3, —6, 1, —5, 2) está em W, enquanto que (2, 4, 6, 7, 8) não está.

Exemplo 9. Seja F um subcorpo do corpo C dos números complexos e seja V o espaço vetorial das 2×2 matrizes sôbre F. Seja W_1 o subconjunto de V constituído por tôdas as matrizes da forma

$$\begin{bmatrix} x & y \\ z & 0 \end{bmatrix}$$

onde x, y, z são escalares arbitrários em F. Finalmente, seja W_2 o subconjunto de V constituído por tôdas as matrizes da forma

$$\begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix}$$

onde x e y são escalares arbitrários em F. Então, W_1 e W_2 são subespaços de V. Além disso

$$V = W_1 + W_2$$

pois

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & d \end{bmatrix}.$$

O subespaço $W_1 \cap W_2$ consiste de tôdas as matrizes da forma

$$\begin{bmatrix} x & 0 \\ 0 & 0 \end{bmatrix}.$$

Exemplo 10. Seja A uma $m \times n$ matriz sôbre um corpo F. Os **vetores-linhas** de A são os vetores em F^n dados por $\alpha_i = (A_{i1}, \ldots, A_{in})$, $i = 1, \ldots, m$. O subespaço de F^n gerado pelos vetores-linhas de A é denominado o **espaço-linha** de A. O subespaço considerado no Exemplo 8 é o espaço-linha da matriz

$$A = \begin{bmatrix} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Êle também é o espaço-linha da matriz

$$B = \begin{bmatrix} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ -4 & -8 & 1 & -8 & 0 \end{bmatrix}.$$

Exemplo 11. Seja V o espaço das funções polinomiais sôbre F. Seja S o subconjunto de V constituído pelas funções polinomiais $f_0, f_1, f_2, \ldots$, definidas por

$$f_n(x) = x^n, \text{ n } = 0, 1, 2, \ldots$$

Então V é o subespaço gerado pelo conjunto S.

Exercícios

1. Quais dos seguintes conjuntos de vetores $\alpha = (a_1, \ldots, a_n)$ em R^n são subespaços de R^n? ($n \geq 3$)

(a) todos α tais que $a_1 \geq 0$;
(b) todos α tais que $a_1 + 3a_2 = a_3$;
(c) todos α tais que $a_2 = a_1^2$;
(d) todos α tais que $a_1 a_2 = 0$;
(e) todos α tais que x a_2 seja racional.

2. Seja V o espaço vetorial (real) de tôdas as funções f de R em R. Quais dos seguintes conjuntos de funções são subespaços de V?

(a) tôdas f tais que $f(x^2) = f(x^2)$;
(b) tôdas f tais que $f(0) = f(1)$;
(c) tôdas f tais que $f(3) = 1 + f(-5)$;
(d) tôdas f tais que $f(-1) = 0$;
(e) tôdas f que são contínuas.

3. O vetor $(3, -1, 0, -1)$ está no subespaço de R^4 gerado pelos vetores $(2, -1, 3, 2)$, $(-1, 1, 1, -3)$ e $(1, 1, 9, -5)$?

4. Seja W o conjunto de todos os $(x_1, x_2, x_3, x_4, x_5)$ em R^5 que satisfazem

$$\begin{aligned} 2x_1 - x_2 + \tfrac{4}{3}x_3 - x_4 &= 0 \\ x_1 + \tfrac{2}{3}x_3 - x_5 &= 0 \\ 9x_1 - 3x_2 + 6x_3 - 3x_4 - 3x_5 &= 0. \end{aligned}$$

Determinar um conjunto finito de vetores que gere W.

5. Seja F um corpo e seja n um inteiro positivo ($n \geq 2$). Seja V o espaço vetorial das $n \times n$ matrizes sôbre F. Quais dos seguintes conjuntos de matrizes A em V são subespaços de V?

(a) tôdas A inversíveis;
(b) tôdas A não-inversíveis;
(c) tôdas A tais que $AB = BA$, onde B é uma certa matriz fixa em V;
(d) tôdas A tais que $A^2 = A$.

6. (a) Demonstrar que os únicos subespaços de R^1 são R^1 e o subespaço nulo.

(b) Demonstrar que um subespaço de R^2 ou é R^2, ou é o subespaço nulo ou então consiste de todos os múltiplos escalares de um certo vetor fixo em R^2. (O último tipo de subespaço é (intuitivamente) uma reta pela origem.)

(c) Você é capaz de descrever os subespaços de R^3?

7. Sejam W_1 e W_2 subespaços de um espaço vetorial V tais que a reunião de W_1 e W_2 também seja um subespaço. Demonstrar que um dos espaços W_i está contido no outro.

8. Seja V o espaço vetorial das funções de R em R; seja V_p o subconjunto das funções pares, $f(-x) = f(x)$; seja V_i o subconjunto das funções ímpares, $f(-x) = -f(x)$.

(a) Demonstrar que V_p e V_i são subespaços de V.

(b) Demonstrar que $V_p + V_i = V$.

(c) Demonstrar que $V_p \cap V_i = \{0\}$.

9. Sejam W_1 e W_2 subespaços de um espaço vetorial V tais que $W_1 + W_2 = V$ e $W_1 \cap W_2 = \{0\}$. Determinar que para cada vetor α em V existem vetores *bem determinados* α_1 em W_1 e α_2 em W_2, tais que $\alpha = \alpha_1 + \alpha_2$.

2.3 Bases e Dimensão

Passamos agora à tarefa de atribuir uma dimensão a certos espaços vetoriais. Apesar de associarmos usualmente "dimensão" a algo geométrico, precisamos encontrar uma definição algébrica adequada da dimensão de um espaço vetorial. Isto será feito através do conceito de uma base para o espaço.

Definição. *Seja* V *um espaço vetorial sôbre* F. *Um subconjunto* S *de* V *é dito* **linearmente dependente** (*ou, simplesmente,* **dependente**) *se existem vetores distintos* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *em* S *e escalares* $c_1, c_2, \ldots, c_n$ *em* F, *não todos nulos, tais que*

$$c_1\alpha_1 + c_2\alpha_2 + \ldots + c_n\alpha_n = 0.$$

Um conjunto que não é linearmente dependente é dito **linearmente independente**. *Se o conjunto* S *contém apenas um número finito de vetores* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *dizemos, às vêzes, que* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *são dependentes (ou independentes) em vez de dizer que* S *é dependente (ou independente). Decorrem fàcilmente da definição as conseqüências seguintes:*

(a) *Todo conjunto que contém um conjunto linearmente dependente é linearmente dependente.*

(b) *Todo subconjunto de um conjunto linearmente independente é linearmente independente.*

(c) *Todo conjunto que contém o vetor nulo é linearmente dependente, pois* $1 \,.\, 0 = 0$.

(d) *Um conjunto* S *de vetores é linearmente independente se e sòmente se todo subconjunto finito de* S *é linearmente independente, isto é, se e sòmente se para quaisquer vetores distintos* $\alpha_1, \ldots, \alpha_n$ *em* S $c_1\alpha_1 + \ldots + c_n\alpha_n = 0$ *implica que cada* $c_i = 0$.

Exemplo 12. Seja F um subcorpo do corpo dos números complexos. Em F_3 os vetores

$$\begin{aligned} \alpha_1 &= (\ \ 3,\ 0,\ -3) \\ \alpha_2 &= (-1,\ 1,\ \ 2) \end{aligned}$$

$$\begin{aligned} \alpha_3 &= (\ \ 4,\ 2,\ -2) \\ \alpha_4 &= (\ \ 2,\ 1,\ \ \ 1) \end{aligned}$$

são linearmente dependentes, pois

$$2\alpha_1 + 2\alpha_2 - \alpha_3 + 0 \,.\, \alpha_4 = 0.$$

Os vetores

$$\begin{aligned} \epsilon_1 &= (1,\ 0,\ 0) \\ \epsilon_2 &= (0,\ 1,\ 0) \\ \epsilon_3 &= (0,\ 0,\ 1) \end{aligned}$$

são linearmente independentes.

Definição. *Seja* V *um espaço vetorial. Uma* **base** *de* V *é um conjunto linearmente independente de vetores em* V *que gera* V.

Exemplo 13. Seja F um corpo e, em F_n, seja S o subconjunto constituído dos vetores $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ definidos por

$$\begin{aligned} \epsilon_1 &= (1,\ 0,\ 0, \ldots,\ 0) \\ \epsilon_2 &= (0,\ 1,\ 0, \ldots,\ 0) \\ &\ \ \ldots\ldots\ldots\ldots \\ \epsilon_n &= (0,\ 0,\ 0, \ldots,\ 1). \end{aligned}$$

Sejam $x_1, x_2, \ldots, x_n$ escalares em F e coloquemos $\alpha = x_1\epsilon_1 + x_2\epsilon_2 + \ldots + x_n\epsilon_n$. Então

$$\alpha = (x_1, x_2, \ldots, x_n). \tag{2-12}$$

Isto mostra que $\epsilon_1, \ldots, \epsilon_n$ geram F^n. Como $\alpha = 0$ se e sòmente se $x_1 = x_2 = \ldots = x_n = 0$, os vetores $\epsilon_1, \ldots, \epsilon_n$ são linearmente independentes. O conjunto $S = \{\epsilon_1, \ldots, \epsilon_n\}$ é portanto uma base de F^n. Denominamos esta base particular a **base canônica** de F^n.

Exemplo 14. Seja F um subcorpo do corpo dos números complexos. Usando a notação do Exemplo 11 consideraremos o subespaço V do espaço das funções polinomiais sôbre F que é gerado pelas funções f_0, f_1, f_2. Suponhamos que c_0, c_1, c_2 sejam escalares em F tais que

$$c_0f_0 + c_1f_1 + c_2f_2 = 0.$$

Isto significa que para cada x em F,

$$c_0 + c_1x + c_2x^2 = 0.$$

Tomando $x = 0$, vemos que $c_0 = 0$ e, fazendo $x = 1$ e $x = -1$, obtemos as equações

$$\begin{aligned} c_1 + c_2 &= 0 \\ -c_1 + c_2 &= 0. \end{aligned}$$

Somando e subtraindo, concluímos que $2c_2 = 0$ e $2c_1 = 0$, donde concluímos que $c_1 = 0$ e $c_2 = 0$. Assim as funções f_0, f_1, f_2 são linearmente independentes e formam uma base de V. Posteriormente, mostraremos que o conjunto infinito constituído por tôdas as funções f_n, $n = 0, 1, 2, \ldots$, é uma base do espaço de tôdas as funções polinomiais sôbre F. Quando houvermos feito isto, teremos um exemplo de uma base infinita para um espaço vetorial. Notemos que, apesar de

$$\{f_0, f_1, f_2, f_3, \ldots\}$$

ser um conjunto infinito que é uma base para o espaço das funções polinomiais sôbre F, isto não quer dizer que estejamos considerando combinações lineares infinitas. Cada função polinomial será uma combinação linear de um certo número *finito* das funções f_n.

Teorema 4. *Seja* V *um espaço vetorial gerado par um conjunto finito de vetores* $\beta_1, \beta_2, \ldots, \beta_m$. *Então, todo conjunto independente de vetores em* V *é finito e contém no máximo* m *elementos.*

Demonstração. Para demonstrar o teorema basta mostrar que todo subconjunto S de V que contém mais de m vetores é linearmente dependente. Seja S um tal conjunto. Em S existem vetores distintos $\alpha_1, \alpha_2, \ldots, \alpha_n$ com $n > m$. Como $\beta_1, \ldots, \beta_m$ geram V existem escalares A_{ij} em F tais que

$$\alpha_j = \sum_{i=1}^{m} A_{ij}\beta_i.$$

Para n escalares arbitrários $x_1, x_2, \ldots, x_n$ temos

$$\begin{aligned} x_1\alpha_1 + \ldots + x_n\alpha_n &= \sum_{j=1}^{n} x_j\alpha_j \\ &= \sum_{j=1}^{n} x_j \sum_{i=1}^{m} A_{ij}\beta_i \\ &= \sum_{j=1}^{n} \sum_{i=1}^{m} (A_{ij}x_j)\beta_i \\ &= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} A_{ij}x_j \right) \beta_i. \end{aligned}$$

Como $n > m$, o Teorema 6 do Capítulo 1 implica que existem escalares $x_1, x_2, \ldots, x_n$ não todos nulos, tais que

$$\sum_{j=1}^{n} A_{ij}x_j = 0, \qquad 1 \leq i \leq m.$$

Logo, $x_1\alpha_1 + x_2\alpha_2 + \ldots + x_n\alpha_n = 0$. Isto mostra que S é um conjunto linearmente dependente.

Definição. *Um espaço vetorial* V *é de* **dimensão finita** *se êle possui uma base finita.*

Corolário 1. *Se* V *é um espaço vetorial de dimensão finita, então duas quaisquer bases de* V *têm o mesmo número (finito) de elementos.*

Demonstração. Sendo de dimensão finita, V possui uma base finita

$$\{\beta_1, \beta_2, \ldots, \beta_m\}.$$

Pelo Teorema 4, tôda base de V é finita e contém no máximo m elementos. Assim, se $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ é uma base, $n \leq m$. Pela mesma razão, $m \leq n$. Logo $m = n$.

Definição. *Se* V *é um espaço vetorial de dimensão finita,* a **dimensão** *de* V *é definida como sendo o número de elementos de uma base de* V. *Indicaremos a dimensão de um espaço vetorial* V *de dimensão finita por dim* V.

Exemplo 15. Se F é um corpo, a dimensão de F^n é n, pois a base canônica de F^n contém n vetores.

Corolário 2. *Seja* V *um espaço vetorial* n*-dimensional. Então*

(a) *todo conjunto de vetores em* V *que contém mais de* n *vetores é linearmente dependente.*

(b) *nenhum conjunto contendo menos de* n *vetores pode gerar* V.

Lema. *Seja* S *um subconjunto linearmente independente de um espaço vetorial* V. *Suponhamos que* β *seja um vetor em* V *que não esteja no subespaço gerado por* S. *Então o conjunto obtido acrescentando-se* β *a* S *é linearmente independente.*

Demonstração. Suponhamos que $\alpha_1, \ldots, \alpha_m$ sejam vetores distintos em S e que

$$c_1\alpha_1 + \ldots + c_m\alpha_m + b\beta = 0.$$

Então $b = 0$, caso contrário

$$\beta = \left(\frac{c_2}{b}\right)\alpha_1 + \ldots + \left(-\frac{c_m}{b}\right)\alpha_m$$

e β estaria no subespaço gerado por S. Assim, $c_1\alpha_1 + \ldots + c_m\alpha_m = 0$, e como S é um conjunto linearmente independente, cada $c_i = 0$.

Teorema 5. *Se* W *é um subespaço de um espaço vetorial* V *de dimensão finita, todo subconjunto de* W *que é linearmente independente é finito e é parte de uma base (finita) de* W.

Demonstração. Suponhamos que S_0 seja um subconjunto de W linearmente independente. Se S é um subconjunto de W linearmente independente contendo S_0, então S também é um subconjunto de W linearmente independente; como V é de dimensão finita, S contém no máximo dim V elementos. Portanto, existe um subconjunto S de W linearmente independente que é maximal e contém S_0. Como S é um subconjunto de W linearmente independente e maximal contendo S_0, o lema anterior mostra que W é o subespaço gerado por S. Logo, S é uma base de W e o conjunto original S_0 é parte de uma base de W.

Corolário 1. *Se* W *é um subespaço próprio de um espaço vetorial* V *de dimensão finita, então* W *é de dimensão finita e dim* W $<$ *dim* V.

Demonstração. Podemos supor que W contém um vetor $\alpha \neq 0$. Pelo Teorema 5 e sua demonstração, existe uma base de W que contém α e no máximo dim V elementos. Logo W é de dimensão finita e dim $W \leq$ dim V. Como W é subespaço próprio, existe um vetor β em V que não está em W. Acrescentando β a uma base arbitrária de W obtemos um subconjunto de V linearmente independente. Portanto dim $W <$ dim V.

Corolário 2. *Num espaço vetorial* V *de dimensão finita todo conjunto não-vazio de vetores linearmente independentes é parte de uma base.*

Corolário 3. *Seja* A *uma* n $\times$ n *matriz sôbre um corpo* F *e suponhamos que os vetores-linhas de* A *formem um conjunto de vetores de* F^n *linearmente independentes. Então* A *é inversível.*

Demonstração. Sejam $\alpha_1, \alpha_2, \ldots, \alpha_n$ os vetores-linhas de A e suponhamos que W seja o subespaço de F^n gerado por $\alpha_1, \alpha_2, \ldots \alpha_n$. Como $\alpha_1, \alpha_2, \ldots, \alpha_n$ são linearmente independentes, a dimensão de W é n. O Corolário 1 mostra agora que $W = F^n$. Logo, existem escalares B_{ij} em F tais que

$$\epsilon_i = \sum_{j=1}^{n} B_{ij}\alpha_j, \qquad 1 \leq i \leq n$$

onde $\{\epsilon_1, \epsilon_2, \ldots, \epsilon_n\}$ é a base canônica de F^n. Portanto, para a matriz B com elementos B_{ij}, temos

$$BA = I.$$

Teorema 6. *Se* W_1 *e* W_2 *são subespaços de dimensão finita de um espaço vetorial* V, *então* $W_1 + W_2$ *é de dimensão finita e*

$$\dim W_1 + \dim W_2 = \dim (W_1 \cap W_2) + \dim (W_1 + W_2).$$

Demonstração. Pelo Teorema 5 e seus carolários, $W_1 \cap W_2$ tem uma base finita $\{\alpha_1, \ldots, \alpha_k\}$ que é parte de uma base

$$\{\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_m\} \text{ de } W_1$$

e parte de uma base

$$\{\alpha_1, \ldots, \beta_k, \gamma_1, \ldots, \gamma_n\} \text{ de } W_2.$$

O subespaço $W_1 + W_2$ é gerado pelos vetores

$$\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_m, \gamma_1, \ldots, \gamma_1$$

e êstes vetores formam um conjunto independente. De fato, suponhamos que

$$\Sigma\, x_i\alpha^i + \Sigma\, y_j\beta_j + \Sigma\, z_r\gamma_r = 0.$$

Então

$$-\Sigma\, z_r\gamma_r = \Sigma\, x_i\alpha_i + \Sigma\, y_j\beta_j$$

o que mostra que $\Sigma\, z_r\gamma_r$ pertence a W_1. Como $\Sigma\, z_r\gamma_r$ pertence também a W_2, segue que

$$\Sigma\, z_r\gamma_r = \Sigma\, c_i\alpha_i$$

para certos escalares $c_1, \ldots, c_k$. Por ser o conjunto

$$\{\alpha_1, \ldots, \alpha_k, \gamma_1, \ldots, \gamma_n\}$$

independente, cada um dos escalares $z_r = 0$. Portanto,

$$\Sigma\, x_i\alpha_i + \Sigma\, y_j\beta_j = 0$$

e como

$$\{\alpha_1, \ldots, \alpha_k, \beta_1, . \;\; ., \beta_m\}$$

também é um conjunto independente, cada $x_i = 0$ e cada $y_j = 0$. Assim,

$$\{\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_m, \gamma_1, \ldots, \gamma_n\}$$

é uma base de $W_1 + W_2$. Finalmente

$$\begin{aligned}\dim W_1 + \dim W_2 &= (k + m) + (k + n)\\ &= k + (m + k + n)\\ &= \dim(W_1 \cap W_2) + \dim(W_1 + W_2).\end{aligned}$$

Exercícios

1. Demonstrar que, se dois vetores são linearmente dependentes, um dêles é um múltiplo escalar do outro.

2. Os vetores

$$\begin{aligned}\alpha_1 &= (1, 1, 2, 4) & \alpha_2 &= (2, -1, -5, 2)\\ \alpha_3 &= (1, -1, -4, 0) & \alpha_4 &= (2, 1, 1, 6)\end{aligned}$$

são linearmente independentes em R^4?

3. Determinar uma base do subespaço de R^4 gerado pelos quatro vetores do Exercício 2.

4. Mostrar que os vetores

$$\alpha_1 = (1, 0, -1), \qquad \alpha_2 = (1, 2, 1), \qquad \alpha_3 = (0, -3, 2)$$

formam uma base de R^3. Exprimir cada um dos vetores da base canônica como combinações lineares de α_1, α_2, e α_3.

5. Determinar três vetores em R^3 que sejam linearmente dependentes e tais que dois quaisquer dêles sejam linearmente independentes.

6. Seja V o espaço vetorial das 2×2 matrizes sôbre o corpo F. Demonstrar que V tem dimensão 4 mostrando uma base de V que tenha 4 elementos.

7. Seja V o espaço vetorial do Exercício 6. Seja W_1 o conjunto das matrizes da forma

$$\begin{bmatrix} x & -x \\ y & z \end{bmatrix}$$

e seja W_2 o conjunto das matrizes da forma

$$\begin{bmatrix} a & b \\ -a & c \end{bmatrix}.$$

(a) Demonstrar que W_1 e W_2 são subespaços de V.
(b) Determinar as dimensões de W_1, W_2, $W_1 + W_2$ e $W_1 \cap W_2$.

8. Novamente, seja V o espaço das 2×2 matrizes sôbre F. Determinar uma base $\{A_1, A_2, A_3, A_4\}$ de V tal que $A_j^2 = A_j$ para cada j.

9. Seja V um espaço vetorial sôbre um subcorpo F do corpo dos números complexos. Suponhamos que α, β e γ sejam vetores de V linearmente independentes. Demonstrar que $(\alpha + \beta)$, $(\beta + \gamma)$ e $(\gamma + \alpha)$ são linearmente independentes.

10. Seja V um espaço vetorial sôbre o corpo F. Suponhamos que exista um número finito de vetores $\alpha_1, \ldots, \alpha_r$ de V que gerem V. Demonstrar que V é de dimensão finita.

11. Seja V o conjunto das 2×2 matrizes A com elementos *complexos* satisfazendo $A_{11} + A_{22} = 0$.

(a) Mostrar que V é um espaço vetorial sôbre o corpo dos números *reais*, com as operações usuais de adição de matrizes e multiplicação de uma matriz por um escalar.

(b) Determinar uma base dêsse espaço vetorial.

(c) Seja W o conjunto de tôdas as matrizes A em V tais que $A_{21} = -\bar{A}_{12}$ (a barra indica conjugação completa). Demonstrar que W é um subespaço de V e determinar uma base de W.

12. Demonstrar que o espaço das $m \times n$ matrizes sôbre o corpo F tem dimensão mn, mostrando uma base para êste espaço.

13. Discutir o Exercício 9, para o caso de V ser um espaço vetorial sôbre o corpo formado por dois elementos descritos no Exercício 5, Seção 1.1 (p. 5).

14. Seja V o conjunto dos números reais. Consideremos V como um espaço vetorial sôbre o corpo dos números *racionais*, com as operações usuais. Demonstrar que êste espaço vetorial *não é* de dimensão finita.

2.4 Coordenadas

Uma das características úteis de uma base $\mathcal{B}$ de um espaço n-dimensional V é essencialmente que ela nos permite introduzir coordenadas em V análogas às "coordenadas naturais" x_i de um vetor $\alpha = (x_1, \ldots, x_n)$ do espaço F^n. Em assim sendo, as coordenadas de um vetor α de V em relação à base $\mathcal{B}$ serão os escalares que servem para exprimir α como uma combinação linear dos vetores da base. Assim, gostaríamos de considerar as coordenadas naturais de um vetor α de F^n como sendo definidas por α e pela base canônica de F^n; contudo, ao adotarmos êste ponto de vista precisamos ter um certo cuidado. Se

$$\alpha = (x_1, \ldots, x_n) = \Sigma\, x_i\epsilon_i$$

e $\mathcal{B}$ é a base canônica de F^n, como são as coordenadas de α determinadas por $\mathcal{B}$ e α? Uma maneira de formular a resposta é esta: Um dado vetor α é expresso de maneira única como uma combinação linear dos vetores da base canônica, e a i-ésima coordenada x_i de α é o coeficiente de ϵ_i nesta expressão. Sob êste ponto de vista podemos dizer qual é a i-ésima coordenada, pois temos uma ordenação "natural" dos vetores da base canônica, isto é, temos uma regra para determinar qual é o "primeiro" vetor da base, qual é o "segundo" e assim por diante. Se $\mathcal{B}$ é uma base arbitrária do espaço n-dimensional V, não teremos provàvelmente nenhuma ordenação natural para os vetores de $\mathcal{B}$ e será portanto necessário impormos uma certa ordem sôbre êsses vetores antes de podermos definir "a i-ésima coordenada de α em relação a $\mathcal{B}$".

Se S é um conjunto com n elementos, o que é uma ordenação dos elementos S? Existem muitas definições dêste conceito, apesar de diferirem apenas superficialmente. Adotaremos a seguinte: Uma **ordenação** do conjunto S, de n elementos, é uma função do conjunto dos inteiros positivos $1, \ldots, n$ sôbre o conjunto S. Portanto uma ordenação do conjunto é simplesmente uma regra para nos dizer que elemento deve ser considerado como o primeiro elemento de S, que elemento é segundo, etc. Uma **base ordenada** de um espaço vetorial V de dimensão finita é uma base $\mathcal{B}$ de V, mais uma cordenação fixa dos elementos (vetores) de $\mathcal{B}$. Freqüentemente o que mais convém é descrever essa base ordenada enumerando os vetores de $\mathcal{B}$ de uma maneira bem definida. Assim, diremos que

$$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$$

é uma base ordenada de V, se ficar claro qual vetor do conjunto β é o i-ésimo, α_i.

Suponhamos agora que V seja um espaço vetorial de dimensão finita sôbre o corpo F e que

$$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$$

seja uma base ordenada de V. Dado α em V, existe uma única n-upla $(x_1, \ldots, x_n)$ de escalares tal que

$$\alpha = \sum_{i=1}^{n} x_i\alpha_i.$$

A n-upla é única, pois, se tivéssemos

$$\alpha = \sum_{i=1}^{n} z_i\alpha_i$$

então

$$\sum_{i=1}^{n} (x_i - z_i)\alpha_i = 0$$

e a independência linear dos α_i nos diria que $x_i - z_i = 0$ para cada i. Denominaremos x_i a i-ésima **coordenada de α em relação à base ordenada**

$$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}.$$

Se

$$\beta = \sum_{i=1}^{n} y_i\alpha_i$$

então

$$\alpha + \beta = \sum_{i=1}^{n} (x_i + y_i)\alpha_i$$

de modo que a i-ésima coordenada de $(\alpha + \beta)$ em relação a esta base ordenada é $(x_i + y_i)$. Anàlogamente, a i-ésima coordenada de $(c\alpha)$ é cx_i. Devemos também notar que tôda n-upla $(x_1, \ldots, x_n)$ de F^n é a n-upla de coordenadas de algum vetor de V, a saber, o vetor

$$\sum_{i=1} x_i\alpha_i.$$

Resumindo, cada base ordenada de V determina uma correspondência bijetora

$$\alpha \rightarrow (x_1, \ldots, x_n)$$

entre o conjunto dos vetores de V e o conjunto das n-uplas de F^n. Esta correspondência tem a propriedade de que o correspondente de $(\alpha + \beta)$ é a soma em F^n dos correspondentes de α e β, e que o correspondente de $(c\alpha)$ é o produto em F^n do escalar c pelo correspondente de α.

Poder-se-ia perguntar neste ponto por que não tomar simplesmente uma base ordenada de V e descrever cada vetor de V por sua correspondente n-upla de coordenadas, visto que teríamos então a conveniência de operar apenas com n-uplas. Isto faria malograr nosso objetivo, por duas razões. Primeiro, como indica a nossa definição de espaço vetorial, estamos tentando aprender a raciocinar com espaços vetoriais como sistemas algébricos. Segundo, mesmo nos casos em que usamos coordenadas, os resultados importantes decorrem de nossa habilidade de mudar o sistema de coordenadas, isto é, mudar a base ordenada.

Freqüentemente o mais conveniente será usar a **matriz das coordenadas de α em relação à base ordenada $\mathcal{B}$**:

$$X = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

em vez da n-upla $(x_1, \ldots, x_n)$ das coordenadas. Para indicar que esta matriz de coordenadas depende da base, usaremos o símbolo

$$[\alpha]_{\mathcal{B}}$$

para a matriz das coordenadas do vetor α em relação à base ordenada $\mathcal{B}$. Esta notação será particularmente útil ao passarmos agora a descrever o que acontece com as coordenadas de um vetor α quando passamos de uma base ordenada a outra.

Suponhamos então que V seja n-dimensional e que

$$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\} \quad \text{e} \quad \mathcal{B}' = \{\alpha'_1, \ldots, \alpha'_n\}$$

sejam duas bases ordenadas de V. Existem escalares P_{ij}, bem determinados, tais que

$$\alpha'_j = \sum_{i=1}^{n} P_{ij}\alpha_i, \qquad 1 \leq j \leq n. \tag{2—13}$$

Sejam $x'_1, \ldots, x'_n$ as coordenadas de um dado vetor α em relação à base ordenada $\mathcal{B}'$. Então

$$\begin{aligned}
\alpha &= x'_1\alpha'_1 + \ldots + x'_n\alpha'_n \\
&= \sum_{j=1}^{n} x'_j\alpha'_j \\
&= \sum_{j=1}^{n} x'_j \sum_{i=1}^{n} P_{ij}\alpha_i \\
&= \sum_{j=1}^{n} \sum_{i=1}^{n} (P_{ij}x'_j)\alpha_i \\
&= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} P_{ij}x'_j \right) \alpha_i.
\end{aligned}$$

Portanto, obtemos a relação

$$\alpha = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} P_{ij}x'_j \right) \alpha_i. \tag{2—14}$$

Como as coordenadas $x_1, x_2, \ldots, x_n$ de α em relação à base ordenada $\mathcal{B}$ são determinadas de modo único, decorre de (2—14) que

$$x_i = \sum_{j=1}^{n} P_{ij}x'_j, \qquad 1 \leq i \leq n. \tag{2—15}$$

Seja P a $n \times n$ matriz cujo elemento i, j é o escalar P_{ij} e sejam X e X' as matrizes das coordenadas do vetor α em relação às bases ordenadas $\mathcal{B}$ e $\mathcal{B}'$. Podemos então reformular (2—15) como

$$X = PX'. \tag{2—16}$$

Como $\mathcal{B}$ e $\mathcal{B}'$ são conjuntos linearmente independentes, $X = 0$ se e sòmente se $X' = 0$. Assim, de (2—16) e do Teorema 7 do Capítulo 1, decorre que P é inversível. Logo,

$$X' = P^{-1}X. \tag{2—17}$$

Se usarmos a notação acima introduzida para a matriz das coordenadas de um vetor em relação a uma base ordenada, então (2—16) e (2—17) afirmam que

$$[\alpha]_{\mathcal{B}} = P[\alpha]_{\mathcal{B}'}$$
$$\text{e } [\alpha]_{\mathcal{B}'} = P\mathrm{P}^{1}[\alpha]_{\mathcal{B}}.$$

Portanto, a discussão precedente pode ser resumida como segue.

Teorema 7. *Seja* V *um espaço vetorial* n*-dimensional sôbre o corpo* F *e sejam* $\mathcal{B}$ *e* $\mathcal{B}'$ *duas bases ordenadas de* V. *Então existe uma única* n $\times$ n *matriz* P, *necessàriamente inversível, com elementos em* F, *tal que*

(a) $$[\alpha]_{\mathcal{B}} = P[\alpha]_{\mathcal{B}'}$$
(b) $$[\alpha]_{\mathcal{B}'} = P^{-1}[\alpha]_{\mathcal{B}}$$

para todo vetor α *em* V.

Para completar a análise acima demonstraremos também o resultado que segue.

Teorema 8. *Suponhamos que* P *seja uma* n $\times$ n *matriz inversível sôbre* F. *Seja* V *um espaço vetorial* n*-dimensional sôbre* F *e seja* $\mathcal{B}$ *uma base ordenada de* V. *Então, existe uma única base ordenada* $\mathcal{B}'$ *de* V *tal que*

(a) $$[\alpha]_{\mathcal{B}} = P[\alpha]_{\mathcal{B}'}$$
(b) $$[\alpha]_{\mathcal{B}'} = P^{-1}[\alpha]_{\mathcal{B}}$$

para todo vetor α *em* V.

Demonstração. Seja $\mathcal{B}$ constituída pelos vetores $\alpha_1, \ldots, \alpha_n$. Se $\mathcal{B}' = \{\alpha_1', \ldots, \alpha_n'\}$ é uma base ordenada de V para a qual (a) é válida, é claro que

$$\alpha_j' = \sum_{i=1}^{n} P_{ij}\alpha_i.$$

Assim, basta-nos mostrar que os vetores α_j', definidos por estas equações, formam uma base. Seja $Q = P^{-1}$. Então

$$\begin{aligned}
\sum_j Q_{jk}\alpha_j' &= \sum_j Q_{jk} \sum_i P_{ij}\alpha_i \\
&= \sum_j \Big(\sum_i P_{ij}Q_{jk}\Big)\alpha_i \\
&= \sum_i \Big(\sum_j P_{ij}Q_{jk}\Big)\alpha_i \\
&= \alpha_k.
\end{aligned}$$

Portanto, o subespaço gerado pelo conjunto

$$\mathcal{B}' = \{\alpha'_1, \ldots, \alpha'_n\}$$

contém $\mathcal{B}$, logo é igual a V. Assim, $\mathcal{B}'$ é uma base e, de sua definição e do Teorema 7, é evidente que (a) é válida, logo (b) também o é.

Exemplo 16. Seja F um corpo e seja

$$\alpha = (x_1, x_2, \ldots, x_n)$$

um vetor de F^n. Se $\mathcal{B}$ é a base ordenada canônica de F^n,

$$\mathcal{B} = \{\epsilon_1, \ldots, \epsilon_n\},$$

a matriz das coordenadas do vetor α em relação à base $\mathcal{B}$ é dada por

$$[\alpha]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}.$$

Exemplo 17. Seja R o corpo dos números reais e seja θ um número real fixo. A matriz

$$P = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

é inversível e sua inversa é

$$P^{-1} = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}.$$

Portanto, para cada θ, o conjunto $\mathcal{B}'$ constituído pelos vetores ($\cos\theta$, $\operatorname{sen}\theta$), ($-\operatorname{sen}\theta$, $\cos\theta$) é uma base de R^2; intuitivamente esta base pode ser descrita como sendo a base obtida pela rotação de um ângulo θ da base canônica. Se α é o vetor (x_1, x_2), então

$$[\alpha]_{\mathcal{B}'} = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

ou

$$\begin{aligned} x'_1 &= x_1 \cos\theta + x_2 \operatorname{sen}\theta \\ x'_2 &= -x_1 \operatorname{sen}\theta + x_2 \cos\theta. \end{aligned}$$

Exemplo 18. Seja F um subcorpo do corpo dos números complexos. A matriz

$$P = \begin{bmatrix} -1 & 4 & 5 \\ 0 & 2 & -3 \\ 0 & 0 & 8 \end{bmatrix}$$

é inversível e sua inversa é

$$P^{-1} = \begin{bmatrix} -1 & 2 & \frac{11}{8} \\ 0 & \frac{1}{2} & \frac{3}{16} \\ 0 & 0 & \frac{1}{8} \end{bmatrix}.$$

Portanto, os vetores

$$\begin{aligned} \alpha_1' &= (-1, \ \ 0, \ 0) \\ \alpha_2' &= (\ \ 4, \ \ 2, \ 0) \\ \alpha_3' &= (\ \ 5, -3, \ 8) \end{aligned}$$

formam uma base $\mathcal{B}'$ de F^3. As coordenadas x_1', x_2', x_3' do vetor $\alpha = (x_1, x_2, x_3)$ em relação à base $\mathcal{B}'$ são dados por

$$\begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix} = \begin{bmatrix} -x_1 + 2x_2 + \frac{11}{8}x_3 \\ \frac{1}{2}x_1 + \frac{3}{16}x_3 \\ \frac{1}{8}x_3 \end{bmatrix} = \begin{bmatrix} -1 & 2 & \frac{11}{8} \\ 0 & \frac{1}{2} & \frac{3}{16} \\ 0 & 0 & \frac{1}{8} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Em particular,

$$(3, 2, -8) = -10\alpha_1' - \tfrac{1}{2}\alpha^1 - \alpha^3.$$

Exercícios

1. Mostrar que os vetores

$$\begin{aligned} \alpha_1 &= (1, 1, 0, 0), & \alpha_2 &= (0, 0, 1, 1) \\ \alpha_3 &= (1, 0, 0, 4), & \alpha_4 &= (0, 0, 0, 2) \end{aligned}$$

formam uma base de R_4. Determinar as coordenadas de cada um dos vetores da base canônica em relação à base ordenada $\{x_1, \alpha_2, \alpha_3, \alpha_4\}$.

2. Determinar a matriz das coordenadas do vetor (1, 0, 1) em relação à base de C_3 constituída pelos vetores $(2i, 1, 0)$, $(2, -1, 1)$, $(0, 1 + i\ -i)$, nesta ordem.

3. Seja $\mathcal{B} = \{\alpha_1, \alpha_2, \alpha_3\}$ a base ordenada de R_3 constituída por

$$\alpha_1 = (1, 0, -1), \qquad \alpha_2 = (1, 1, 1), \qquad \alpha_3 = (1, 0, 0).$$

Quais são as coordenadas do vetor (a, b, c) em relação à base ordenada $\mathcal{B}$?

4. Seja W o subespaço de C_3 gerado por $\alpha_1 = (1, 0, i)$ e $\alpha_2 = (1 + i, 1, -1)$.

(a) Mostrar que α_1 e α_2 formam uma base de W.

(b) Mostrar que os vetores $\beta_1 = (1, 1, 0)$ e $\beta_2 = (1, i, 1 + i)$ estão em W e formam outra base de W.

(c) Quais são as coordenadas de α_1 e α_2 em relação à base ordenada $\{\beta_1, \beta_2\}$ de W?

5. Sejam $\alpha = (x_1, x_2)$ e $\beta = (y_1, y_2)$ vetores de R^2 tais que

$$x_1y_1 + x_2y_2 = 0, \qquad x_1^2 + x_2^2 = y_1^2 + y_2^2 = 1.$$

Demonstrar que $\mathcal{B} = \{\alpha, \beta\}$ é uma base de R^2. Determinar as coordenadas do vetor (a, b) em relação à base ordenada $\mathcal{B} = \{\alpha, \beta\}$. (As condições sôbre α e β dizem, geomètricamente, que α e β são perpendiculares e cada um tem comprimento 1.)

6. Seja V o espaço vetorial sôbre o corpo dos números complexos das funções de R em C, isto é, o espaço das funções definidas sôbre a reta real e tomando valores complexos. Sejam $f_1(x) = 1$, $f_2(x) = e^{ix}$, $f_3(x) = e^{-ix}$.

(a) Demonstrar que f_1, f_2 e f_3 são linearmente independentes.

(b) Sejam $g_1(x) = 1$, $g_2(x) = \cos g_3(x) = \operatorname{sen} x$. Determinar uma 3×3 matriz P inversível tal que

$$g_j = \sum_{i=1}^{3} P_{ij}f_i.$$

7. Seja V o espaço vetorial (real) das funções polinomiais de R em R de grau menor ou igual a 2, isto é, o espaço das funções f da forma

$$f(x) = c_0 + c_1x + c_2x^2.$$

Seja t um número fixo e definamos

$$g_1(x) = 1,\ g_2(x) = x + t,\ g_3(x) = (x + t)^2.$$

Demonstrar que $\mathcal{B} = \{g_1, g_2, g_3\}$ é uma base de V. Se

$$f(x) = c_0 + c_1x + c_1x^2$$

quais são as coordenadas de f em relação a esta base ordenada β?

2.5 Resumo de Linha-equivalência

Nesta seção utilizaremos alguns fatos elementares sôbre bases e dimensão de espaços vetoriais de dimensão finita para completar nossa discussão de linha-equivalência de matrizes. Lembramos que se A é uma $m \times n$ matriz sôbre o corpo F, os vetores-linhas de A são os vetores $\alpha_1, \ldots, \alpha_m$ em F^n definidos por

$$\alpha_i = (A_{i1}, \ldots, A_{in})$$

e que o espaço-linha de A é o subespaço de F^n gerado por êstes vetores. O **pôsto-linha** de A é a dimensão do espaço-linha de A.

Se P é uma $k \times m$ matriz sôbre F, então o produto $B = PA$ é uma $k \times m$ matriz cujos vetores-linhas $\beta_1, \ldots, \beta_k$ são combinações lineares

$$\beta_i = P_{i1}\alpha_1 + \ldots + P_{im}\alpha_m$$

dos vetores-linhas de A. Portanto, o espaço-linha de B é um subespaço do espaço-linha de A. Se P é uma $m \times m$ matriz inversível, então B é linha-equivalente a A de modo que a simetria da linha-equivalência, ou a equação $A = P^{-1}B$, implica que o espaço-linha de A também é um subespaço do espaço-linha de B.

Teorema 9. *Matrizes linha-equivalentes possuem o mesmo espaço-linha.*

Vemos assim que para estudar o espaço-linha de A podemos estudar o espaço-linha de uma matriz linha-reduzida à forma em escada que seja linha-equivalente a A. É o que passamos a fazer.

Teorema 10. *Seja* R *uma matriz não-nula linha-reduzida à forma em escada. Então os vetores-linhas não-nulos de* R *formam uma base do espaço-linha de* R.

Demonstração. Sejam $\rho_1, \ldots, \rho_r$ os vetores-linhas não-nulos de R:

$$\rho_i = (R_{i1}, \ldots, R_{in}).$$

Êsses vetores certamente geram o espaço-linha de R; precisamos apenas demonstrar que êles são linearmente independentes. Como R é uma matriz linha-reduzida à forma em escada, existem inteiros positivos $k_1, \ldots, k_r$ tais que, para $i \leq r$, tem-se

$$\text{(2—18)} \qquad \begin{array}{ll} \text{(a)} & R(i,j) = 0, \quad \text{se} \quad j < k_i \\ \text{(b)} & R(i, k_j) = \delta_{ij} \\ \text{(c)} & k_i < \ldots < k_r. \end{array}$$

Suponhamos que $\beta = (b_1 \ldots, b_n)$ seja um vetor do espaço-linha de R:

$$\text{(2—19)} \qquad \beta = c_1\rho_1 + \ldots + c_r\rho_r.$$

Afirmamos então que $c_j = b_{k_j}$. De fato, por (2—18)

$$\begin{aligned} b_{k_j} &= \sum_{i=1}^{r} c_i R(i,k_j) \\ &= \sum_{i=1}^{r} c_i \delta_{ij} \\ \text{(2—20)} \qquad &= c_j. \end{aligned}$$

Em particular, se $\beta = 0$, isto é, se $c_1\rho_1 + \ldots + c_r\rho_r = 0$, então c_j é necessàriamente a k_j-ésima coordenada do vetor nulo, de modo que $c_j = 0$ para $j = 1, \ldots, r$. Assim $\rho_1, \ldots, \rho_r$ são linearmente independentes.

Teorema 11. *Sejam* m *e* n *inteiros positivos e seja* F *um corpo. Suponhamos que* W *seja um subespaço de* F^n *e que dim* W $\leq$ m. *Então, existe exatamente uma* m $\times$ n *matriz sôbre* F, *linha-reduzida à forma em escada, cujo espaço-linha é* W.

Demonstração. Existe pelo menos uma $m \times n$ matriz linha-reduzida à forma em escada cujo espaço-linha é W. Como dim $W \leq m$, podemos tomar m vetores $\alpha_1, \ldots, \alpha_m$ em W que geram W. Seja A a $m \times n$ matriz com vetores-linhas $\alpha_1, \ldots, \alpha_m$ e seja R uma matriz linha-reduzida à forma em escada e linha-equivalente a A. Então, o espaço-linha de R é W.

Seja agora R uma qualquer $m \times n$ matriz linha-reduzida à forma em escada e com espaço-linha W. Mostraremos que R é determinada de modo único pelo subespaço W. A descrição de R em têrmos de W será feita como segue. Consideremos todos os vetores $\beta = (b_1, \ldots, b_n)$ em W. Se $\beta \neq 0$, então a primeira coordenada não-nula de β deve ocorrer em uma certa coluna t:

$$\beta = (0, \ldots, 0, b_t, \ldots, b_n), \qquad b_t \neq 0$$

Sejam $k_1, \ldots, k_r$ os inteiros positivos t tais que exista algum $\beta \neq 0$ em W, cuja primeira coordenada não-nula ocorra na coluna t. Coloquemos os $k_1, \ldots, k_r$ na ordem $k_1 < k_2 < \ldots < k_r$. Para cada inteiro positivo k_s existirá um e sòmente um vetor ρ_s em W tal que a k_s-ésima coordenada de ρ_s seja 1 e a k_i-ésima coordenada de ρ_s seja 0 para $i \neq s$. Então, R é a $m \times n$ matriz cujos vetores-linhas são $\rho_1, \ldots, \rho_r, 0, \ldots, 0$.

Tendo indicado como R será determinada a partir de W, procedemos como segue. Consideremos uma $m \times n$ matriz R arbitrária, linha-reduzida à forma em escada e com espaço-linha W. Se $k_1, \ldots, k_r$ são as colunas distinguidas de R (2—18) e $\rho_1, \ldots, \rho_r$, são os vetores-linhas não-nulos de R, demonstraremos que $k_1, \ldots \rho_r$, são exatamente como foram descritos no último parágrafo. Isto mostrará que R é a única $m \times n$ matriz linha-reduzida à forma em escada cujo espaço-linha é W.

Pelo Teorema 10, os vetores-linhas não-nulos $\rho_1, \ldots, \rho_r$ formam uma base de W. Na demonstração do Teorema 10 observamos que se $\beta = (b_1, \ldots, b_n)$ está em W.

$$\beta = c_1\rho_1 + \ldots + c_r\rho_r,$$

então $c_i = b_{k_i}$, isto é

$$\beta = \sum_{i=1}^{r} b_{k_i}\rho_i. \tag{2—21}$$

Assim, todo vetor β está determinado se se conhecem as coordenadas b_{k_i}, $i = 1, \ldots, r$. Por exemplo, ρ_s é o único vetor em W cuja k_s-ésima coordenada é 1 e cuja k_i-ésima é 0 para $i \neq s$.

Suponhamos que β esteja em W e $\beta \neq 0$. Afirmamos que a primeira coordenada não-nula de β ocorre em uma das colunas k_s. Como

$$\beta = \sum_{i=1}^{r} b_{k_i}\rho_i$$

e $\beta \neq 0$, podemos escrever

$$\beta = \sum_{j=s}^{r} b_{k_i}\rho_i, \qquad b_{k_s} \neq 0. \tag{2—22}$$

Das condições (2—18) tem-se que $R_{ij} = 0$ se $i > s$ e $j \leq k_s$. Portanto,

$$\beta = (0, \ldots, 0, b_{k_s}, \ldots b_n), \; b_{k_s} \neq 0$$

e a primeira coordenada não-nula de β ocorre na coluna k_s.

Notemos que para cada k_s, $s = 1, \ldots, r$, existe um vetor em W cuja primeira coordenada não-nula ocorre na coluna k_s, a saber, o vetor ρ_s. Demonstramos assim que $k_1, \ldots, k_r$ são exatamente os inteiros t tais que algum vetor não-nulo em W possui sua primeira coordenada não-nula na coluna t. Já ressaltamos o fato de que ρ_s é o único vetor em W cuja k_i-ésima coordenada é δ_{is}.

Corolário. *Cada* m $\times$ n *matriz* A *é linha-equivalente a exatamente uma matriz linha-reduzida à forma em escada.*

Demonstração. Sabemos que A é linha-equivalente a pelo menos uma matriz R linha-reduzida à forma em escada. Se A é linha-equivalente a uma outra tal matriz R', então R é linha-equivalente a R'; logo, R e R' possuem o mesmo espaço-linha e são necessàriamente idênticas.

Corolário. *Sejam* A *e* B m $\times$ n *matrizes sôbre o corpo* F. *Então* A *e* B *são linha-equivalentes se, e sòmente se, possuem o mesmo espaço-linha.*

Demonstração. Sabemos que se A e B são linha-equivalentes, então possuem o mesmo espaço-linha. Suponhamos então que A e B possuam o mesmo espaço-linha. Ora, A é linha-equivalente a uma matriz R linha-reduzida à forma em escada e B é linha-equivalente a uma matriz R' linha-reduzida à forma em escada. Como A e B

têm o mesmo espaço-linha, R e R' têm o mesmo espaço-linha. Portanto, $R = R'$ e A é linha-equivalente a B.

Resumindo — se A e B são $m \times n$ matrizes sôbre o corpo **F**, as seguintes afirmações são equivalentes:

(i) A e B são linha-equivalentes.
(ii) A e B possuem o mesmo espaço-linha.
(iii) $B = PA$, sendo P uma $m \times n$ matriz inversível.

Uma quarta afirmação equivalente é que os sistemas homogêneos $AX = 0$ e $BX = 0$ têm as mesmas soluções; contudo, apêsar de sabermos que a linha-equivalência de A e B implica que êsses sistemas têm as mesmas soluções, o melhor parece ser deixar a demonstração da recíproca para mais tarde.

2.6 Cálculos Concernentes a Subespaços

Gostaríamos agora de mostrar como as operações elementares fornecem um método padronizado de responder a certas perguntas concretas concernentes a subespaços de F^n. Já deduzimos os fatos de que precisaremos. Êles são aqui reunidos para conveniência do leitor. A discussão aplica-se a qualquer espaço vetorial n-dimensional sôbre o corpo F se se toma uma base ordenada $\mathfrak{B}$, fixa, e se descreve cada vetor α de V pela n-upla $(x_1, \ldots, x_n)$ que dá as coordenadas de α em relação à base ordenada $\mathfrak{B}$.

Suponhamos que nos sejam dados m vetores $\alpha_1, \ldots, \alpha_m$ em F^n. Consideremos as seguintes perguntas:

(1) Como se pode saber se os vetores $\alpha_1, \ldots, \alpha_m$ são linearmente dependentes? De maneira mais geral, como se determina a dimensão do subespaço W gerado por êstes vetores?

(2) Dado β em F^n, como se pode saber se β é uma combinação linear de $\alpha_1, \ldots, \alpha_m$, isto é, se β está no subespaço W?

(3) Como se pode fazer uma descrição explícita do subespaço W?

A terceira pergunta é um tanto vaga, pois não especifica o que se quer dizer com uma "descrição explícita"; no entanto, esclareceremos êste ponto fazendo o tipo de descrição que temos em mente. Com esta descrição, as perguntas (1) e (2) podem ser respondidas imediatamente.

Seja A a $m \times n$ matriz com vetores-linhas α_i:

$$\alpha_i = (A_{i1}, \ldots, A_{in}).$$

Efetuemos uma seqüência de operações elementares sôbre linhas, começando com A e terminando com uma matriz R linha-reduzida à forma em escada. Já explicamos anteriormente como fazer isto. Neste ponto, a dimensão de W (o espaço-linha de A) é evidente, pois esta dimensão é simplesmente o número de vetores-linhas não-nulos de R. Se $\rho_1, \ldots, \rho_r$ são os vetores-linhas não-nulos de R, então $\mathcal{B} = \{\rho_1, \ldots, \rho_r\}$ é uma base de W. Se a primeira coordenada não-nula de ρ_i é a k_i-ésima, então temos, para $i \leq r$,

(a) $$R(i, j) = 0, \quad \text{se} \quad j < k_i$$

(b) $$R(i, k_j) = \delta_{ij}$$

(c) $$k_1 < \ldots < k_r.$$

O subespaço W consiste de todos os vetores

$$\begin{aligned}\beta &= c_1\rho_1 + \ldots + c_r\rho_r \\ &= \sum_{i=1}^{r} c_i(R_{i1}, \ldots, R_{in}).\end{aligned}$$

As coordenadas $b_1, \ldots, b_n$ de um tal vetor β são então

$$b_j = \sum_{i=1}^{r} c_i R_{ij}. \tag{2-23}$$

Em particular, $b_{k_j} = c_j$, e, se $\beta = (b_1,\ldots,b_n)$ é uma combinação linear dos ρ_i, então tem de ser a particular combinação linear

$$\beta = \sum_{i=1}^{r} b_{k_i}\rho_i. \tag{2-24}$$

As condições sôbre β para que (2-24) valha são:

$$b_j = \sum_{i=1}^{r} b_{k_i}R_{ij}, \qquad j = 1, \ldots, n. \tag{2-25}$$

Ora, (2-25) é a descrição explicíta do subespaço W gerado por $\alpha_1, \ldots, \alpha_m$, isto é, o subespaço consiste dos vetores β em F^n cujas coordenadas satisfazem (2-25). Que tipo de descrição é (2-25)? Em primeiro lugar, ela descreve W como o conjunto das soluções $\beta = (b_1, \ldots, b_n)$ do sistema (2-25) de equações lineares homogêneas. Êste sistema de equações é de uma natureza muito particular, pois $(n - r)$ das coordenadas como combinações lineares das r coordenadas distinguidas $b_{k_1}, \ldots, b_{k_r}$. Tem-se completa liberdade de escolha das coordenadas b_{k_i}, isto é, se $c_1, \ldots, c_r$ são r escalares *arbitrários*, existe um e sòmente um vetor β em W cuja k_i-ésima coordenada é c_i.

O ponto importante aqui é o seguinte: Dados os vetores α_i, a linha-redução é um método direto de determinar os inteiros r,

$k_1, \ldots, k_r$ e os escalares R_{ij} que dão a descrição (2-25) do subespaço gerado por $\alpha_1, \ldots, \alpha_m$. Deve-se observar, como no Teorema 11, que todo subespaço W de F^n possui uma descrição do tipo (2-25). Devemos destacar também algumas coisas sôbre a pergunta (2). Já mostramos na seção 1.4 como se pode encontrar uma $m \times m$ matriz inversível P tal que $R = PA$. O conhecimento de P nos permite determinar os escalares $x_1, \ldots, x_m$ tais que

$$\beta = x_1\alpha_1 + \ldots + x_m\alpha_m$$

quando isto é possível. Como os vetores-linhas de R são dados por

$$\rho_i = \sum_{j=1}^{m} P_{ij}\alpha_j$$

temos que se β é uma combinação linear dos α_j, então

$$\begin{aligned}\beta &= \sum_{i=1}^{r} b_{k_i}\rho_i \\ &= \sum_{i=1}^{r} b_{k_i} \sum_{j=1}^{m} P_{ij}\alpha_j \\ &= \sum_{j=1}^{m} \sum_{i=1}^{r} b_{k_i}P_{ij}\alpha_j\end{aligned}$$

e portanto

$$x_j = \sum_{i=1}^{r} b_{k_i}P_{ij}$$

é uma escolha possível para os x_j (podem existir muitas).

O problema de saber se $\beta = (b_1, \ldots, b_n)$ é uma combinação linear dos α_i, e, em caso afirmativo, quais são os escalares x_i, pode também ser considerado perguntando-se se o sistema de equações

$$\sum_{i=1}^{m} A_{ij}x_i = b_j, \qquad j = 1, \ldots, n$$

admite soluções e quais são elas. A matriz dos coeficientes dêste sistema de equações é a $n \times m$ matriz B com vetores-*colunas* $\alpha_1, \ldots, \alpha_m$. No Capítulo 1 discutimos o uso de operações elementares sôbre linhas na resolução de um sistema de equações $BX = Y$. Consideremos um exemplo no qual adotamos êste último ponto de vista ao respondermos a perguntas sôbre subespaços de F^n.

Exemplo 19. Proponhamos o seguinte problema: seja W o subespaço de R^4 gerado pelos vetores

$$\begin{aligned} \alpha_1 &= (1, 2, 2, 1) \\ \alpha_2 &= (0, 2, 0, 1) \\ \alpha_3 &= (-2, 0, -4, 3). \end{aligned}$$

(a) Demonstrar que α_1, α_2, α_3 formam uma base de W, isto é, que êstes vetores são linearmente independentes.

(b) Seja $\beta = (b_1, b_2, b_3, b_4)$ um vetor em W. Quais são as coordenadas de β em relação à base ordenada $\{\alpha_1, \alpha_2, \alpha_3\}$?

(c) Sejam

$$\begin{aligned} \alpha'_1 &= (1, 0, 2, 0) \\ \alpha'_2 &= (0, 2, 0, 1) \\ \alpha'_3 &= (0, 0, 0, 3). \end{aligned}$$

Mostrar que α'_1, α'_2, α'_3 formam uma base de W.

(d) Se β está em W, indiquemos por X a matriz das coordenadas de β em relação à base dos α e por X' a matriz das coordenadas de β em relação à base dos α'. Determinar a 3×3 matriz P tal que $X = PX'$ para todo β em W.

Para responder a estas diversas perguntas, formemos a 4×3 matriz B com vetores-colunas α_1, α_2, α_3:

$$B = \begin{bmatrix} 1 & 0 & -2 \\ 2 & 2 & 0 \\ 2 & 0 & -4 \\ 1 & 1 & 3 \end{bmatrix}.$$

Perguntamos para quais y_1, y_2, y_3, y_4 o sistema $BX = Y$ admite solução.

$$\begin{bmatrix} 1 & 0 & -2 & y_1 \\ 2 & 2 & 0 & y_2 \\ 2 & 0 & -4 & y_3 \\ 1 & 1 & 3 & y_4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -2 & y_1 \\ 0 & 2 & 4 & y_2 - 2y_1 \\ 0 & 0 & 0 & y_3 - 2y_1 \\ 0 & 1 & 5 & y_4 - y_1 \end{bmatrix} \rightarrow$$

$$\begin{bmatrix} 1 & 0 & -2 & y_1 \\ 0 & 0 & -6 & y_2 - 2y_4 \\ 0 & 1 & 5 & y_4 - y_1 \\ 0 & 0 & 0 & y_3 - 2y_1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & y_1 - \frac{1}{3}y_2 + \frac{2}{3}y_4 \\ 0 & 0 & 1 & \frac{1}{6}(2y_4 - y_2) \\ 0 & 1 & 0 & -y_1 + \frac{5}{6}y_2 - \frac{2}{3}y_4 \\ 0 & 0 & 0 & y_3 - 2y_1 \end{bmatrix}$$

Portanto, a condição para que o sistema $BX = Y$ admita uma solução é $y_3 = 2y_1$. Então, $\beta = (b_1, b_2, b_3, b_4)$ está em W se, e sòmente se, $b_3 = 2b_1$. Se β está em W, então $\beta = x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3$ onde necessàriamente,

$$\begin{aligned} x_1 &= b_1 - \tfrac{1}{3}b_2 + \tfrac{2}{3}b_4 \\ x_2 &= -b_1 + \tfrac{5}{6}b_2 - \tfrac{2}{3}b_4 \\ x_3 &= - \tfrac{1}{6}b_2 + \tfrac{1}{3}b_4 \end{aligned} \tag{2-26}$$

(a) As equações acima mostram certamente que α_1, α_2, α_3 são linearmente independentes.

(b) Se β está em W, as coordenadas de β em relação à base ordenada $\{\alpha_1, \alpha_2, \alpha_3\}$ são os escalares x_1, x_2, x_3 dados por (2-26).

(c) Os vetores dados α'_1 α'_2, α'_3 satisfazem $y_3 = 2y_1$, portanto estão em W. Deveria ser óbvio para o leitor, à primeira vista, que êsses vetores são linearmente independentes.

(d) A matriz P é determinada exprimindo-se α'_1, α'_2, α'_3 como combinações lineares dos α_i. As equações (2-26) nos dizem como fazer isto. Por exemplo, com $\beta = \alpha'_1$ temos $b_1 = 1$, $b_2 = 0$, $b_3 = 2$, $b_4 = 0$, e

$$\begin{aligned} x_1 &= 1 - \tfrac{1}{3}(0) + \tfrac{2}{3}(0) = 1 \\ x_2 &= -1 + \tfrac{5}{6}(0) - \tfrac{2}{3}(0) = -1 \\ x_3 &= - \tfrac{1}{6}(0) + \tfrac{1}{3}(0) = 0. \end{aligned}$$

Assim, $\alpha'_1 = \alpha_1 - \alpha_2$. Anàlogamente, obtemos $\alpha'_2 = \alpha_2$ e $\alpha'_3 = 2\alpha_1 - 2\alpha_2 + \alpha_3$. Logo

$$P = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Exercícios

1. Responder as perguntas do Exemplo 19 usando operações elementares sôbre as linhas da 3×4 matriz de vetores-linhas α_i.

2. Sejam

$$\alpha_1 = (1, 1, -2, 1), \quad \alpha_2 = (3, 0, 4, -1), \quad \alpha_3 = (-1, 2, 5, 2).$$

Sejam

$$\alpha = (4, -5, 9, -7), \quad \beta = (3, 1, -4, 4), \quad \gamma = (-1, 1, 0, 1).$$

(a) Quais dos vetores α, β, γ estão no subespaço de R^4 gerado pelos α_i?
(b) Quais dos vetores α, β, γ estão no subespaço de C^4 gerado pelos α_i?
(c) Isto sugere algum teorema?

3. Consideremos os vetores em R^4 definidos por

$$\alpha_1 = (-1, 0, 1, 2), \quad \alpha_2 = (3, 4, -2, 5), \quad \alpha_3 = (1, 4, 0, 9).$$

Determinar um sistema de equações lineares homogêneas para o qual o espaço das soluções seja exatamente o subespaço de R^4 gerado pelos três vetores dados.

4. Em C^3 sejam

$$\alpha_1 = (1, 0, -i), \quad \alpha_2 = (1 + i, 1 - i, 1), \quad \alpha_3 = (i, i, i).$$

Demonstrar que êstes vetores formam uma base de C^3. Quais são as coordenadas do vetor (a, b, c) em relação a esta base?

5. Fazer uma descrição explícita do tipo (2-25) para os vetores

$$\beta = (b_1, b_2, b_3, b_4, b_5)$$

em R^5 que são combinações lineares dos vetores

$$\alpha_1 = (1, 0, 2, 1, -1), \quad \alpha_2 = (-1, 2, -4, 2, 0),$$
$$\alpha_3 = (2, -1, 5, 2, 1), \text{ e } \alpha_4 = (2, 1, 3, 5, 2).$$

6. Usar a nossa discussão de linha-redução para demonstrar o que segue. Se A é uma $m \times n$ matriz sôbre o corpo F com pôsto r, então a dimensão do espaço-solução de A é $(n - r)$. (*Sugestão*: bastará demonstrar êste fato para uma matriz R linha-reduzida à forma em escada. Para uma tal R, o sistema $RX = 0$ exprime r das coordenadas $x_{k_1}, \ldots, x_{k_r}$ em função das $(n - r)$ coordenadas restantes $u_1, \ldots, u_{n-r}$. Construir a solução X_j obtida atribuindo-se o valor 1 a u_j e 0 aos outros u_i. Demonstrar que $X_1, \ldots, X_{n-r}$; formam uma base do espaço-solução.)

7. Seja A uma $m \times n$ matriz sôbre o corpo F e consideremos o sistema de equações $AX = Y$. Demonstrar que êste sistema de equações admite uma solução se, e sòmente se, o pôsto-linha de A é igual ao pôsto-linha da matriz completa do sistema.

CAPÍTULO 3

TRANSFORMAÇÕES LINEARES

3.1 Transformações Lineares

Introduziremos agora as transformações lineares, objetos que estudaremos na maior parte do restante dêste livro. O leitor poderá achar útil ler (ou reler) a discussão sôbre funções no Apêndice, visto que usaremos livremente a terminologia daquela discussão.

Definição. *Sejam* V *e* W *espaços vetoriais sôbre o corpo* F. *Uma* **transformação linear** *de* V *em* W *é uma função* T *de* V *em* W *tal que*

$$T(c\alpha+\beta) = c(T\alpha) + T\beta$$

para todos α e β em V *e todos escalares* c *em F.*

Exemplo 1. Se V é um espaço vetorial arbitrário, a transformação idêntica I, definida por $I\alpha = \alpha$, é uma transformação linear de V em V. A **transformação nula** 0, definida por $0\alpha = 0$, é uma transformação linear de V em V.

Exemplo 2. Seja F um corpo e seja V o espaço das funções polinomiais f de F em F, dadas por

$$f(x) = c_0 + c_1x + \ldots + c_kx^k.$$

Seja

$$(Df)\,(x) = c_1 + 2c_2x + \ldots + kc_kx^{k-1}.$$

Então, D é uma transformação linear de V em V — a transformação derivação.

Exemplo 3. Seja V o espaço das $n \times 1$ matrizes - (colunas) sôbre um corpo F e seja A uma $m \times n$ matriz fixa sôbre F. A função T sôbre V, definida por $T(X) = AX$, é uma transformação linear de V no espaço W das $m \times 1$ matrizes sôbre F.

Exemplo 4. Seja V o espaço das $m \times n$ matrizes sôbre um corpo F, seja P uma $m \times m$ matriz fixa sôbre F e seja Q uma $n \times n$ matriz fixa sôbre F. Definamos uma função T de V em V por $T(A) = PAQ$. Então T é uma transformação linear de V em V, pois

$$\begin{aligned} T(cA + B) &= P(cA + B)\,Q \\ &= (cPA + PB)Q \\ &= cPAQ + PBQ \\ &= cT(A) + T(B). \end{aligned}$$

Exemplo 5. Seja R o corpo dos números reais e seja V o espaço das funções de R em R que são *contínuas*. Definamos T por

$$(Tf)(x) = \int_0^x f(t)\,dt.$$

Então T é uma transformação linear de V em V. A função Tf, além de contínua, possui a primeira derivada contínua. A linearidade da integração é uma de suas propriedades fundamentais.

O leitor não deverá encontrar nenhuma dificuldade para verificar que as transformações definidas nos Exemplos 1, 2, 3 e 5 são transformações lineares. Ampliaremos consideràvelmente nossa lista de exemplos à medida que aprendermos mais sôbre transformações lineares.

É importante notar que se T é uma transformação linear de V em W, então $T(0) = 0$; pode-se ver isto a partir da definição, pois

$$T(0) = T(0 + 0) = T(0) + T(0).$$

Êste ponto freqüentemente causa confusão à pessoa que estuda álgebra linear pela primeira vez, desde que ela provàvelmente tomou contato com uma utilização ligeiramente diferente do têrmo "função linear". Um comentário breve deverá eliminar a confusão. Suponhamos que V seja o espaço vetorial R^1. Uma transformação linear de V em V é então um tipo especial de função com valores reais definida sôbre a reta real R. Num curso de cálculo, ter-se-ia provàvelmente a denominação linear para uma tal função se seu gráfico fôsse uma reta. Uma transformação linear de R^1 em R^1, de acôrdo com nossa definição, será uma função de R em R, cujo gráfico é uma reta *que passa pela origem*.

Além da propriedade $T(0) = 0$, destaquemos outra propriedade de uma transformação linear arbitrária T. Uma tal transformação "conserva" combinações lineares; isto é, se $\alpha_1, \ldots, \alpha_n$ são vetores em V e $c_1, \ldots, c_n$ são escalares, então

$$T(c_1\alpha_1 + \ldots + c_n\alpha_n) = c_1(T\alpha_1) + \ldots + c_n(T\alpha_n).$$

Isto decorre imediatamente da definição. Por exemplo,

$$\begin{aligned} T(c_1\alpha_1 + c_2\alpha_2) &= c_1(T\alpha_1) + T(c_2\alpha_2) \\ &= c_1(T\alpha_1) + c_2(T\alpha_2). \end{aligned}$$

Teorema 1. *Seja* V *um espaço vetorial de dimensão finita sôbre o corpo* F *e seja* $\{\alpha_1, \ldots, \alpha_n\}$ *uma base ordenada de* V. *Seja* W *um espaço vetorial sôbre o mesmo corpo* F *e sejam* $\beta_1, \ldots, \beta_n$ *vetores arbitrários em* W. *Então, existe exatamente uma transformação linear* T *de* V *em* W *tal que*

$$T\alpha_j = \beta_j, \quad j = 1, \ldots, n.$$

Demonstração. Para demonstrar que existe pelo menos uma transformação linear T com $T\alpha_j = \beta_j$ procedemos como segue. Dado α em V, existe uma única n-upla $(x_1, \ldots, x_n)$ tal que

$$\alpha = x_1\alpha_1 + \ldots + x_n\alpha_n.$$

Para êste vetor α, definamos

$$T\alpha = x_1\beta_1 + \ldots + x_n\beta_n.$$

Então, T é uma regra bem definida para se associar a cada vetor α em V um vetor $T\alpha$ em W. Pela definição, é evidente que $T\alpha_j = \beta_j$ para todo j. Para ver que T é linear, seja

$$\beta = y_1\alpha_1 + \ldots + y_n\alpha_n$$

em V e c escalar arbitrário. Ora,

$$c\,\alpha + \beta = (cx_1 + y_1)\alpha_1 + \ldots + (cx_n + y_n)\alpha_n$$

portanto, pela definição,

$$T(c\alpha + \beta) = (cx_1 + y_1)\beta_1 + \ldots + (cx_n + y_n)\beta_n.$$

Por outro lado,

$$\begin{aligned} c(T\alpha) + T\beta &= c\sum_{i=1}^{n} x_i\beta_i + \sum_{i=1}^{n} y_i\beta_i \\ &= \sum_{i=1}^{n} (cx_i + y_i)\beta_i \end{aligned}$$

e assim

$$T(c\alpha + \beta) = c(T\alpha) + T\beta.$$

Se U é uma transformação linear de V em W com $U\alpha_j = \beta_j$, $j = 1, \ldots, n$, então, para o vetor $\alpha = \sum_{i=1}^{n} x_i\alpha_i$, temos

$$\begin{aligned} U\alpha &= U\left(\sum_{i=1}^{n} x_i\alpha_i\right) \\ &= \sum_{i=1}^{n} x_i(U\alpha_i) \\ &= \sum_{i=1}^{n} x_i\beta_i \end{aligned}$$

de modo que U é exatamente a regra T que definimos acima. Isto mostra que a transformação linear T com $T\alpha_j = \beta_j$ é única.

Teorema 2. *Sejam* V *e* W *espaços vetoriais sôbre o corpo* F *e seja* T *uma transformação linear de* V *em* W. *Então, a imagem de* T *é um subespaço de* W. *O conjunto de vetores* α *em* V *tais que* $T\alpha = 0$ *também é um subespaço de* V, *dito o* **núcleo** (*ou espaço nulo*) *de* T.

Demonstração. Indiquemos por R_T a imagem de T, isto é, o conjunto dos vetores β em W tais que $\beta = T\alpha$ para algum α em V. Sejam β_1 e β_2 em R_T e seja c em F. Ora, existem vetores α_1 e α_2 em V tais que $T\alpha_i = \beta_i$, $i = 1, 2$. Como T é linear

$$T(c\alpha_1 + \alpha_2) = c\beta_1 + \beta_2$$

o que mostra que $c\beta_1 + \beta_2$ também está na imagem de T. Portanto, R_T é um subespaço de W.

Seja N_T o conjunto dos α em V tais que $T\alpha = 0$. Se α_1 e α_2 estão em N_T e c é um escalar arbitrário, então

$$\begin{aligned} T(c\alpha_1 + \alpha_2) &= c(T\alpha_1) + T\alpha_2 \\ &= c(0) + 0 \\ &= 0 \end{aligned}$$

de modo que $c\alpha_1 + \alpha_2$ está novamente em N_T. Logo, N_T é um subespaço.

No exemplo 1, a imagem da transformação idêntica é todo o espaço V, ao passo que seu núcleo é o subespaço nulo. A imagem da transformação nula é o subespaço nulo e seu núcleo é todo o espaço V. No Exemplo 2, a imagem da transformação derivação é todo o espaço das funções polinomiais e seu núcleo é o subespaço das funções constantes (se F é um subcorpo do corpo dos números complexos). No Exemplo 3, a imagem da transformação $X \to AX$

é o conjunto das $m \times 1$ matrizes Y tais que o sistema de equações $AX = Y$ admite solução, isto é, tôdas Y que sejam combinações lineares das colunas de A. O núcleo desta transformação é o conjunto das $n \times 1$ matrizes X tais que $AX = 0$, isto é, o espaço-solução de A. No Exemplo 4, a imagem e o núcleo de T são um tanto difíceis de descrever, exceto pela repetição de suas definições. No Exemplo 5, a imagem de T consiste das g que têm a primeira derivada contínua e tais que $g(0) = 0$. O núcleo neste caso é o subespaço nulo.

Se T é uma transformação linear de V em W e $\alpha_1, \ldots, \alpha_n$ são vetores que geram V, então é evidente que os vetores $T\alpha_1, \ldots, T\alpha_n$ geram a imagem de T. Em particular, se V fôr de dimensão finita, a imagem de T será um subespaço de W de dimensão finita.

Definição. *Seja* T *uma transformação linear de* V *em* W, *sendo* V *de dimensão finita. O* **pôsto** *de* T *é a dimensão de imagem de* T. *A* **nulidade** *de* T *é a dimensão do núcleo de* T.

Teorema 3. *Sejam* V *e* W *espaços vetoriais sôbre o corpo* F *e seja* T *uma transformação linear de* V *em* W. *Suponhamos que* V *seja de dimensão finita. Então*

$$\text{pôsto}\,(T) + \text{nulidade}\,(T) = \dim V.$$

Demonstração. Seja $\{\alpha_1, \ldots, \alpha_k\}$ uma base de N, o núcleo de T. Existem vetores $\alpha_{k+1}, \ldots, \alpha_n$ em V tais que $\{\alpha_1, \ldots, \alpha_n\}$ seja uma base de V. Demonstraremos agora que $\{T\alpha_{k+1}, \ldots, T\alpha_n\}$ é uma base da imagem de T. Os vetores $T\alpha_1, \ldots, T\alpha_n$ certamente geram a imagem de T e, como $T\alpha_j = 0$, para $j \leq k$, vemos que $T\alpha_{n+1}, \ldots, T\alpha_n$ geram a imagem. Para ver que êsses vetores são independentes, suponhamos que existam escalares c_i tais que

$$\sum_{i=k+1}^{n} c_i(T\alpha_i) = 0.$$

Isto diz que

$$T\left(\sum_{i=k+1}^{n} c_i\alpha_i\right) = 0$$

e, conseqüentemente, o vetor $\alpha = \sum_{i=k+1}^{n} c_i\alpha_i$ está no núcleo de T. Como $\alpha_1, \ldots, \alpha_k$ formam uma base de N, existem, necessàriamente, escalares $b_1, \ldots, b_k$ tais que

$$\alpha = \sum_{i=1}^{k} b_i\alpha_i.$$

Assim

$$\sum_{i=1}^{k} b_i\alpha_i - \sum_{j=k+1}^{n} c_j\alpha_j = 0$$

e, como $\alpha_1, \ldots, \alpha_n$ são linearmente independentes, devemos ter

$$b_1 = \ldots = b_k = c_{k+1} = \ldots = c_n = 0.$$

Se r é o pôsto de T, o fato de $T\alpha_{k+1}, \ldots, T\alpha_n$ formarem uma base da imagem de T nos diz que $r = n - k$. Como k é a nulidade de T e n é a dimensão de V, está completa a demonstração.

Exercícios

1. Quais das seguintes funções T de R^2 em R^2 são transformações lineares?

(a) $T(x_1, x_2) = (1 + x_1, x_2)$;
(b) $T(x_1, x_2) = x_2, x_1)$;
(c) $T(x_1, x_2) = (0, x_2)$;
(d) $T(x_1, x_2) = (\text{sen } x_1, x_2)$;
(e) $T(x_1, x_2) = (x_1 - x_2, 0)$.

2. Verificar que as transformações definidas em (a) e (b) são transformações lineares de R^2 em R^3. Determinar, para cada uma delas, a imagem, pôsto, núcleo e nulidade.

(a) $T(x_2, x_2) = (x_1 + x_2, x_1 - x_2, x_2)$;
(b) $T(x_1, x_2) = (x_1 - x_2, x_2 - x_1, -x_1)$.

3. Descrever explìcitamente (como nos Exercícios 1 e 2) a transformação linear T de F^2 em F^2 tal que $T\epsilon_1 = (a, b)$, $T\epsilon_2 = (c, d)$.

4. Seja F um subcorpo do corpo dos números complexos e seja T a função de F^3 em F^3 definida por

$$T(x_1, x_2, x_3) = (x_1 - x_2 + 2x_3, 2x_1 + x_2, -x_1 - 2x_2 + 2x_3).$$

(a) Verificar que T é uma transformação linear.

(b) Se (a, b, c) é um vetor em F^3, quais as condições sôbre a, b, c, para que o vetor esteja na imagem de T? Qual é o pôsto de T?

(c) Quais são as condições sôbre a, b e c para que o vetor esteja no núcleo de T? Qual é a nulidade de T?

5. Descrever explìcitamente uma transformação linear de R^3 em R^3 cuja imagem seja o subespaço gerado por $(1, 0, -1)$ e $(1, 2, 2)$.

6. Seja V o espaço vetorial das $n \times n$ matrizes sôbre o corpo F e seja B uma $n \times n$ matriz fixa. Se

$$T(A) = AB - BA$$

verificar que T é uma transformação linear de V em V.

7. Seja V o conjunto dos números complexos considerado como um espaço vetorial sôbre o corpo dos números *reais* (operações usuais). Determinar uma função de V em V que seja uma transformação linear sôbre o

espaço vetorial acima, mas que não seja uma transformação linear sôbre C^1, isto é, que não seja linear complexa.

8. Seja V o espaço das $n \times 1$ matrizes sôbre F e seja W o espaço das $m \times 1$ matrizes sôbre F. Seja A uma $m \times n$ matriz fixa sôbre F e seja T a transformação linear de V em W definida por $T(X) = AX$. Demonstrar que T é a transformação nula se e sòmente se A é a matriz nula.

9. Seja V um espaço vetorial n-dimensional sôbre o corpo F e seja T uma transformação linear de V em V tal que a imagem e o núcleo de T sejam idênticos. Demonstrar que n é par. (Dar um exemplo de uma tal transformação linear.)

10. Seja V um espaço vetorial e T uma transformação linear de V em V. Demonstrar que as duas afirmações seguintes sôbre T são equivalentes:

(a) A interseção da imagem de T com o núcleo de T é o subespaço nulo de V.

(b) Se $T(T\alpha) = 0$, então $T\alpha = 0$.

11. Usar o Teorema 3 e o Exercício 6 da Seção 2.5 para demonstrar que para qualquer $m \times n$ matriz A, o pôsto-linha de A é igual ao pôsto-coluna de A. (Esbôço da demonstração: Seja V o espaço das $n \times 1$ matrizes sôbre F e seja W o espaço das $m \times 1$ matrizes. Seja T a transformação linear de V em W definida por $T(X) = AX$. Usar o Teorema 3 para demonstrar que a soma do pôsto-coluna de A com a dimensão do espaço-solução de A é n. Agora concluir a demonstração com o citado exercício do Capítulo 2.)

3.2 A Álgebra das Transformações Lineares

No estudo das transformações lineares de V em W, é de importância fundamental o fato de que o conjunto dessas transformações herda uma estrutura natural de espaço vetorial. O conjunto das transformações lineares de um espaço V em si mesmo possui uma estrutura algébrica mais rica pois a composição usual de funções fornece uma "multiplicação" dessas transformações. Nesta seção exploraremos essas idéias.

Teorema 4. *Sejam* V *e* W *espaços vetoriais sôbre o corpo* F. *Sejam* T *e* U *transformações lineares de* V *em* W. *A função* (T + U) *definida por*

$$(T + U)(\alpha) = T\alpha + U\alpha$$

é uma transformação linear de V *em* W. *Se* c *é um elemento qualquer de* F, *a função* (cT) *definida por*

$$(cT)(\alpha) = c(T\alpha)$$

é uma transformação linear de V *em* W. *O conjunto das transformações lineares de* V *em* W, *munido da adição e multiplicação escalar acima definida, é um espaço vetorial sôbre o corpo* F.

Demonstração. Suponhamos que T e U sejam transformações lineares de V em W e definamos $(T + U)$ como acima. Então

$$\begin{aligned}(T + U)(c\alpha + \beta) &= T(c\alpha + \beta) + U(c\alpha + \beta)\\ &= c(T\alpha) + T\beta + c(U\alpha) + U\beta\\ &= c(T\alpha + U\alpha) + (T\beta + U\beta)\\ &= c(T + U)(\alpha) + (T + U)(\beta)\end{aligned}$$

o que mostra que $(T + U)$ é uma transformação linear. Anàlogamente,

$$\begin{aligned}(cT)(d\alpha + \beta) &= c[T(d\alpha + \beta)]\\ &= c[d(T\alpha) + T\beta]\\ &= cd(T\alpha) + c(T\beta)\\ &= d[c(T\alpha)] + c(T\beta)\end{aligned}$$

mostrando que (cT) é uma transformação linear.

Para verificar que o conjunto das transformações lineares de V em W (munido destas duas operações) é um espaço vetorial, é necessário verificar diretamente cada uma das condições sôbre a adição de vetores e a multiplicação escalar. Deixamos a parte principal disto a cargo do leitor e contentamo-nos com êste comentário: O vetor nulo dêste espaço será a transformação nula, que leva todo vetor de V no vetor nulo de W; cada uma das propriedades das duas operações decorre diretamente da propriedade correspondente das operações no espaço W.

Talvez devamos mencionar outra maneira de considerar êste teorema. Se se define soma e múltiplo escalar como fizemos acima, então o conjunto de *tôdas* as funções de V em W torna-se um espaço vetorial sôbre o corpo F. Isto nada tem a ver com o fato de V ser um espaço vetorial, mas apenas com o fato de V ser um conjunto não-vazio. Quando V é um espaço vetorial podemos definir uma transformação linear de V em W e o Teorema 4 diz que as transformações lineares formam um subespaço do espaço de tôdas as funções de V em W.

Indicaremos o espaço das transformações lineares de V em W por $L(V, W)$. Lembramos novamente ao leitor que $L(V, W)$ está definido sòmente para V e W espaços vetoriais sôbre o mesmo corpo.

Teorema 5. *Seja* V *um espaço vetorial* n*-dimensional sôbre o corpo* F *e seja* W *um espaço vetorial* m*-dimensional sôbre* F. *Então o espaço* L(V, W) *é de dimensão finita e tem dimensão* mn.

Demonstração. Sejam

$$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\} \quad \text{e} \quad \mathcal{B}' = \{\beta_1, \ldots, \beta_m\}$$

bases ordenadas de V e W, respectivamente. Para cada par de inteiros (p, q) com $1 \leq p \leq m$ e $1 \leq q \leq n$, definamos uma transformação linear $E^{p,q}$ de V em W por

$$E^{p,q}(\alpha_i) = \begin{cases} 0; & \text{se } i \neq q \\ \beta_p, & \text{se } i \neq q \end{cases} \qquad \begin{matrix} p \neq q \\ q = k \end{matrix}$$
$$= \delta_{iq}\beta_p.$$

De acôrdo com o Teorema 1, existe uma única transformação linear de V em W que satisfaz estas condições. Afirmamos que às mn transformações $E^{p,q}$ formam uma base de $L(V, W)$.

Seja T uma transformação linear de V em W. Para cada j, $1 \leq j \leq n$, sejam $A_{ij}, \ldots, A_{mj}$ as coordenadas do vetor $T\alpha_j$ em relação à base ordenada $\mathcal{B}'$, isto é,

$$(3\text{—}1) \qquad T\alpha_j = \sum_{p=1}^{m} A_{pj}\beta_p.$$

Desejamos mostrar que

$$(3\text{—}2) \qquad T = \sum_{p=1}^{m} \sum_{q=1}^{m} A_{pq}E^{p,q}.$$

Seja U a transformação linear no segundo membro de (3—2). Então para cada j

$$\begin{aligned} U\alpha_j &= \sum_p \sum_q A_{pq}E^{p,q}(\alpha_j) \\ &= \sum_p \sum_q A_{pq}\delta_{jq}\beta_p \\ &= \sum_{p=1}^{m} A_{pj}\beta_p \\ &= T\alpha_j \end{aligned}$$

e, conseqüentemente, $U = T$. Agora (3—2) mostra que as $E^{p,q}$ geram $L(V, W)$; precisamos demonstrar que elas são independentes. Mas isto é evidente pelo que fizemos acima, pois, se a transformação

$$U = \sum_p \sum_q A_{pq}E^{p,q}$$

é a transformação nula, então $U\alpha_j = 0$ para cada j, portanto

$$\sum_{p=1}^{m} A_{pj}\beta_p = 0$$

e a independência dos β_p implica que $A_{pj} = 0$ para todos p e j.

Teorema 6. *Sejam* V, W *e* Z *espaços vetoriais sôbre o corpo* F. *Seja* T *uma transformação linear de* V *em* W *e* U *uma transformação linear de* W *em* Z. *Então, a função composta* UT, *definida por* (UT) $(\alpha) = U(T(\alpha))$ *é uma transformação linear de* V *em* Z.

Demonstração.

$$\begin{aligned}(UT)(c\alpha + \beta) &= U[T(c\alpha + \beta)]\\ &= U(cT\alpha + T\beta)\\ &= c[U(T\alpha)] + U(T\beta)\\ &= c(UT)(\alpha) + (UT)(\beta).\end{aligned}$$

No que segue, estaremos primordialmente preocupados com transformações lineares de um espaço vetorial em si mesmo. Como teríamos de escrever a todo instante "T é uma transformação linear de V em V", substituiremos isto por "T é um operador linear sôbre V".

Definição. *Se* V *é um espaço vetorial sôbre o corpo* F, *um* **operador linear sôbre** V *é uma transformação linear de* V *em* V.

Aplicando o Teorema 6 para $V = W = Z$, de modo que U e T sejam operadores lineares sôbre o espaço V, vemos que a composta UT é ainda um operador linear sôbre V. Assim, o espaço $L(V, V)$ possui uma "multiplicação", definida sôbre si por meio da composição. Neste caso, o operador UT também está definido e devemos notar que em geral $UT \neq TU$, isto é, $UT - TU \neq 0$. Devemos notar particularmente o fato de que se T é um operador linear sôbre V, então podemos compor T com T. Usaremos a notação $T^2 = TT$ e, em geral, $T^n = T \ldots T$ (n vêzes) para $n = 1, 2, 3, \ldots$. Definimos $T^0 = I$ se $T \neq 0$.

Exemplo 6. Seja F um corpo e V o espaço vetorial das funções polinomiais de F em F. Seja D o operador derivação definido no Exemplo 2 e T o operador linear "multiplicação por x":

$$(Tf)(x) = xf(x).$$

Então $DT \neq TD$. Na verdade, o leitor deverá achar fácil verificar que $DT - TD = I$, o operador idêntico.

Mesmo não sendo comutativa, a "multiplicação" que temos sôbre $L(V, V)$ está bastante relacionada com as operações de espaço vetorial de $L(V, V)$.

Lema. *Seja* V *um espaço vetorial sôbre o corpo* F; *sejam* U, T_1 *e* T_2 *operadores lineares sôbre* V; *seja* c *um elemento de* F.

(a) $IU = UI = U$;
(b) $U(T_1 + T_2) = UT_1 + UT_2$; $(T_1 + T_2)U = T_1U + T_2U$;
(c) $c(UT_1) = (cU)T_1 = U(cT_1)$.

Demonstração. (a) Esta propriedade da função idêntica é óbvia. Enunciamo-la aqui por mera questão de ênfase.

(b)
$$\begin{aligned}[U(T_1 + T_2)](\alpha) &= U[(T_1 + T_2)(\alpha)]\\ &= U(T_1\alpha + T_2\alpha)\\ &= U(T_1\alpha) + U(T_2\alpha)\\ &= (UT_1)(\alpha) + (UT_2)(\alpha)\end{aligned}$$

de modo que $U(T_1 + T_2) = UT_1 + UT_2$. Além disso

$$\begin{aligned}[(T_1 + T_2)U](\alpha) &= (T_1 + T_2)(U\alpha)\\ &= T_1(U\alpha) + T_2(U\alpha)\\ &= (T_1U)(\alpha) + (T_2U)(\alpha)\end{aligned}$$

de modo que $(T_1 + T_2)U = T_1U + T_2U$. (O leitor pode notar que para as demonstrações destas duas leis distributivas não foi usado o fato de T_1 e T_2 serem lineares, e que para a demonstração da segunda não foi usado tampouco o fato de U ser linear.)

(c) Deixamos a demonstração da parte (c) a cargo do leitor.

Êste lema e uma parte do Teorema 5 nos dizem que o espaço vetorial $L(V, V)$, munido da operação de composição, é o que se conhece por uma álgebra linear com elemento unidade. Discutiremos isto no Capítulo 4.

Teorema 7. *Sejam* V *e* W *espaços vetoriais sôbre o corpo* F *e seja* T *uma transformação linear de* V *em* W. *Se* T *é injetora e sobrejetora então a função inversa* T^{-1} *é uma transformação linear de* W *sôbre V.*

Demonstração. Lembramos ao leitor que T **injetora** significa que $T\alpha \neq T\beta$ sempre que $\alpha \neq \beta$ e que T **sobrejetora** significa que a imagem de T é todo o espaço W. Quando T é injetora e sobrejetora, existe uma função inversa T^{-1}, determinada de modo único, que leva W sôbre V tal que $T^{-1}T$ é a função idêntica de V e TT^{-1} é a função idêntica de W. O que estamos demonstrando aqui é que, se uma função linear T é inversível, então a inversa T^{-1} também é linear.

Sejam β_1 e β_2 vetores em W e seja c um escalar. Queremos mostrar que

$$T^{-1}(c\beta_1 + \beta_2) = cT^{-1}\beta_1 + T^{-1}\beta_2.$$

Seja $\alpha_i = T^{-1}\beta_i$, $i = 1, 2$, isto é, seja α_i o único vetor em V tal que $T\alpha_i = \beta_i$. Como T é linear,

$$\begin{aligned} T(c\alpha_1 + \alpha_2) &= cT\alpha_1 + T\alpha_2 \\ &= c\beta_1 + \beta_2. \end{aligned}$$

Assim $c\alpha_1 + \alpha_2$ é o único vetor em V que é levado por T em $c\beta_1 + \beta_2$, portanto

$$\begin{aligned} T^{-1}(c\beta_1 + \beta_2) &= c\alpha_1 + \alpha_2 \\ &= c(T^{-1}\beta_1) + T^{-1}\beta_2 \end{aligned}$$

e T^{-1} é linear.

Exemplo 7. Se V é o espaço das funções polinomiais sôbre o corpo F, D o operador derivação e T o operador "multiplicação por x" do Exemplo 6, temos a seguinte situação: A imagem de D é todo o espaço V, logo D é sobrejetor; contudo, D não é injetor, pois tôdas as funções constantes são levadas em 0 por D. A imagem de T é o conjunto das funções polinomiais f para as quais $f(0) = 0$, logo T não é sobrejetor; T é injetor, como se pode verificar fàcilmente.

Exemplo 8. Seja F um corpo e seja T o operador linear sôbre F^2 definido por

$$T(x_1, x_2) = (x_1 + x_2, x_1).$$

Então, T é injetor, pois se $T(y_1, y_2) = T(x_1, x_2)$, temos

$$\begin{aligned} x_1 + x_2 &= y_1 + y_2 \\ x_1 &= y_1 \end{aligned}$$

de modo que $x_1 = y_1$ e $x_2 = y_2$. Vejamos também que T é sobrejetor; de fato, seja (z_1, z_2) um vetor arbitrário em F^2. Para mostrar que (z_1, z_2) está na imagem de T precisamos determinar x_1 e x_2 tais que

$$\begin{aligned} x_1 + x_2 &= z_1 \\ x_1 &= z_2 \end{aligned}$$

e a solução é evidentemente $x_1 = z_2$, $x_2 = z_1 - z_2$. Êste último cálculo nos dá até mesmo a fórmula explícita para T^{-1}, a saber,

$$T^{-1}(z_1, z_2) = (z_2, z_1 - z_2).$$

Se T é uma transformação linear de V em W diremos que T é **não-singular** se o núcleo de T consistir apenas do vetor nulo. O leitor certamente observará que isto é equivalente à afirmação de que T é injetora, porque quando T é linear $T\alpha = T\beta$ se, e sòmente se, $T(\alpha - \beta) = 0$. Nesta linguagem, T é inversível (injetora e sobrejetora) se e sòmente se T é não-singular e a imagem de T é todo o espaço W.

Teorema 8. *Seja* T *uma transformação linear de* V *em* W. *Então* T *é não-singular se e sòmente se* T *leva todo subconjunto linearmente independente de* V *sôbre um subconjunto linearmente indèpendente de* W.

Demonstração. Suponhamos primeiro que T seja não-singular. Seja S um subconjunto linearmente independente de V. Se $\alpha_1, \ldots, \alpha_k$ são vetores em S, então os vetores $T\alpha_1, \ldots, T\alpha_k$ são linearmente independentes, pois se

$$c_1(T\alpha_1) + \ldots + c_k(T\alpha_k) = 0$$

então

$$T(c_1\alpha_1 + \ldots + c_k\alpha_k) = 0$$

e como T é não-singular

$$c_1\alpha_1 + \ldots + c_k\alpha_k = 0$$

do que segue que cada $c_i = 0$ pois S é um conjunto independente. Êste argumento mostra que a imagem de S por meio de T é independente.

Suponhamos que T leve subconjuntos independentes sôbre subconjuntos independentes. Seja α um vetor não-nulo em V. Então o conjunto S constituído apenas pelo vetor α é independente. A imagem de S é o conjunto constituído apenas pelo vetor $T\alpha$ e êste conjunto é independente. Portanto $T\alpha \neq 0$, pois o conjunto constituído apenas pelo vetor nulo é dependente. Isto mostra que o núcleo de T é o subespaço nulo, isto é, T é não-singular.

Teorema 9. *Sejam* V *e* W *espaços vetoriais de dimensão finita sôbre o corpo* F *tais que dim* V = *dim* W. *Se* T *é uma transformação linear de* V *em* W, *as seguintes afirmações são equivalentes:*

(i) T *é inversível.*

(ii) T *é não-singular.*

(iii) *A imagem de* T *é* W.

(iv) *Se* $\{\alpha_1, \ldots, \alpha_n\}$ *é uma base arbitrária de* V, *então* $\{T\alpha_1, T\alpha_n\}$ *é uma base de* W.

(v) *Existe pelo menos uma base* $\{\alpha_1, \ldots, \alpha_n\}$ *de* V *tal que* $\{T\alpha_1, \ldots, T\alpha_n\}$ *seja uma base de* W.

Demonstração. (i) $\to$ (ii). Se T é inversível, T é não-singular. (ii) $\to$ (iii). Suponhamos que T seja não-singular. Seja $\{\alpha_1, \ldots, \alpha_n\}$ uma base de V. Pelo Teorema 8, $\{T\alpha_1, \ldots, T\alpha_n\}$ é um conjunto linearmente independente de vetores em W e como a dimensão de W também é n, êste conjunto de vetores é uma base de W. Agora seja β um vetor arbitrário em W. Existem escalares $c_1, \ldots, c_n$ tais que

$$\begin{aligned}\beta &= c_1(T\alpha_1) + \ldots + c_n(T\alpha_n)\\ &= T(c_1\alpha_1 + \ldots + c_n\alpha_n)\end{aligned}$$

o que mostra que β está na imagem de T. (iii) $\to$ (iv). Suponhamos agora que T seja sobrejetora. Se $\{\alpha_1, \ldots, \alpha_n\}$ é uma base arbitrária de V, os vetores $T\alpha_1, \ldots, T\alpha_n$ geram a imagem de T, que é todo o espaço W por hipótese. Como a dimensão de W é n, êstes n vetores precisam ser linearmente independentes, isto é, precisam formar uma base de W. (iv) $\to$ (v). Não requer nenhum comentário. (v) $\to$ (i). Suponhamos que exista alguma base $\{\alpha_1, \ldots, \alpha_n\}$ de V tal que $\{T\alpha_1, \ldots, T\alpha_n\}$ seja uma base de W. Como os $T\alpha_i$ geram W, é evidente que a imagem de T coincide com W. Se $\alpha = c_1\alpha_1 + \ldots + c_n\alpha_n$ está no núcleo de T, então

$$T(c_1\alpha_1 + \ldots + c_n\alpha_n) = 0$$

ou

$$c_1(T\alpha_1) + \ldots + c_n(T\alpha_n) = 0$$

e como os $T\alpha_i$ são independentes, cada $c_i = 0$ e assim $\alpha = 0$. Mostramos que a imagem de T é W e que T é não-singular, logo T é inversível.

Se V é um espaço vetorial de dimensão finita, o Teorema 9 nos diz o que segue a respeito de operadores lineares sôbre V. Se T é um operador linear sôbre V que é injetor (não-singular) então a imagem de T é, necessàriamente, todo V, logo tem que ser inversível. Se a imagem de T é todo V então T tem que ser não-singular, logo inversível. Estas duas afirmações possuem uma demonstração mais simples que a que fizemos e que usa o Toeroma 3. Seja r o pôsto de T e seja k a nulidade de T. Pelo Teorema 3, $r + k = n$. A afirmação de que T é não-singular significa que $k = 0$, ao passo que a afirmação de que a imagem de T é V significa que $r = n$. Como $r + k = n$ estas afirmações sôbre T são òbviamente equivalentes. O leitor é aconselhado a compreender ambas as demonstrações. Note-se que o Exemplo 7 mostra que nem a injeção nem a sobrejeção implicam uma à outra para operadores em um espaço que não seja de dimensão finita.

Lema *Seja* V *um espaço vetorial sôbre o corpo* F, *e sejam* U *e* T *operadores lineares inversíveis sôbre* V. *Então* UT *é inversível e* $(UT)^{-1} = T^{-1}U^{-1}$.

Demonstração. Verificar simplesmente que

$$(UT)(T^{-1}U^{-1}) = (T^{-1}U^{-1})(UT) = I.$$

O conjunto dos operadores lineares inversíveis sôbre um espaço V, com a operação de composição, fornece um belo exemplo do que é conhecido em álgebra por um "grupo". Apesar de que não teremos tempo para discutir grupos com quaisquer promenores, daremos pelo menos a definição.

Definição. *Um* **grupo** *consiste do seguinte:*

(1) *Um conjunto* G;

(2) *Um regra (ou operação) que associa a cada par de elementos* x, y *em* G *um elemento* xy *em* G *de uma maneira tal que*

(a) x(yz) = (xy)z, *para todos* x, y e z *em* G (*associatividade*);

(b) *existe um elemento* e *em* G *tal que* ex = xe = x, *para todo* x *em* G;

(c) *a cada elemento* x *em* G *corresponde um elemento* x^{-1} *em* G *tal que* $xx^{-1} = x^{-1}x = e$.

O lema acima nos diz que a composição $(U, T) \rightarrow UT$ associa a cada par de operadores lineares inversíveis sôbre um espaço V outro operador inversível sôbre V. A composição é uma operação associativa. O operador idêntico I satisfaz $IT = TI = T$ para todo T e para um T inversível existe (pelo Teorema 7) um operador linear inversível T^{-1} tal que $TT^{-1} = T^{-1}T = I$. Portanto o conjunto dos operadores lineares inversíveis sôbre V, munido desta operação, é um grupo. O conjunto das $n \times n$ matrizes inversíveis com a multiplicação de matrizes como a operação é outro exemplo de um grupo. Um grupo é dito **comutativo** se satisfaz a condição $xy = yx$ para todos x e y. Os dois exemplos que demos acima não são, em geral, grupos comutativos. Freqüentemente indica-se a operação num grupo comutativo por $(x, y) \rightarrow x + y$ em lugar de $(x, y) \rightarrow xy$ e usa-se então o símbolo 0 para o elemento "unidade" e. O conjunto dos vetores de um espaço vetorial, com a operação de adição de vetores, é um grupo comutativo. Um corpo pode ser descrito como um conjunto com duas operações, denominadas adição e multiplicação, que é um grupo comutativo em relação à adição e no qual os elementos não-nulos formam um grupo comutativo em relação à multiplicação, valendo a lei distributiva $x(y + z) = xy + xz$.

Exercícios

1. Sejam T e U os operadores lineares sôbre R^2 definidos por

$$T(x_1, x_2) = (x_2, x_1) \text{ e } U(x_1, x_2) = (x_1, 0).$$

(a) Como você descreveria T e U geomètricamente?
(b) Dar regras como as que definem T e U para cada uma das transformações $(U + T)$, UT, TU, T^2 U^2.

2. Seja T o (único) operador linear sôbre C^3 para o qual

$$T\epsilon_1 = (1, 0, i), \quad T\epsilon_2 = (0, 1, 1), \quad T\epsilon_3 = (i, 1, 0).$$

T é inversível?

3. Seja T o operador linear sôbre R^3 definido por

$$T(x_1, x_2, x_3) = (3x_1, x_1 - x_2, 2x_1 + x_2 + x_3).$$

T é inversível? Em caso afirmativo, determinar uma regra para T^{-1} como a que define T.

4. Para o operador linear T do Exercício 3, demonstrar que

$$(T^2 - I)(T - 3I) = 0.$$

5. Seja V o espaço vetorial complexo das 2×2 matrizes como elementos complexos. Seja

$$B = \begin{bmatrix} 1 & -1 \\ -4 & 4 \end{bmatrix}$$

e seja T o operador linear sôbre V definido por $T(A) = BA$. Qual é o pôsto de T? Descrever T^2.

6. Seja T uma transformação linear de R^3 em R^2 e seja U uma transformação linear de R^2 em R^3. Demonstrar que a transformação linear UT não é inversível. Generalizar o teorema.

7. Determinar dois operadores lineares T e U sôbre R^2 tais que $TU = 0$ mas $UT \neq 0$.

8. Seja V um espaço vetorial sôbre o corpo F e T um operador linear sôbre V. Se $T^2 = 0$, o que se pode dizer sôbre a relação entre a imagem de T e o núcleo de T? Dar um exemplo de um operador linear T sôbre R^2 tal que $T^2 = 0$ mas $T \neq 0$.

9. Seja T um operador linear sôbre o espaço vetorial V de dimensão finita. Suponhamos que exista um operador linear U sôbre V tal que $TU = I$. Demonstrar que T é inversível e $U = T^{-1}$. Dar um exemplo que mostre que isto é falso quando V não é de dimensão finita. (*Sugestão*: Seja $T = D$, o operador derivação sôbre o espaço das funções polinomiais.)

10. Seja V um espaço vetorial de dimensão finita e seja B uma base ordenada de V. Construir a base $\{E^{p,q}\}$ para $L(V, V)$ como na demonstração do Teorema 5, usando $\mathcal{B}' = \mathcal{B}$. Então $E^{p,q}E^{r,s} = ?$

11. Seja V um espaço vetorial de dimensão finita e seja T um operador linear sôbre V. Suponhamos que pôsto(T^2) = pôsto(T). Demonstrar que a imagem e o núcleo de T são disjuntos, isto é, possuem em comum apenas o vetor nulo.

12. Sejam p, m e n inteiros positivos e F um corpo. Seja V o espaço das $m \times n$ matrizes sôbre F e W o espaço das $p \times n$ matrizes sôbre F. Seja B uma $p \times m$ matriz fixa e seja T a transformação linear de V em W definida por $T(A) = BA$. Demonstrar que T é inversível se, e sòmente se, $p = m$ e B é uma $m \times m$ matriz inversível.

3.3 Isomorfismo

Se V e W são espaços vetoriais sôbre o corpo F, uma transformação linear bijetora (injetora e sobrejetora) T de V em W é denominada um **isomorfismo de V em W**. Se existir um isomorfismo de V em W, diremos que V é **isomorfo a W**.

Notemos que V é trivialmente isomorfo a V, pois o operador idêntico é um isomorfismo de V em V. Além disso, se V é isomorfo a W por meio de um isomorfismo T, então W é isomorfo a V uma vez que T^{-1} é um isomorfismo de W em V. O leitor deverá achar fácil verificar que se V é isomorfo a W e W é isomorfo a Z, então V é isomorfo a Z. Em suma, o isomorfismo é uma relação de equivalência sôbre a classe dos espaços vetoriais. Se existir um isomorfismo de V em W, poderemos às vêzes dizer que V e W são isomorfos, em vez de dizer que V é isomorfo a W. Isto não causará confusão alguma porque V é isomorfo a W se, e sòmente se, W é isomorfo a V.

Teorema 10. *Todo espaço vetorial n-dimensional sôbre o corpo* F *é isomorfo ao espaço F^n.*

Demonstração. Seja V um espaço n-dimensional sôbre o corpo F e seja $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ uma base ordenada de V. Definamos uma função T de V em F^n, como segue: Se α está em V, seja $T\alpha$ a n-upla $(x_1, \ldots, x_n)$ das coordenadas de α em relação à base ordenada $\mathcal{B}$, isto é, a n-upla tal que

$$\alpha = x_1\alpha_1 + \ldots + x_n\alpha_n.$$

Em nossa discussão de coordenadas no Capítulo 2, verificamos que esta T é linear, injetora e leva V sôbre F^n.

Para muitos objetivos freqüentemente consideram-se espaços vetoriais isomorfos como sendo "o mesmo", apesar de que os vetores e as operações nos espaços possam ser bem diferentes, isto é, freqüentemente identificamos espaços isomorfos. Não tentaremos fazer uma longa discussão sôbre esta idéia no momento mas deixaremos a compreensão do isomorfismo e do sentido no qual espaços isomorfos são "o mesmo" crescerem à medida que continuemos nosso estudo de espaços vetoriais.

Faremos alguns comentários breves. Suponhamos que T seja um isomorfismo de V em W. Se S é um subconjunto de V, o Teorema 8 nos diz que S é linearmente independente se, e sòmente se, o conjunto $T(S)$ em W é independente. Portanto, ao decidirmos se S é independente não importa se consideramos S ou $T(S)$. A partir disto vê-se que um isomorfismo "conserva a dimensão", isto é, todo subespaço de V de dimensão finita tem a mesma dimensão que sua imagem por meio de T. Eis uma ilustração muito simples dessa idéia. Suponhamos que A seja uma $m \times n$ matriz sôbre o corpo F. Na verdade demos duas definições do espaço-solução da matriz A. O primeiro é o conjunto das n-uplas $(x_1, \ldots, x_n)$ em F^n que satisfazem cada uma das equações do sistema $AX = 0$. O segundo é o conjunto das $n \times 1$ matrizes colunas X tais que $AX = 0$. O primeiro espaço-solução é portanto um subespaço de F^n e o segundo é um subespaço do espaço de tôdas as $n \times 1$ matrizes sôbre F. Agora existe um isomorfismo evidente entre F^n e as $n \times 1$ matrizes colunas sôbre F, a saber,

$$(x_1, \ldots, x_n) \to \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

Por meio dêste isomorfismo, o primeiro espaço-solução de A é levado sôbre o segundo espaço-solução. Êstes espaços têm a mesma dimensão, portanto se quisermos demonstrar um teorema sôbre a dimensão do espaço-solução, não importará qual espaço resolvamos discutir. Na verdade, o leitor provàvelmente não objetaria se resolvêssemos identificar F^n com o espaço das $n \times 1$ matrizes. Poderemos fazê-lo quando fôr conveniente, e quando não o fôr não o faremos.

Exercícios

1. Seja V o conjunto dos números complexos e seja F o corpo dos números reais. Com as operações usuais, V é um espaço vetorial sôbre F. Descrever explìcitamente um isomorfismo dêste espaço em R^2.

2. Seja V um espaço vetorial sôbre o corpo dos números complexos e suponhamos que exista um isomorfismo T de V em C^3. Sejam $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ vetores em V tais que

$$T\alpha_1 = (1, 0, i) \qquad T\alpha_2 = (-2, 1 + i, 0),$$
$$T\alpha_3 = (-1, 1, 1), \qquad T\alpha_4 = (\sqrt{2}, i, 3).$$

(a) α_1 está no subespaço gerado por α_2 e α_3?

(b) Seja W_1 o subespaço gerado por α_1 e α_2 e seja W_2 o subespaço gerado por α_3 e α_4. Qual é a interseção de W_1 com W_2?

(c) Determinar uma base do subespaço de V gerado pelos quatro vetores α_j.

3. Seja W o conjunto das 2×2 matrizes hermitianas, complexas, isto é, o conjunto das 2×2 matrizes complexas A tais que $A_{ij} = \overline{A_{ji}}$ (a barra indica conjugação complexa). Como destacamos no Exemplo 6 do Capítulo 2, W é um espaço vetorial sôbre o corpo dos números *reais*, em relação às operações usuais. Verificar que

$$(x, y, z, t) \rightarrow \begin{bmatrix} t + x & y + iz \\ y - iz & t - x \end{bmatrix}$$

é um isomorfismo de R^4 em W.

4. Seja V o conjunto dos números complexos considerado como um espaço vetorial sôbre o corpo dos números reais (Exercício 1). Definamos uma função T de V no espaço das 2×2 matrizes reais, como segue. Se $z = x + iy$ com x e y números reais, então

$$T(z) = \begin{bmatrix} x + 7y & 5y \\ -10y & x - 7y \end{bmatrix}.$$

(a) Verificar que T é uma transformação linear (real) injetora de V no espaço das 2×2 matrizes.

(b) Verificar que $T(z_1 z_2) = T(z_1)T(z_2)$.

(c) Como você descreveria a imagem de T?

5. Sejam V e W espaços vetoriais de dimensão finita sôbre o corpo F. Demonstrar que V e W são isomorfos se, e sòmente se, $\dim V = \dim W$.

6. Sejam V e W espaços vetoriais sôbre o corpo F e seja U um isomorfismo de V em W. Demonstrar que $T \rightarrow UTU^{-1}$ é um isomorfismo de $L(V, V)$ em $L(W, W)$.

3.4 Representação de Transformações por Matrizes

Seja V um espaço vetorial n-dimensional sôbre o corpo F e seja W um espaço vetorial m-dimensional sôbre F. Sejam $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ uma base ordenada de V e $\mathcal{B}' = \{\beta_1, \ldots, \beta_m\}$ uma base ordenada de W. Se T é uma transformação linear arbitrária de V em W, então T é determinada por seu efeito sôbre os vetores α_j. Cada um dos n vetores $T\alpha_j$ pode ser expresso de modo único como uma combinação linear

$$T\alpha_j = \sum_{i=1}^{m} A_{ij}\beta_i \tag{3-3}$$

dos β_i, sendo os escalares $A_{1j}, \ldots, A_{mj}$ as coordenadas de $T\alpha_j$ em relação à base ordenada $\mathcal{B}'$. Conseqüentemente, a transformação T é determinada pelos mn escalares A_{ij} por meio das fórmulas (3-3). A $m \times n$ matriz A definida por $A(i, j) = A_{ij}$ é denominada **a matriz de T em relação ao par de bases ordenadas** $\mathcal{B}$ e $\mathcal{B}'$. Nosso trabalho imediato será o de compreender explìcitamente como a matriz A determina a transformação linear T.

Se $\alpha = x_1\alpha_1 + \ldots + x_n\alpha_n$ é um vetor em V, então

$$\begin{aligned} T\alpha &= T\left(\sum_{j=1}^{n} x_j\alpha_j\right) \\ &= \sum_{j=1}^{n} x_j(T\alpha_j) \\ &= \sum_{j=1}^{n} c_j \sum_{i=1}^{m} A_{ij}\,\beta_i \\ &= \sum_{i=1}^{m}\left(\sum_{j=1}^{n} A_{ij}x_j \quad \beta_i.\right) \end{aligned}$$

Se X é a matriz das coordenadas de α em relação à base ordenada $\mathcal{B}$, o cálculo acima mostra que AX é a matriz das coordenadas do vetor $T\alpha$ em relação à base ordenada $\mathcal{B}'$, uma vez que o escalar

$$\sum_{j=1}^{n} A_{ij}x_j$$

é o elemento da i-ésima linha da matriz coluna AX. Observemos também que se A é uma $m \times n$ matriz arbitrária sôbre o corpo F então

$$T\left(\sum_{j=1}^{n} x_j\alpha_j\right) = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} A_{ij}x\right)_j\beta_i \tag{3-4}$$

define uma transformação linear T de V em W, cuja matriz é A, em relação a $\mathcal{B}$, $\mathcal{B}'$. Resumindo formalmente:

Teorema 11. *Seja* V *um espaço vetorial* n*-dimensional sôbre o corpo* F *e* W *um espaço vetorial* m*-dimensional sôbre F. Seja* $\mathcal{B}$ *uma uma base ordenada de* V *e* $\mathcal{B}'$ *uma base ordenada de* W. *Para cada transformação linear* T *de* V *em* W, *existe uma* m $\times$ n *matriz* A *sôbre o corpo* F, *a matriz de* T *em relação a* $\mathcal{B}$, $\mathcal{B}'$, *tal que*

$$[T\alpha]_{\mathcal{B}'} = A[\alpha]_{\mathcal{B}}$$

para todo vetor α *em* V. *Além disso,* T $\rightarrow$ A *é uma correspondência bijetora entre o conjunto das transformações lineares de* V *em* W *e o conjunto das* m $\times$ n *matrizes sôbre o corpo* F.

Suponhamos agora que T e U sejam transformações lineares de V em W e que a matriz de T em relação a $\mathcal{B}$, $\mathcal{B}'$ seja A e a matriz de V em relação a $\mathcal{B}$, $\mathcal{B}'$ seja B. Qual é a matriz correspondente a $(T + U)$? Isto é respondido fàcilmente, pois se $\alpha = x_1\alpha_1 + \ldots + x_n\alpha_n$

$$\begin{aligned}(T + U)(\alpha) &= T\alpha + U\alpha \\ &= \sum_{i=1}^{m}\left(\sum_{j=1}^{n} A_{ij}x_j\right)\beta_i + \sum_{i=1}^{m}\left(\sum_{j=1}^{n} B_{ij}x_j\right)\beta_i \\ &= \sum_{i=1}^{m}\left(\sum_{j=1}^{n}(A_{ij} + B_{ij})x_j\right)\beta_i.\end{aligned}$$

Assim a matriz de $T + U$ em relação a $\mathcal{B}$, $\mathcal{B}'$ é a soma das matrizes A e B. Anàlogamente, pode-se verificar com facilidade que se c é um escalar arbitrário então a matriz de (cT) é cA. Estas duas observações nos dizem que a correspondência entre transformações lineares e matrizes definida por $\mathcal{B}$ e $\mathcal{B}'$ é linear.

Usando o que sabemos sôbre multiplicação de matrizes, a linearidade da representação por matrizes pode ser vista como segue: Suponhamos que a matriz de T em relação a $\mathcal{B}$, $\mathcal{B}'$ seja A e que a matriz de U em relação ao mesmo par seja B. Se α é um vetor arbitrário em V,

$$\begin{aligned}[T\alpha]_{\mathcal{B}'} &= A[\alpha]_{\mathcal{B}} \\ [U\alpha]_{\mathcal{B}'} &= B[\alpha]_{\mathcal{B}}\end{aligned}$$

e como sabemos que

$$(cA + B)X = cAX + BX$$

temos $$[cT\alpha + U\alpha]_{\mathcal{B}'} = (cA + B)[\alpha]_{\mathcal{B}}$$

Portanto, deve-se ter que $cA + B$ é a matriz de $cT + U$ em relação ao par $\mathcal{B}$, $\mathcal{B}'$.

Teorema 12. *Seja* V *um espaço vetorial* n-*dimensional sôbre o corpo* F *e seja* W *um espaço vetorial* m-*dimensional sôbre* F. *Para cada par de bases ordenadas* $\mathcal{B}$, $\mathcal{B}'$ *de* V *e* W, *respectivamete, a função que associa a uma transformação linear* T *sua matriz em relação a* $\mathcal{B}$, $\mathcal{B}'$ *é um isomorfismo entre o espaço* L(V, W) *e o espaço das* m $\times$ n *matrizes sôbre o corpo* F.

Demonstração. Observamos acima que a função em questão é linear e, como está enunciado no Teorema 11, esta função é injetora e leva $L(V, W)$ sôbre o conjunto das $m \times n$ matrizes.

Estaremos particularmente interessados na representação por matrizes de transformações lineares de um espaço em si mesmo, isto é, operadores lineares sôbre um espaço V. Neste caso, é mais conveniente usar a mesma base ordenada em cada caso, isto é, tomar $\mathcal{B} = \mathcal{B}'$. A matriz representante será então denominada simplesmente a **matriz de T em relação à base ordenada** $\mathcal{B}$. Como êste conceito será muito importante para nós, recordaremos sua definição. Se T é um operador linear sôbre o espaço vetorial V de dimensão finita e $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ é uma base ordenada de V, a matriz de T em relação a $\mathcal{B}$ é a $n \times n$ matriz A cujos elementos A_{ij} são definidos pelas equações

$$(3\text{-}5) \qquad T\alpha_j = \sum_{i=1}^{n} A_{ij}\alpha_i, \quad j = 1, \ldots, n.$$

Deve-se ter sempre em mente que esta matriz que representa T depende da base ordenada $\mathcal{B}$ e que existe uma matriz que representa T em relação a cada base ordenada de V. (Para transformações de um espaço em outro a matriz depende de duas bases ordenadas, uma de V e uma de W.) Para não esquecermos esta dependência, usaremos a notação

$$[T]_{\mathcal{B}}$$

para a matriz do operador linear T em relação à base ordenada β. A maneira como esta matriz e a base ordenada descrevem T é que, para cada α em V,

$$[T\alpha]_{\mathcal{B}} = [T]_{\mathcal{B}}[\alpha]_{\mathcal{B}}.$$

Exemplo 9. Seja V o espaço das $n \times 1$ matrizes-colunas sôbre o corpo F; seja W o espaço das $m \times 1$ matrizes sôbre F; seja A uma $m \times n$ matriz sôbre F, fixa. Seja T a transformação linear de V em W definida por $T(X) = AX$. Seja $\mathcal{B}$ a base ordenada de V análoga à base canônica em F^n; isto é, o i-ésimo vetor em $\mathcal{B}$ é a $n \times 1$ matriz X_i com 1 linha i e com todos os outros elementos nulos. Seja $\mathcal{B}'$ a correspondente base ordenada de W, isto é; o j-ésimo vetor em $\mathcal{B}'$ é a $m \times 1$ matriz Y_j com 1 na linha j e com todos os outros elementos nulos. Então a matriz de T em relação ao par $\mathcal{B}$, $\mathcal{B}'$ é a própria matriz A. Isto é evidente, pois a matriz AX_j é a j-ésima coluna de A.

Exemplo 10. Seja F um corpo e seja T o operador sôbre F^2 definido por

$$T(x_1, x_2) = (x_1, 0).$$

É fácil ver que T é um operador linear sôbre F^2. Seja $\mathcal{B}$ a base ordenada canônica de F^2, $\mathcal{B} = \{\epsilon_1, \epsilon_2\}$. Ora,

$$T\epsilon_1 = T(1, 0) = (1, 0) = 1\epsilon_1 + 0\epsilon_2$$
$$T\epsilon_2 = T(0, 1) = (0, 0) = 0\epsilon_1 + 0\epsilon_2$$

de modo que a matriz de T em relação à base ordenada $\mathcal{B}$ é

$$[T]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Exemplo 11. Seja V o espaço das funções polinomiais de R em R da forma

$$f(x) = c_0 + c_1x + c_2x^2 + c_3x^3$$

isto é, o espaço das funções polinomiais de grau menor ou igual a 3. O operador derivação D do Exemplo 2 leva V em V, pois D diminui o grau. Seja $\mathcal{B}$ a base ordenada de V formada pelas quatro funções f_1, f_2, f_3, f_4 definidas por $f_j(x) = x^{j-1}$. Então

$$\begin{aligned}
(Df_1)(x) &= 0, & Df_1 &= 0f_1 + 0f_2 + 0f_3 + 0f_4 \\
(Df_2)(x) &= 1, & Df_2 &= 1f_1 + 0f_2 + 0f_3 + 0f_4 \\
(Df_3)(x) &= 2x, & Df_3 &= 0f_1 + 2f_2 + 0f_3 + 0f_4 \\
(Df_4)(x) &= 3x^2, & Df_4 &= 0f_1 + 0f_2 + 3f_3 + 0f_4
\end{aligned}$$

de modo que a matriz de D em relação à base ordenada $\mathcal{B}$ é

$$[D]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Vimos o que acontece às matrizes representantes quando as transformações são somadas, a saber, que as matrizes se somam. Gostaríamos agora de perguntar o que acontece quando compomos transformações. Mais especìficamente, sejam V, W e Z espaços vetoriais sôbre o corpo F, de dimensões n, m e p, respectivamente. Seja T uma transformação linear de V em W e U uma transformação linear de W em Z. Suponhamos que existam bases ordenadas

$$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}, \mathcal{B}' = \{\beta_1, \ldots, \beta_m\}, e\ \mathcal{B}'' = \{\gamma_1, \ldots, \gamma_p\}$$

para os espaços V, W e Z, respectivamente. Seja A a matriz de T em relação ao par $\mathcal{B}$, $\mathcal{B}'$ e B a matriz de U em relação ao par $\mathcal{B}'$, $\mathcal{B}''$. É fácil ver então que a matriz C da transformação UT em relação ao par $\mathcal{B}$, $\mathcal{B}''$ é o produto de B por A, pois se α é um vetor arbitrário em V

$$[T\alpha]_{\mathcal{B}'} = A[\alpha]_{\mathcal{B}}$$

$$[U(T\alpha)]_{\mathcal{B}''} = B[T\alpha]_{\mathcal{B}'}$$

e então $$[(UT)(\alpha)]_{\mathcal{B}''} = BA[\alpha]_{\mathcal{B}}$$

logo, pela definição e unicidade da matriz representante, temos, necessàriamente, $C = BA$. Isto também pode ser visto efetuando os cálculos

$$\begin{aligned}(UT)(\alpha_j) &= U(T\alpha_j)\\ &= U\left(\sum_{k=1}^{m} A_{kj}\beta_k\right)\\ &= \sum_{k=1}^{m} A_{kj}(U\beta_k)\\ &= \sum_{k=1}^{m} A_{kj}\sum_{i=1}^{p} B_{ik}\gamma_i\\ &= \sum_{i=1}^{p}\left(\sum_{k=1}^{m} B_{ik}A_{kj}\right)\gamma_i\end{aligned}$$

de modo que temos

$$C_{ij} = \sum_{k=1}^{m} B_{ik}A_{kj}. \tag{3-6}$$

A definição (3-6) de multiplicação de matrizes foi motivada por meio de operações sôbre as linhas de uma matriz. Vê-se aqui que uma motivação bastante forte para a definição encontra-se na composição de transformações lineares. Resumindo formalmente:

Teorema 13. *Sejam* V, W *e* Z *espaços vetoriais de dimensão finita sôbre o corpo* F; *seja* T *uma transformação linear de* V *em* W *e* U *uma transformação linear de* W *em* Z. *Se* $\mathcal{B}$, $\mathcal{B}'$ *e* $\mathcal{B}''$ *são bases ordenadas dos espaços* V, W *e* Z, *respectivamente, se* A *é matriz de* T *em relação ao par* $\mathcal{B}$, $\mathcal{B}'$ *e* B *é a matriz de* U *em relação ao par* $\mathcal{B}'$, $\mathcal{B}''$, *então a matriz da composta* UT *em relação ao par* $\mathcal{B}$, $\mathcal{B}''$ *é a matriz produto* C = BA.

É importante notar que se T e U são operadores lineares sôbre um espaço V e se estamos usando apenas uma base ordenada $\mathcal{B}$, então o Teorema 11 toma a forma simples $[UT]_{\mathcal{B}} = [U]_{\mathcal{B}}\,[T]_{\mathcal{B}}$. Assim, neste caso, a correspondência que $\mathcal{B}$ determina entre operadores lineares e matrizes é não sòmente um isomorfismo de espaço vetorial mas conserva também produtos. Uma conseqüência simples disto é que o operador linear T é inversível se, e sòmente se, $[T]_{\mathcal{B}}$ é uma matriz inversível. De fato, o operador idêntico I é representado pela matriz unidade em relação a qualquer base ordenada, portanto

$$UT = TU = I$$

é equivalente a

$$[U]_{\mathcal{B}}[T]_{\mathcal{B}} = [T]_{\mathcal{B}}[U]_{\mathcal{B}} = I.$$

Evidentemente, quando T é inversível

$$[T^{-1}]_{\mathfrak{B}} = [T]_{\mathfrak{B}}^{-1}.$$

Gostaríamos agora de perguntar o que acontece com as matrizes representantes quando mudamos a base ordenada. Para efeito de simplicidade, consideraremos esta equação apenas para operadores lineares sôbre um espaço V, de modo que possamos usar uma única base ordenada. A questão específica é a seguinte: seja T um operador linear sôbre o espaço de dimensão finita V e sejam a

$$\mathfrak{B} = \{\alpha_1, \ldots, \alpha_n\} \text{ e } \mathfrak{B}' = \{\alpha'_1, \ldots, \alpha'_n\}$$

duas bases ordenadas de V. Qual é a relação entre as matrizes $[T]_{\mathfrak{B}}$ e $[T]_{\mathfrak{B}'}$? Como observamos no Capítulo 2, existe uma única $n \times n$ matriz (inversível) P tal que

(3-7) $$[\alpha]_{\mathfrak{B}} = P[\alpha]_{\mathfrak{B}'}$$

para todo vetor α em V. Por definição

(3-8) $$[T\alpha]_{\mathfrak{B}} = [T]_{\mathfrak{B}}[\alpha]_{\mathfrak{B}}.$$

Aplicando (3-7) ao vetor $T\alpha$ temos

(3-9) $$[T\alpha]_{\mathfrak{B}} = P[T\alpha]_{\mathfrak{B}'}.$$

Combinando (3-7), (3-8) e (3-9), obtemos

$$[T]_{\mathfrak{B}} P[\alpha]_{\mathfrak{B}'} = P[T\alpha]_{\mathfrak{B}'}$$

ou $$P^{1}[T]_{\mathfrak{B}} P[\alpha]_{\mathfrak{B}'} = [T\alpha]_{\mathfrak{B}'}$$

e então é necessário que

(3-10) $$[T]_{\mathfrak{B}'} = P^{1}[T]_{\mathfrak{B}}P.$$

Isto responde nossa pergunta.

Antes de enunciarmos formalmente êste resultado, observemos um fato. Existe um único operador linear U que leva $\mathfrak{B}$ sôbre $\mathfrak{B}'$ definido por

$$U\alpha_j = \alpha'_j, \quad j = 1, \ldots, n.$$

Êste operador U é inversível uma vez que leva uma base de V sôbre uma base de V. A matriz P (acima) é exatamente a matriz do operador U em relação à base ordenada $\mathfrak{B}$. De fato, P é definida por

$$\alpha'_j = \sum_{i=1}^{n} P_{ij}\alpha_i$$

e como $U\alpha_j = \alpha'_j$, esta equação pode ser escrita como

$$U\alpha_j = \sum_{i=1}^{n} P_{ij}\alpha_i.$$

Portanto $P = [U]_{\mathfrak{B}'}$ por definição.

Teorema 14. *Seja* V *um espaço vetorial de dimensão finita sôbre o corpo* F *e sejam*

$$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\} \text{ e } \mathcal{B}' = \{\alpha'_1, \ldots, \alpha'_n\}$$

bases ordenadas de V. *Suponhamos que* T *seja um operador linear sôbre* V. *Se* P *é a* n × n *matriz que exprime as coordenadas de cada vetor de* V *em relação a* $\mathcal{B}$ *em têrmos de suas coordenadas em relação a* $\mathcal{B}'$, *então*

$$[T]_{\mathcal{B}'} = P^{-1}[T]_{\mathcal{B}}P.$$

Alternativamente, se U *é o operador inversível sôbre* V *definido por* $U\alpha_j = \alpha'_j$, $j = 1, \ldots, n$, *então*

$$[T]_{\mathcal{B}'} = [U]_{\mathcal{B}}^{-1}[T]_{\mathcal{B}}[U]_{\mathcal{B}}.$$

Exemplo 12. Seja T o operador linear sôbre R^2 definido por $T(x_1, x_2) = (x_1, 0)$. No Exemplo 10 mostramos que a matriz de T em relação à base ordenada canônica $\mathcal{B} = \{\epsilon_1, \epsilon_2\}$ é

$$[T]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Suponhamos que $\mathcal{B}'$ seja a base ordenada de R^2 formada pelos vetores $\epsilon'_1 = (1, 1)$, $\epsilon'_2 = (2, 1)$. Então

$$\begin{aligned} \epsilon'_1 &= \epsilon_1 + \epsilon_2 \\ \epsilon'_2 &= 2\epsilon_1 + \epsilon_2 \end{aligned}$$

de modo que P é a matriz

$$P = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}.$$

Efetuando cálculos simples obtemos

$$P^{-1} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$$

Assim

$$\begin{aligned} [T]_{\mathcal{B}'} &= P^{-1}[T]_{\mathcal{B}}P \\ &= \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -1 & -2 \\ 1 & 2 \end{bmatrix}. \end{aligned}$$

Podemos verificar fàcilmente que isto está correto porque

$$\begin{aligned} T\epsilon'_1 &= (1, 0) = -\epsilon'_1 + \epsilon'_2 \\ T\epsilon'_2 &= (2, 0) = -2\epsilon'_1 + 2\epsilon'_2. \end{aligned}$$

Exemplo 13. Seja V o espaço das funções polinomiais de R em R, de "grau" menor ou igual a 3. Como no Exemplo 11, seja D o operador derivação sôbre V e seja

$$\mathcal{B} = \{f_1, f_2, f_3, f_4\}$$

a base ordenada de V definida por $f_i(x) = x^{i-1}$. Seja t um número real e definamos $g_i(x) = (x + t)^{i-1}$, isto é,

$$\begin{aligned} g_1 &= f_1 \\ g_2 &= tf_1 + f_2 \\ g_3 &= t^2f_1 + 2tf_2 + f_3 \\ g_4 &= t^3f_1 + 3t^2f_2 + 3tf_3 + f_4. \end{aligned}$$

Como se pode ver fàcilmente, a matriz

$$P = \begin{bmatrix} 1 & t & t^2 & t^3 \\ 0 & 1 & 2t & 3t^2 \\ 0 & 0 & 1 & 3t \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

é inversível com

$$P^{-1} = \begin{bmatrix} 1 & -t & t^2 & -t^3 \\ 0 & 1 & -2t & 3t^2 \\ 0 & 0 & 1 & -3t \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

portanto decorre que $\mathcal{B}' = \{g_1, g_2, g_3, g_4\}$ é uma base ordenada de V. No Exemplo 11, ficamos sabendo que a matriz de D em relação à base ordenada $\mathcal{B}$ é

$$[D]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

A matriz de D em relação à base ordenada $\mathcal{B}'$ é portanto

$$\begin{aligned} P^{-1}[D]_{\mathcal{B}}P &= \begin{bmatrix} 1 & -t & t^2 & t^3 \\ 0 & 1 & -2t & 3t^2 \\ 0 & 0 & 1 & -3t \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & t & t^2 & t^3 \\ 0 & 1 & 2t & 3t^2 \\ 0 & 0 & 1 & 3t \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -t & t^2 & t^3 \\ 0 & 1 & -2t & 3t^2 \\ 0 & 0 & 1 & -3t \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2t & 3t^2 \\ 0 & 0 & 2 & 6t \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Assim D é representado pela mesma matriz em relação às bases ordenadas $\mathcal{B}$ e $\mathcal{B}'$. Evidentemente, isto pode ser visto um pouco mais diretamente pois

$$\begin{aligned} Dg_1 &= 0 \\ Dg_2 &= g_1 \\ Dg_3 &= 2g_2 \\ Dg_4 &= 3g_3. \end{aligned}$$

Êste exemplo ilustra um fato interessante. Se se conhece a matriz de um operador linear em relação a alguma base ordenada $\mathcal{B}$ e quer-se determinar a matriz em relação a outra base ordenada $\mathcal{B}'$, freqüentemente o que mais convém é efetuar a mudança de coordenadas usando a matriz inversível P; contudo, pode ser muito mais fácil determinar a matriz representante recorrendo diretamente à sua definição.

Definição. *Sejam* A *e* B n $\times$ n *matrizes* (*quadradas*) *sôbre o corpo* F. *Dizemos que* B *é* **semelhante a** A **sôbre** F *se existe uma* n $\times$ n *matriz inversível* P *sôbre* F *tal que* B $= P^{-1}$AP.

De acôrdo com o Teorema 14, temos o seguinte: Se V é um espaço vetorial n-dimensional sôbre F e $\mathcal{B}$ e $\mathcal{B}'$ são duas bases ordenadas de V, então, para cada operador linear T sôbre V, a matriz $B = [T]_{\mathcal{B}'}$, é semelhante à matriz $A = [T]_{\mathcal{B}}$. O argumento também vale no outro sentido. Suponhamos que A e B sejam $n \times n$ matrizes e que B seja semelhante a A. Seja V um espaço vetorial n-dimensional arbitrário sôbre F e seja $\mathcal{B}$ uma base ordenada de V. Seja T o operador linear sôbre V que é representada em relação à base $\mathcal{B}$ por A. Se $B = P^{-1}AP$, seja $\mathcal{B}'$ a base ordenada de V obtida de $\mathcal{B}$ por meio de P, isto é,

$$\alpha_j' = \sum_{i=1}^{n} P_{ij}\alpha_i.$$

Então, a matriz de T em relação à base ordenada $\mathcal{B}'$ será B.

Assim, a afirmação de que B é semelhante a A significa que em cada espaço n-dimensional sôbre F as matrizes A e B representam a mesma transformação linear em relação a duas bases ordenadas (possìvelmente) distintas.

Notemos que tôda $n \times n$ matriz A é semelhante a si mesma, bastando tomar $P = I$; se é semelhante a A, então A é semelhante a B, pois, $B = P^{-1}AP$ implica que $A = (P^{-1})^{-1}BP^{-1}$; se B é semelhante a A e C é semelhante a B, então C é semelhante a A, pois $B = P^{-1}AP$ e $C = Q^{-1}BQ$ implicam que $C = (PQ)^{-1}A(PQ)$. Assim, a semelhança é uma relação de equivalência sôbre o conjunto das matrizes sôbre o corpo F. Notemos também que a única matriz semelhante à matriz unidade I é a própria I e que a única matriz semelhante à matriz nula é a própria matriz nula.

Exercícios

1. Seja T o operador linear sôbre C^2 definido por $T(x_1, x_2) = (x_1, 0)$. Seja $\mathcal{B}$ a base ordenada canônica de C^2 e seja $\mathcal{B}' = \{\alpha_1, \alpha_2\}$ a base ordenada definida por $\alpha_1 = (1, i)$, $\alpha_2 = (-i, 2)$.

(a) Qual é a matriz de T em relação ao par $\mathcal{B}$, $\mathcal{B}'$?
(b) Qual é a matriz de T em relação ao par $\mathcal{B}'$, $\mathcal{B}$?
(c) Qual é a matriz de T em relação à base ordenada $\mathcal{B}'$?
(d) Qual é a matriz de T em relação à base ordenada $\{\alpha_2, \alpha_1\}$?

2. Seja T a transformação linear de R^3 em R^2 definida por

$$T(x_1, x_2, x_3) = (x_1 + x_2, 2x_3 - x_1).$$

(a) Se $\mathcal{B}$ é a base ordenada canônica de R^3 e $\mathcal{B}'$ é a base ordenada canônica de R^2, qual é a matriz de T em relação ao par $\mathcal{B}$, $\mathcal{B}'$?

(b) Se $\mathcal{B} = \{\alpha_1, \alpha_2, \alpha_3\}$ e $\mathcal{B}' = \{\beta_1, \beta_2\}$, sendo

$$\alpha_1 = (1, 0, -1),\ \alpha_2 = (1,1,1),\ \alpha_3 = (1, 0, 0),\ \beta_1 = (0, 1),\ \beta_2 = (1, 0)$$

qual é a matriz de T em relação ao par $\mathcal{B}$, $\mathcal{B}'$?

3. Seja T um operador linear sôbre F^n, seja A a matriz de T em relação à base ordenada canônica de F^n e seja W o subespaço de F^n gerado pelos vetores-colunas de A. Qual é a relação de W com T?

4. Seja V um espaço vetorial bidimensional sôbre o corpo F e seja Q uma base ordenada de V. Se T é um operador linear sôbre V e

$$[T]_{\mathcal{B}} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

demonstrar que $T^2 - (a + d)T + (ad - bc)I = 0$.

5. Seja T o operador linear sôbre R^3, cuja matriz em relação à base ordenada canônica é

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & 3 & 4 \end{bmatrix}.$$

Determinar uma base da imagem de T e uma base do núcleo de T.

6. Seja T o operador linear sôbre R^2 definido por

$$T(x_1, x_2) = (-x_2, x_1).$$

(a) Qual é a matriz de T em relação à base ordenada canônica de R^2?

(b) Qual é a matriz de T em relação à base ordenada $\mathcal{B} = \{\alpha_1, \alpha_2\}$ sendo $\alpha_1 = (1, 2)$ e $\alpha_2 = (1, -1)$?

(c) Demonstrar que para todo número real c o operador $(T - cI)$ é inversível.

(d) Demonstrar que se $\mathcal{B}$ é uma base ordenada qualquer de R^2 e $[T]_{\mathcal{B}} = A$, então $A_{12}A_{21} \neq 0$.

7. Seja T o operador linear sôbre R^3 definido por

$$T(x_1, x_2, x_3) = (3x_1 + x_3, -2x_1 + x_2, -x_1 + 2x_2 + 4x_3).$$

(a) Qual é a matriz de T em relação à base ordenada canônica de R^3?
(b) Qual é a matriz de T em relação à base ordenada

$$\{\alpha_1, \alpha_2, \alpha_3\}$$

sendo $\alpha_1 = (1, 0, 1)$, $\alpha_2 = (-1, 2, 1)$, e $\alpha_3 = (2, 1, 1)$?

(c) Demonstrar que T é inversível e dar uma regra para T^{-1} como a que define T.

8. Seja θ um número real. Demonstrar que as duas matrizes seguintes são semelhantes sôbre o corpo dos números complexos:

$$\begin{bmatrix} \cos\theta & -\text{sen}\,\theta \\ \text{sen}\,\theta & \cos\theta \end{bmatrix} \qquad \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix}$$

(*Sugestão*: Seja T o operador linear sôbre C^2 que é representado pela primeira matriz em relação à base ordenada canônica. Determinar então vetores α_1 e α_2 tais que $T\alpha_1 = e^{i\theta}\alpha_1$, $T\alpha_2 = e^{-i\theta}\alpha_2$, e $\{\alpha_1, \alpha_2\}$ seja uma base.)

9. Seja V um espaço vetorial de dimensão finita sôbre o corpo F e sejam S e T operadores lineares sôbre V. Perguntamos: quando é que existem bases ordenadas $\mathcal{B}$ e $\mathcal{B}'$ de V tais que $[S]_{\mathcal{B}} = [T]_{\mathcal{B}'}$? Demonstrar que tais bases existem se, e sòmente se, existe um operador linear inversível U sôbre V tal que $T = USU^{-1}$. (*Esbôço de demonstração*: Se $[S]_{\mathcal{B}} = [T]_{\mathcal{B}'}$, seja U o operador que leva $\mathcal{B}$ sôbre $\mathcal{B}'$; mostrar que $S = UTU^{-1}$. Reciprocamente se $T = USU^{-1}$ para algum U inversível, seja $\mathcal{B}$ uma base ordenada arbitrária de V e seja Q' sua imagem por meio de U. Mostrar então que $[S]_{\mathcal{B}} = [T]_{\mathcal{B}'}$.)

10. Vimos que o operador linear T sôbre R^2 definido por $T(x_1, x_2) = (x_1, 0)$ é representado em relação à base ordenada canônica pela matriz

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Êste operador satisfaz $T^2 = T$. Demonstrar que se S é um operador linear sôbre R^2 tal que $S^2 = S$, então $S = 0$ ou $S = I$ ou então existe uma base ordenada $\mathcal{B}$ de R^2 tal que $[S]_{\mathcal{B}} = A$ (acima).

11. Seja W o espaço das $n \times 1$ matrizes-colunas sôbre um corpo F. Se A é uma $n \times n$ matriz sôbre F, então A define um operador linear L_A sôbre W por meio da multiplicação à esquerda: $L_A(X) = AX$. Demonstrar que todo operador linear sôbre W é a multiplicação à esquerda por alguma $n \times n$ matriz, isto é, é L_A para algum A.

Suponhamos agora que V seja um espaço vetorial-dimensional sôbre o corpo F e seja $\mathcal{B}$ uma base ordenada de V. Para cada α em V definamos $U\alpha = [\alpha]_{\mathcal{B}}$. Demonstrar que U é um isomorfismo de V em W. Se T é um operador linear sôbre V, então UTU^{-1} é um operador linear sôbre W. Isto significa que UTU^{-1} é a multiplicação à esquerda por alguma $n \times n$ matriz A. Qual é a matriz A?

12. Seja V um espaço vetorial n-dimensional sôbre o corpo F e seja $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ uma base ordenada de V.

(a) De acôrdo com o Teorema 1, existe um único operador linear T sôbre V tal que

$$T\alpha_j = \alpha_{j+1}, \quad j = 1, \ldots, n-1, \quad T\alpha_n = 0.$$

Qual é a matriz A de T em relação à base ordenada $\mathcal{B}$?

(b) Demonstrar que $T^n = 0$ mas $T^{n-1} \neq 0$.

(c) Seja S um operador linear arbitrário sôbre V tal que $S^n = 0$ mas $S^{n-1} \neq 0$. Demonstrar que existe uma base ordenada $\mathcal{B}'$ de V tal que a matriz de S em relação à base ordenada $\mathcal{B}'$ é a matriz A da parte (a).

(d) Demonstrar que se M e N são $n \times n$ matrizes sôbre F tais que $M^n = N^n = 0$ mas $M^{n-1} \neq 0 \neq N^{n-1}$, então M e N são semelhantes.

13. Sejam V e W espaços vetoriais de dimensão finita sôbre o corpo F e Seja T uma transformação linear de V em W. Se

$$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\} \text{ e } \mathcal{B}' = \{\beta_1, \ldots, \beta_m\}$$

são bases ordenadas de V e W, respectivamente, definamos as transformações lineares $E^{p,q}$ como na demonstração do Teorema 5: $E^{p,q}(\alpha_i) = \delta_{iq}\beta_p$. Então as $E^{p,q}$, $1 \leq p \leq m$, $1 \leq q \leq n$, formam uma base de (LV, W), e portanto

$$T = \sum_{p=1}^{m} \sum_{q=1}^{n} A_{pq}E^{p,q}$$

para certos escalares A_{pq} (as coordenadas de T em relação a esta base de $L(V, W)$. Mostrar que a matriz com elementos $A(p,q) = A_{pq}$ é exatamente a matriz de T em relação ao par $\mathcal{B}$, $\mathcal{B}'$.

3.5 Funcionais Lineares

No estudo de um espáço vetorial V sôbre o corpo F, as transformações lineares de V em F têm interêsse suficiente para merecerem um nome especial, **funcionais lineares.** Num sentido rigoroso, não deveríamos falar de transformações lineares de V em F, porque F não é um espaço vetorial. Deveríamos falar de transformações lineares de V em F^1; no entanto, da mesma forma que dizemos que V é um espaço vetorial sôbre F, diremos que F é um espaço vetorial sôbre F, isto é, usaremos o símbolo F tanto para o corpo como para o espáço F^1.

Definição. *Se* V *é um espaço vetorial sôbre o corpo* F, *um* **funcional linear sôbre** V *é uma transformação linear de* V *em* F. *O espaço* L(V, F) *de tôdas as funcionais lineares será indicado por* V* *e denominado o* **espaço dual** *de* V.

Se se quer começar do início, um funcional linear sôbre V é uma função f de V em F tal que

$$f(c\alpha + \beta) = cf(\alpha) + f(\beta)$$

para quaisquer vetores α e β em V e qualquer escalar c em F. O conceito de "funcional linear" é importante e tem muitas aplicações. Teremos tempo de tocar em apenas algumas destas aplicações. Descreveremos abaixo pelo menos a forte relação entre funcionais lineares, equações lineares e coordenadas.

Exemplo 14. Seja V um espaço vetorial de dimensão finita sôbre o corpo F e seja $\mathcal{B}$ uma base ordenada de V. A função f_i que associa a cada vetor α em V a i-ésima coordenada de α em relação à base ordenada $\mathcal{B}$ é um funcional linear sôbre V. Discutiremos abaixo êste exemplo com alguns detalhes.

Exemplo 15. Eis um exemplo importante de funcional linear. Seja n um inteiro positivo e F um corpo. Seja V o espaço vetorial das $n \times n$ matrizes sôbre F. Se A está em V, o **traço** de A é o escalar

$$\text{tr}A = A_{11} + A_{22} + \ldots + A_{nn}.$$

A função traço é um funcional linear sôbre V, pois

$$\begin{aligned}\text{tr}(cA + B) &= \sum_{i=1}^{n} (cA_{ii} + B_{ii})\\ &= c \sum_{i=1}^{n} A_{ii} + \sum_{i=1}^{n} B_{ii}\\ &= c[\text{tr}(A)] + \text{tr}(B).\end{aligned}$$

Exemplo 16. Seja V o espaço vetorial das funções polinomiais do corpo F em F. Seja t um elemento de F. Se definirmos

$$L_t(p) = p(t)$$

então L_t será um funcional linear sôbre V. Geralmente descreve-se isto dizendo que, para cada t, o cálculo do valor em t é um funcional linear sôbre o espaço das funções polinomiais.

Se V é de dimensão finita podemos obter uma descrição bastante explícita do espaço dual V^*. Pelo Teorema 5 sabemos alguma coisa sôbre o espaço V^*, a saber, que

$$\dim V^* = \dim V.$$

Seja $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ uma base de V. De acôrdo com o Teorema 1, existe (para cada i) um único funcional linear f_i sôbre V tal que

$$f_i(\alpha_j) = \delta_{ij}. \tag{3—11}$$

Desta maneira obtemos a partir de $\mathcal{B}$ um conjunto de n funcionais lineares distintos $f_1, \ldots, f_n$ sôbre V. Êstes funcionais também são linearmente independentes. De fato, suponhamos que

$$f = \sum_{i=1}^{n} c_i f_i. \tag{3—12}$$

Então

$$\begin{aligned}f(\alpha_j) &= \sum_{i=1}^{n} c_i f_i(\alpha_j)\\ &= \sum_{i=1}^{n} c_i \delta_{ij}\\ &= c_j.\end{aligned}$$

Em particular, se f é o funcional nulo, $f(\alpha_j) = 0$ para cada j, logo, os escalares c_j são todos nulos. Ora, $f_1, \ldots, f_n$ são n funcionais linearmente independentes e como sabemos que V^* tem dimensão n, devemos ter que $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ é uma base de V^*. Esta base é dita a **base dual** de $\mathcal{B}$.

Teorema 15. *Seja* V *um espaço vetorial de dimensão finita sôbre o corpo* F *e seja* $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ *uma base de* V. *Então existe uma única base dual* $\mathcal{B}'' = \{f_1, \ldots, f_n\}$ *de* V* *tal que* $f_i(\alpha_j) = \delta_{ij}$. *Para cada funcional linear* f *sôbre* V temos

$$f = \sum_{i=1}^{n} f(\alpha_i)f_i \tag{3—13}$$

e para cada vetor α *em* V *temos*

$$\alpha = \sum_{i=1}^{n} f_i(\alpha)\alpha_i. \tag{3—14}$$

Demonstração. Demonstramos acima que existe uma única base que é "dual" a $\mathcal{B}$. Se f é um funcional linear sôbre V então f é alguma combinação linear (3—12) dos f_i e, como observamos por (3—12), os escalares c_j são dados necessàriamente por $c_j = f(\alpha_j)$. Anàlogamente, se

$$\alpha = \sum_{i=1}^{n} x_i\alpha_i$$

é um vetor em V, então

$$\begin{aligned} f_j(\alpha) &= \sum_{i=1}^{n} x_i f_j(\alpha_i) \\ &= \sum_{i=1}^{n} x_i\delta_{ij} \\ &= x_j \end{aligned}$$

de modo que a única expressão de α como combinação linear dos α_i é

$$\alpha = \sum_{i=1}^{n} f_i(\alpha)\alpha_i.$$

(3—14) nos fornece agora uma maneira satisfatória de descrever o que é a base dual. Ela nos diz que se $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ é uma base ordenada de V e $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ é a base dual, então fi é exatamente a função que associa a cada vetor α em V a i-ésima coordenada de α em relação à base ordenada $\mathcal{B}$. Assim podemos também chamar as f_i de funções coordenadas para $\mathcal{B}$. A fórmula (3—13) quando combinada com (3—14) nos diz o seguinte: Se f está em V^* e $f(\alpha_i) = a_i$, então se

$$\alpha = x_1\alpha_1 + \ldots + x_n\alpha_n$$

temos

$$f(\alpha) = a_1x_1 + \ldots + a_nx_n. \tag{3—15}$$

Em outras palavras, se fixamos uma base ordenada $\mathcal{B}$ de V e descrevemos cada vetor em V por sua n-upla de coordenadas $(x_1, \ldots, x_n)$ em relação a $\mathcal{B}$, então, todo funcional linear sôbre V é da forma (3—15). O leitor certamente notará que esta é exatamente a representação de f que obtemos aplicando a f nossa representação matricial geral de transformações lineares; isto é,

$$[a_1\ a_2\ \ldots\ a_n]$$

é a matriz de f em relação às bases ordenadas $\mathcal{B}$ de V e $\{1\}$ de F. É particularmente importante notar que todo funcional linear f sôbre o espaço F^n das n-uplas é da forma

$$f(x_1, \ldots, x_n) = a_1x_1 + \ldots + a_nx_n$$

onde $a_1, \ldots, a_n$ são n escalares fixos em F.

Uma questão que surge naturalmente neste ponto é se tôda base de V^* é a base dual de alguma base de V. Isto ocorre e, para ver por que, investigaremos ràpidamente o espaço $V^{**} = (V^*)^*$ dual do espaço V^*. Se α é um vetor arbitrário em V, a função $L\alpha$ sôbre V^* definida por

$$L_\alpha(f) = f(\alpha) \tag{3—16}$$

é um funcional linear sôbre V^*, pois

$$\begin{aligned} L_\alpha(cf + g) &= (cf + g)(\alpha) \\ &= cf(\alpha) + g(\alpha) \\ &= cL_\alpha(f) + L_\alpha(g). \end{aligned}$$

Demonstraremos agora que todo funcional linear sôbre V^* é da forma (3—16) para algum α em V.

Lema. *Seja* V *um espaço vetorial de dimensão finita sôbre o corpo* F. *Se α é um vetor não-nulo em* V, *existe um funcional linear* f *sôbre* V *tal que* $f(\alpha) \neq 0$.

Demonstração. Como $\alpha \neq 0$ existe uma base $\{\alpha_1, \ldots, \alpha_n\}$ de V tal que $\alpha_2 = \alpha$. Se $\{f_1, \ldots, f_n\}$ é a base dual, então $f_1(\alpha) \neq 0$.

Teorema 16. *Seja V um espaço vetorial de dimensão finita sôbre o corpo F. Para cada vetor α em V definamos*

$$L_\alpha(f) = f(\alpha), f \text{ em } V^*.$$

A aplicação $\alpha \to L_\alpha$ é então um isomorfismo de V em V^{**}.

Demonstração. Mostramos que para todo α a função L_α é linear. Suponhamos que α e β estejam em V, c esteja em F e seja $\gamma = c\alpha + \beta$. Então, para todo f em V^*

$$\begin{aligned} L_\gamma(f) &= f(\gamma) \\ &= f(c\alpha + \beta) \\ &= cf(\alpha) + f(\beta) \\ &= cL_\alpha(f) + L_\beta(f), \end{aligned}$$

portanto
$$L_\gamma = cL_\alpha + L_\beta.$$

Isto mostra que a aplicação $\alpha \to L_\alpha$ é uma transformação linear de V em V^{**}. Esta transformação é não singular, pois de acôrdo com o Lema acima $L_\alpha = 0$ se, e sòmente se, $\alpha = 0$. Portanto $\alpha \to L_\alpha$ é uma transformação linear não-singular de V em V^{**}, e como

$$\dim V^{**} = \dim V^* = \dim V$$

o Teorema 9 nos diz que esta transformação é inversível, sendo portanto um isomorfismo de V em V^{**}.

Corolário. *Seja V um espaço vetorial de dimensão finita sôbre o corpo F. Se L é um funcional linear sôbre o espaço V* dual de V, então existe um único vetor α em V tal que*

$$L(f) = f(\alpha)$$

para todo f em V.*

Corolário. *Seja V um espaço vetorial de dimensão finita sôbre o corpo F. Tôda base de V^* é a dual de alguma base de V.*

Demonstração. Seja $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ uma base de V^*. Pelo Teorema 15, existe uma base $\{L_1, \ldots, L_n\}$ de V^{**} tal que

$$L_i(f_j) = \delta_{ij}.$$

Usando o corolário acima, para cada i existe um vetor α_i em V tal que

$$L_i(f) = f(\alpha_i)$$

para tôda f em V^*, isto é, tal que $L_i = L_{\alpha_i}$. Decorre imediatamente que $\{\alpha_1, \ldots, \alpha_n\}$ é uma base de V e que B^* é a dual desta base.

Exemplo 17. Seja V o espaço vetorial das funções polinomiais de R em R, de grau menor ou igual a 2. Sejam t_1, t_2 e t_3 três quaisquer números reais *distintos* e seja

$$L_i(p) = p(t_i).$$

Então L_1, L_2 e L_3 são funcionais lineares sôbre V. Êstes funcionais são linearmente independentes; de fato, suponhamos que

$$L = c_1L_1 + c_2L_2 + c_3L_3.$$

Se $L = 0$, isto é, se $L(p) = 0$ para todo p_2 em V, então, aplicando L às "funções" polinomiais particulares 1, x, x^2, obtemos

$$\begin{aligned} c_1 + c_2 + c_3 &= 0 \\ t_1c_1 + t_2c_2 + t_3c_3 &= 0 \\ t_1^2c_1 + t_2^2c_2 + t_3^2c_3 &= 0. \end{aligned}$$

Disto decorre que $c_1 = c_2 = c_3 = 0$, pois (como mostram alguns cálculos simples) a matriz

$$\begin{bmatrix} 1 & 1 & 1 \\ t_1 & t_2 & t_3 \\ t_1^2 & t_2^2 & t_3^2 \end{bmatrix}$$

é inversível quando t_1, t_2 e t_3 são distintos. Portanto, os L_i são independentes e como V tem dimensão 3, êstes funcionais formam uma base de V^*. Qual é a base de V cuja dual é esta? Tal base $\{p_1, p_2, p_3\}$ de V precisa satisfazer

$$L_i(p_j) = \delta_{ij}$$

ou

$$p_j(t_i) = \delta_{ij}.$$

Estas funções polinomiais são, como se vê fàcilmente,

$$\begin{aligned} p_1(x) &= \frac{(x - t_2)(x - t_3)}{(t_1 - t_2)(t_1 - t_3)} \\ p_2(x) &= \frac{(x - t_1)(x - t_3)}{(t_1 - t_2)(t_1 - t_3)} \\ p_3(x) &= \frac{(x - t_1)(x - t_2)}{(t_3 - t_1)(t_3 - t_2)} \end{aligned}$$

A base $\{p_1, p_2, p_3\}$ de V é interessante porque, de acôrdo com (3—14), temos, para cada p em V,

$$p = p(t_1)p_1 + p(t_2)p_2 + p(t_3)p_3.$$

Assim, se c_1, c_2 e c_3 são números reais arbitrários, existe exatamente uma função polinomial p sôbre R de grau menor ou igual a 2 que satisfaz $p(t_j) = c_j$, $j = 1, 2, 3$. Esta função polinomial é $p = c_1p_1 + c_2p_2 + c_3p_3$.

Exercícios

1. Em R^3, sejam $\alpha_1 = (1, 0, 1)$, $\alpha_2 = (0, 1, -2)$, $\alpha_3 = (-1, -1, 0)$.

(a) Se f é um funcional linear sôbre R_3 tal que

$$f(\alpha_1) = 1, \quad f(\alpha_2) = -1, \quad f(\alpha_3) = 3,$$

e se $\alpha = (a, b, c)$, determinar $f(\alpha)$.

(b) Descrever explìcitamente um funcional linear f sôbre R^3 tal que

$$f(\alpha_1) = f(\alpha_2) = 0 \quad \text{mas } f(x_3) \neq 0.$$

(c) Seja f um funcional linear arbitrário tal que

$$f(\alpha_1) = f(\alpha_2) = 0 \text{ e } f(\alpha_3) \neq 0.$$

Se $\alpha = (2, 3, -1)$, mostrar que $f(\alpha) \neq 0$.

2. Seja $\mathcal{B} = \{\alpha_1, \alpha_2, \alpha_3\}$ a base de C^3 definida por

$$\alpha_1 = (1, 0, -1), \alpha_2 = (1, 1, 1), \alpha_3 = (2, 2, 0).$$

Determinar a base dual de $\mathcal{B}$.

3. Se A e B são $n \times n$ matrizes sôbre o corpo F, mostrar que traço (AB) = traço (BA). Mostrar depois que matrizes semelhantes têm o mesmo traço.

4. Seja V o espaço vetorial das funções polinomiais p de R em R que têm grau menor ou igual a 2, ou seja,

$$p(x) = c_0 + c_1x + c_2x^2.$$

Definamos três funcionais lineares sôbre V por

$$f_1(p) = \int_0^1 p(x)\,dx \qquad f_2(p) = \int_0^2 p(x)\,dx \qquad f_3(p) = \int_0^{-1} p(x)\,dx.$$

Mostrar que $\{f_1, f_2, f_3\}$ é uma base de V^* exibindo a base de V^* da qual ela é dual.

5. Se A e B são $n \times n$ matrizes complexas, mostrar que é impossível ter-se $AB - BA = I$.

6. De acôrdo com o Exercício 3, matrizes semelhantes têm o mesmo traço. Assim, podemos definir o traço de um operador linear sôbre um espaço vetorial de dimensão finita como sendo o traço de qualquer matriz que represente o operador em relação a uma base ordenada. Isto está bem definido uma vez que tôdas as matrizes representantes de um operador são semelhantes.

Seja agora V o espaço das 2×2 matrizes sôbre o corpo F e seja P uma 2×2 matriz fixa. Seja T o operador linear sôbre V definido por $T(A) = PA$. Demonstrar que traço (T) = 2 traço (P).

7. Sejam f_1 e f_2 funcionais lineares sôbre um espaço vetorial n-dimensional V. Seja N_j o núcleo de f_j, $j = 1, 2$. Suponhamos que $f_1 \neq 0 \neq f_2$ e que $N_1 \neq N_2$. Determinar as dimensões de cada um dos espaços N_1, N_2, $N_1 \cap N_2$ e $N_1 + N_2$.

8. Sejam m e n inteiros positivos e F um corpo. Sejam $f_1, \ldots, f_m$ funcionais lineares sôbre F^n. Para α em F^n definamos

$$T\alpha = (f_1(\alpha), \ldots, f_m(\alpha)).$$

Mostrar que T é uma transformação linear de F^n em F^m. Mostrar então que tôda transformação linear de F^n em F^m é da forma acima, para certos $f_1, \ldots, f_m$.

9. Seja V um espaço vetorial sôbre o corpo F. Seja f um funcional linear fixo não-nulo sôbre V e seja N o núcleo de f. Fixemos um vetor α_0 em V que não esteja em N. Demonstrar que para cada α em V existe um escalar c e um vetor β em N tal que $\alpha = c\,\alpha_0 + \beta$. Demonstrar que c e β são únicos.

10. Seja V um espaço vetorial de dimensão finita sôbre o corpo F e seja W um subespaço de V. Se f é um funcional linear sôbre W, demonstrar que existe um funcional linear g sôbre V tal que $g(\alpha) = f(\alpha)$ para todo α no subespaço W.

11. Seja F um subcorpo do corpo dos números complexos e seja V um espaço vetorial arbitrário sôbre F. Suponhamos que f e g sejam funcionais lineares sôbre V tais que a função h, definida por $h(\alpha) = f(\alpha)g(\alpha)$ também seja um funcional linear sôbre V. Demonstrar que ou $f = 0$ ou $g = 0$.

12. Mostrar que o funcional traço sôbre $n \times n$ matrizes é o único no seguinte sentido: se W é o espaço das $n \times n$ matrizes sôbre o corpo F e se f é um funcional linear sôbre W tal que $f(AB) = f(BA)$ para tôdas as A e B em W, então f é um múltiplo escalar da função traço. Se, além disso, $f(I) = n$, então f é a função traço.

13. Seja F um subcorpo do corpo dos números complexos e seja V um espaço vetorial de dimensão finita sôbre F. Se $\alpha_1, \ldots, \alpha_m$ são um número finito de vetores em V, todos diferentes do vetor nulo, demonstrar que existe um funcional linear f sôbre V tal que

$$f(\alpha_i) \neq 0, \quad i = 1, \ldots, m.$$

14. Sejam $f_1, \ldots, f_r$ funcionais lineares sôbre um espaço vetorial V e seja N_j o núcleo de f_j, $j = 1, \ldots, r$. Seja f outro funcional linear sôbre V, com núcleo N. Demonstrar que f é uma combinação linear de $f_1, \ldots, f_r$, se e sòmente se, N contém a interseção $\bigcap_{j=1}^{r} N_j$. (Demonstrar primeiro para $r = 1$.)

15. Seja W o espaço das $n \times n$ matrizes sôbre o corpo F e seja W_0 o subespaço gerado pelas matrizes c da forma $C = AB - BA$.

Demonstrar que W_0 é exatamente o subespaço das matrizes que têm traço nulo. (*Sugestão*: Qual é a dimensão do espaço das matrizes de traço nulo? Usar as "matrizes unitárias", isto é, as matrizes com exatamente um elemento não-nulo, para construir um número suficiente de matrizes linearmente independentes da forma $AB - BA$.)

3.6 Anuladores

O objetivo principal desta seção é demonstrar que um subespaço W de um espaço vetorial V de dimensão finita é determinado pelo conjunto dos funcionais lineares f em V^* que se anulam em W. Mais precisamente, se dim $V = n$ e dim $W = r$, demonstraremos que existem $(n - r)$ funcionais lineares $f_1, \ldots, f_{n-r}$ em V^* tais que W consiste exatamente dos vetores α em V que satisfazem $f_j(\alpha) = 0$, $j = 1, \ldots, n - r$.

Definição. *Se* V *é um espaço vetorial sôbre o corpo* F *e* S *é um subconjunto de* V, *o* **anulador** *de* S *é o conjunto* S^0 *dos funcionais lineares* f *sôbre tais* V *que* $f(\alpha) = 0$ *para todo* α *em* S.

Deve ficar claro para o leitor que S^0 é um subespaço de V^*, seja S um subespaço de V ou não. Se S é o conjunto formado apenas pelo vetor nulo, então, $S^0 = V^*$. Se $S = V$, então S^0 é o subespaço nulo de V^*. (Isto é fácil de ver para o caso em que V é de dimensão finita.)

Teorema 17. *Seja* V *um espaço vetorial de dimensão sôbre o corpo* F *e seja* W *um subespaço de* V. *Então*

$$\dim W + \dim W^0 = \dim V.$$

Demonstração. Seja k a dimensão de W e $\{\alpha_1, \ldots, \alpha_k\}$ uma base de W. Tomemos vetores $\alpha_{k+1}, \ldots, \alpha_n$ em V tais que $\{\alpha_1, \ldots, \alpha_n\}$ seja uma base de V. Seja $\{f_1, \ldots, f_n\}$ a base de V^* que é a dual desta base de V. Afirmamos que $\{f_{k+1}, \ldots, f_n\}$ é uma base do anulador de W^0. Certamente f_i pertence a W^0 para $i > k + 1$, porque

$$f_i(\alpha_j) = \delta_{ij}$$

e $\delta_{ij} = 0$ se $i \geq k + 1$ e $j \leq k$; disto decorre que, para $i \geq k + 1$, $f_i(\alpha) = 0$ sempre que α seja uma combinação linear de $\alpha_1, \ldots, \alpha_k$. Os funcionais $f_{k+1}, \ldots, f_n$ são independentes, portanto basta mostrar que êles geram W^0. Suponhamos que f esteja em V^*. Ora,

$$f = \sum_{i=1}^{n} f(\alpha_i) f_i$$

de modo que, se f está em W^0, temos $f(\alpha_i) = 0$ para $i \leq k$ e

$$f = \sum_{i=k+1}^{n} f(\alpha_i) f_i.$$

Mostramos que se k é a dimensão de W e n é a dimensão de V, então a dimensão de W^0 é $(n - k)$.

Um fato razoàvelmente elementar mas importante sôbre um espaço V de dimensão finita é que se W_1 e W_2 são dois subespaços de V tais que $W_1^0 = W_2^0$, então $W_1 = W_2$. Uma maneira de ver isto é investigar o anulador do anulador de um subconjunto S de V. Se S é um subconjunto de V, então, rigorosamente falando, $S^{00} = (S^0)^0$ é um subespaço do espaço bidual V^{**}; contudo, identificando V^{**} com V por meio do isomorfismo $\alpha \rightarrow L_\alpha$, podemos considerar S^{00} como um subespaço de V, isto é, conjunto de todos α em V tais que $f(\alpha) = 0$ para todo f em S^0.

Teorema 18. *Seja* V *um espaço vetorial de dimensão finita sôbre o corpo* F *e seja* W *um subespaço de* V. *Então* $W^{00} = W$.

Demonstração. Certamente W é um subconjunto de W^{00}, pois se α está em W, então $f(\alpha) = 0$ para todo f em W^0. De acôrdo com o Teorema 17,

$$\begin{aligned} \dim W \ + \dim W^0 &= \dim V \\ \dim W^0 + \dim W^{00} &= \dim V^* = \dim V \end{aligned}$$

do que segue que $\dim W = \dim W^{00}$. Como W está contido em W^{00} temos, necessàriamente, que $W = W^{00}$.

Corolário. *Se* S *é um subconjunto arbitrário de* V, *então* S^{00} *é o subespaço gerado por* S.

Demonstração. Seja W o subconjunto gerado por S. Então $S^0 = W^0$; pois se f está em W^0 é evidente que f está em S^0 e, recìprocamente, se f está em S^0 então $f(\alpha) = 0$ sempre que α é uma combinação linear de vetores em S, isto é, f está em W^0. Pelo Teorema 18, $W^{00} = W$ e como $W^0 = S^0$ temos $S^{00} = W$.

Corolário. *Se* W_1 *e* W_2 *são subespaços de* V, *então* $W_2 = W_1$ *se e sòmente se* $W_1^0 = W_2^0$.

Consideremos ràpidamente sistemas de equações lineares homogêneas do ponto de vista de funcionais lineares. Suponhamos ter um sistema de equações lineares

$$\begin{array}{ccc} A_{11}x_1 + \dots + A_{1n}x_n & = & 0 \\ \vdots & & \vdots \\ A_{m1}x_1 + \dots + A_{mn}x_n & = & 0 \end{array}$$

cujas soluções queiramos determinar. Se indicarmos por f_i, $i = 1, \dots, m$, o funcional linear sôbre F^n definido por

$$f_i(x_1, \dots, x_n) = A_{i1}x_1 + \dots + A_{in}x_n$$

então estamos procurando o subespaço de F^n constituído por todos α tais que

$$f_i(\alpha) = 0, \qquad i = 1, \dots, m.$$

Em outras palavras, estamos procurando o subespaço anulado por $f_1, \dots, f_m$. A linha-redução da matriz dos coeficientes nos fornece um método sistemático para determinar êsse subespaço. A n-upla $(A_{i1}, \dots, A_{in})$ dá as coordenadas do funcional linear f_i em relação à base que é dual da base canônica de F^n. O espaço-linha da matriz dos coeficientes pode portanto ser considerado como o espaço de funcionais lineares gerado por $f_1, \dots, f_m$. O espaço-solução é o subespaço anulado por êsse espaço de funcionais. Pelo Teorema 18, isto é o mesmo que dizer que o espaço de funcionais gerado pelas f_i é o anulador do espaço das soluções. O que o Teorema 17 nos diz nesse contexto é que a soma do pôsto-linha de A com a dimensão do espaço-solução é n. Em outras palavras, se A é uma $m \times n$ matriz de pôsto-linha r, então o espaço das soluções do sistema homogêneo $AX = 0$ tem dimensão $(n - r)$.

Agora podemos considerar o sistema de equações do ponto de vista "dual". Isto é, suponhamos que nos sejam dados m vetores em F^n

$$\alpha_i = (A_{i1}, \ldots, A_{in})$$

e queiramos determinar o anulador do subespaço gerado por êsses vetores. Como um funcional linear arbitrário sôbre F^n tem a forma

$$f(x_1, \ldots, x_n) = c_1 x_1 + \ldots + c_n x_n$$

a condição para que f esteja nesse anulador é que

$$\sum_{j=1}^{n} A_{ij} c_j = 0, \qquad i = 1, \ldots, m$$

isto é, que $(c_1, \ldots, c_n)$ seja uma solução do sistema $AX = 0$. Sob êste ponto de vista, a linha-redução nos dá um método sistemático para determinar o anulador do subespaço gerado por um dado conjunto finito de vetores em F^n.

Exemplo 18. Seja W o subespaço de R^5 gerado pelos vetores

$$\alpha_1 = (2, -2, 3, 4, -1), \qquad \alpha_3 = (0, 0, -1, -2, 3),$$
$$\alpha_2 = (-1, 1, 2, 5, 2), \qquad \alpha_4 = (1, -1, 2, 3, 0).$$

Como se descreve W^0, o anulador de W? Formemos a 4×5 matriz A com vetores-linhas $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ e determinemos a matriz R linha-reduzida à forma em escada que é linha-equivalente a A;

$$A = \begin{bmatrix} 2 & -2 & 3 & 4 & -1 \\ -1 & 1 & 2 & 5 & 2 \\ 0 & 0 & -1 & -2 & 3 \\ 1 & -1 & 2 & 3 & 0 \end{bmatrix} \to R = \begin{bmatrix} 1 & -1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Se f é um funcional linear sôbre R^5,

$$f(x_1, \ldots, x_5) = \sum_{j=1}^{5} c_j x_j$$

então f está em W^0 se, e sòmente se, $f(\alpha_i) = 0$, $i = 1, 2, 3, 4$, isto é, se, e sòmente se,

$$\sum_{j=1}^{5} A_{ij} c_j = 0, \qquad 1 \leq i \leq 4.$$

Isto é equivalente a

$$\sum_{j=1}^{5} R_{ij}c_j = 0, \qquad 1 \leq i \leq 3$$

ou

$$\begin{aligned} c_1 - c_2 - c_4 &= 0 \\ c_3 + 2c_4 &= 0 \\ c_5 &= 0. \end{aligned}$$

Obtemos todos êsses funcionais lineares atribuindo valores arbitrários a c_2 e c_4, por exemplo, $c_2 = a$ e $c_4 = b$ e depois calculando os correspondentes $c_1 = a + b$, $c_3 = -2b$, $c_5 = 0$. Portanto, W^0 consiste de todos os funcionais lineares f da forma

$$f(x_1, x_2, x_3, x_4, x_5) = (a + b)x_1 + ax_2 - 2bx_3 + bx_4.$$

A dimensão de W^0 é 2 e pode-se encontrar uma base $\{f_1, f_2\}$ de W^0 tomando primeiro $a = 1$, $b = 0$ e depois $a = 0$, $b = 1$:

$$\begin{aligned} f_1(x_1, \ldots, x_5) &= x_1 + x_2 \\ f_2(x_1, \ldots, x_5) &= x_1 - 2x_3 + x_4. \end{aligned}$$

Um f arbitrário em W^0 é $f = af_1 + bf_2$. O fato de que $W^{00} = W$ significa exatamente que W consiste de todos os vetores $\alpha = (x_1, x_2, x_3, x_4, x_5)$ que satisfazem $f_1(\alpha) = f_2(\alpha) = 0$, isto é, todos α tais que

$$\begin{aligned} x_1 + x_2 &= 0 \\ x_1 - 2x_3 + x_4 &= 0 \end{aligned}$$

O leitor deve notar que esta é a descrição explícita de W que demos no Capítulo 2, (2-25).

Exercícios

1. Sejam $\alpha_1 = (1, 0, -1, 2)$ e $\alpha_2 = (2, 3, 1, 1)$ e seja W o subespaço de R^3 gerado por α_1 e α_2. Quais funcionais lineares f,

$$f(x_1, x_2, x_3, x_4) = c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4,$$

estão no anulador de W?

2. Seja W o subespaço de R^5 gerado pelos vetores

$$\begin{gathered} \alpha_1 = \epsilon_1 + 2\epsilon_2 + \epsilon_3, \quad \alpha_2 = \epsilon_2 + 3\epsilon_3 + 3\epsilon_4 + \epsilon_5 \\ \alpha_3 = \epsilon_1 + 4\epsilon_2 + 6\epsilon_3 + 4_4\epsilon + \epsilon_5. \end{gathered}$$

Determinar uma base de W^0.

3. Seja V o espaço vetorial das 2×2 matrizes sôbre o corpo dos números reais e seja

$$B = \begin{bmatrix} 2 & -2 \\ -1 & 1 \end{bmatrix}.$$

Seja W o subespaço de V que consiste de tôdas A tais que $AB = 0$. Seja f um funcional linear sôbre V que esteja no anulador de W. Suponhamos que $f(I) = 0$ e $f(C) = 3$, sendo I a 2×2 matriz unidade e

$$C = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Determinar $f(B)$.

4. Seja F um subcorpo do corpo dos números complexos. Definamos n funcionais lineares sôbre $F^n (n \geq 2)$ por

$$f_k(x_1, \ldots, x_n) = \sum_{j=1}^{n} (k - j)x_j, \quad 1 \leq k \leq n.$$

Qual é a dimensão do subespaço anulado por $f_1, \ldots, f_n$?

5. Seja n um inteiro positivo e F um corpo. Seja W o conjunto de todos os vetores $(x_1, \ldots, x_n)$ em F^n tais que $x_1 + \ldots + x_n = 0$.

(a) Demonstrar que W^0 consiste dos funcionais lineares f da forma

$$f(x_1, \ldots, x_n) = c \sum_{j=1}^{n} x_j.$$

(b) Mostrar que o espaço W^* dual de W pode ser identificado de maneira "natural" com os funcionais lineares

$$f(x_1, \ldots, x_n) = c_1x_1 + \ldots + c_nx_n$$

sôbre F^n que satisfazem $c_1 + \ldots + c_n = 0$.

6. Se V é um espaço vetorial sôbre o corpo F, definamos um **hiperplano** em V como sendo o núcleo de um funcional linear não-nulo sôbre V. Se V é de dimensão finita, demonstrar que todo subespaço de V é a interseção de um número finito de hiperplanos.

7. Sejam W_1 e W_2 subespaços de um espaço vetorial V de dimensão finita.

(a) Demonstrar que $(W_1 + W_2)^0 = W_1^0 \cap W_2^0$.
(b) Demonstrar que $(W_1 \cap W_2)^0 = W_1^0 + W_2^0$.

3.7 A Transposta de uma Transformação Linear

Suponhamos que existam dois espaços vetoriais V e W sôbre o corpo F e uma transformação linear T de V em W. Então T induz uma transformação linear de W^* em V^*, como segue. Suponhamos que g seja um funcional linear sôbre W e seja

$$f(\alpha) = g(T\alpha) \tag{3-17}$$

para cada α em V. Então (3-17) define uma função f de V em F, a saber, a composta de T, uma função de V em W, com g, uma função de W em F. Como T e g são ambas lineares, o Teorema 6 nos diz que f também é linear, isto é, f é um funcional linear sôbre V. Assim T nos fornece uma função T^t que a cada funcional linear g sôbre W faz corresponder um funcional linear $f = T^tg$ sôbre V, definido por (3-17). Notemos também que T^t é na verdade uma transformação linear de W^* em V^*; de fato, se g_1 e g_2 estão em W^* e c é um escalar

$$\begin{aligned}[T^t(cg_1 + g_2)](\alpha) &= (cg_1 + g_2)(T\alpha)\\ &= cg_1(T\alpha) + g_2(T\alpha)\\ &= c(T^tg_1)(\alpha) + (T^tg_2)(\alpha)\end{aligned}$$

de modo que $T^t(cg_1 + g_2) = cT^tg_1 + T^tg_2$. Façamos um resumo.

Teorema 19. *Sejam* V *e* W *espaços vetoriais sôbre o corpo* F. *Para cada transformação linear* T *de* V *em* W *existe uma única transformação linear* T^t *de* W* *em* V* *tal que*

$$(T^tg)(\alpha) = g(T\alpha)$$

para todos g *em* W* *e* α *em* V.

Denominaremos T^t a **transposta** de T. Esta transformação T^t é freqüentemente denominada a adjunta de T; no entanto, não usaremos essa terminologia.

Teorema 20. *Sejam* V *e* W *espaços vetoriais sôbre o corpo* F *e seja* T *uma transformação linear de* V *em* W. *O núcleo de* T^t *é o anulador da imagem de* T. *Se* V *e* W *são de dimensão finita, então*

(i) *pôsto* (T^t) = *pôsto* (T)

(ii) *a imagem de* T^t *é o anulador do núcleo de* T.

Demonstração. Se g está em W^*, então por definição

$$(T^tg)(\alpha) = g(T\alpha)$$

para todo α em V. A afirmação de que g está no núcleo de T^t significa que $g(T\alpha) = 0$ para todo α em V. Portanto o núcleo de T^t é exatamente o anulador da imagem de T.

Suponhamos que V e W sejam de dimensão finita, digamos dim $V = n$ e dim $W = n$. (i) Seja r o pôsto de T, isto é, a dimensão da imagem de T. Pelo Teorema 17, o anulador da imagem de T tem dimensão $(m - r)$. Pela primeira afirmação dêste Teorema, a nulidade de T^t deve ser $(m - r)$. Mas como T^t é uma transformação linear sôbre um espaço vetorial m-dimensional, o pôsto de T é $m - (m - r) = r$, logo T e T^t têm o mesmo pôsto. (ii) Seja N o núcleo de T. Todo funcional que está na imagem de T^t está no anulador de N; de fato, suponhamos que $f = T^tg$ para algum g em W^*; então, se α está em N

$$f(\alpha) = (T^tg)(\alpha) = g(T\alpha) = g(0) = 0.$$

Ora, a imagem de T^t é um subespaço do espaço N^0, e

$$\dim N^0 = n - \dim N = \text{pôsto } (T) = \text{pôsto } (T^t)$$

de modo que a imagem de T^t deve ser exatamente N^0.

Teorema 21. *Sejam* V *e* W *espaços vetoriais de dimensão finita sôbre o corpo* F. *Seja* $\mathcal{B}$ *uma base ordenada de* W *com base dual* $\mathcal{B}^*$ *e seja* $\mathcal{B}'$ *uma base ordenada de* W *com base dual* β'^*. *Seja* T *uma transformação linear de* V *em* W; *seja* A *a matriz de* T *em relação a* $\mathcal{B}$, $\mathcal{B}'$ *e seja* B *a matriz de* T^t *em relação a* $\mathcal{B}'^*$, $\mathcal{B}^*$. *Então* $B_{ij} = A_{ji}$.

Demonstração. Sejam

$$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}, \qquad \mathcal{B}' = \{\beta_1, \ldots, \beta_m\},$$
$$\mathcal{B}^* = \{f_1, \ldots, f_n\}, \qquad \mathcal{B}'^* = \{g_1, \ldots, g_m\}.$$

Por definição,

$$T\alpha_j = \sum_{i=1}^{n} A_{ij}\beta_i, \qquad j = 1, \ldots, n$$

$$T^tg_j = \sum_{i=1}^{n} B_{ij}f_i, \qquad j = 1, \ldots, m.$$

Por outro lado,

$$\begin{aligned}(T^tg_j)(\alpha_i) &= g_j(T\alpha_i)\\ &= g_j\left(\sum_{k=1}^{m} A_{ki}\beta_k\right)\\ &= \sum_{k=1}^{m} A_{ki}g_j(\beta_k)\\ &= \sum_{k=1}^{m} A_{ki}\delta_{jk}\\ &= A_{ji}.\end{aligned}$$

Para qualquer funcional linear f sôbre V

$$f = \sum_{i=1}^{m} f(\alpha_i)f_i.$$

Aplicando esta fórmula ao funcional $f = T^tg_j$ e usando o fato de que $(g_j)T^t(\alpha_i) = A_{ji}$ temos

$$T^tg_j = \sum_{i=1}^{n} A_{ji}f_i$$

do que decorre imediatamente que $B_{ij} = A_{ji}$.

Definição. *Se* A *é uma* m $\times$ n *matriz sôbre o corpo* F, *a* **transposta** *de* A *é a* n $\times$ m *matriz* A^t *definida por* $A^t_{ij} = A_{ji}$.

O Teorema 21 afirma portanto que se T é uma transformação linear de V em W, cuja matriz em relação a algum par de bases é A, então a transformação transposta T^t é representada em relação ao par de bases duais pela matriz transposta A^t.

Teorema 22. *Seja* A *uma* m $\times$ n *matriz arbitrária sôbre o corpo* F. *Então o pôsto-linha de* A *é igual ao pôsto-coluna de* A.

Demonstração. Seja $\mathcal{B}$ a base ordenada canônica de F^n e $\mathcal{B}'$ a base ordenada canônica de F^m. Seja T a transformação linear de F^n em F^m tal que a matriz de T em relação ao par $\mathcal{B}$, $\mathcal{B}'$ seja A, isto é,

$$T(x_1, \ldots, x_n) = (y_1, \ldots, y_m)$$

onde

$$y_i = \sum_{j=1}^{n} A_{ij}x_j.$$

O pôsto-coluna de A é o pôsto da transformação T, pois a imagem de T consiste de tôdas as m-uplas que são combinações lineares dos vetores-colunas de A.

Em relação às bases duais $\mathcal{B}'^*$ e $\mathcal{B}^*$, a aplicação transposta T^t é representada pela matriz A^t. Como as colunas de A^t são as linhas de A, vemos, pelo mesmo raciocínio, que o pôsto-linha de A (o pôsto-coluna de A^t) é igual ao pôsto de T^t. Pelo Teorema 20, T e T^t têm o mesmo pôsto, logo o pôsto-linha de A é igual ao pôsto-coluna de A.

Vemos agora que se A é uma $m \times n$ matriz sôbre F e T é a transformação linear de F^n em F^m definida acima, então

$$\text{pôsto } (T) = \text{pôsto-linha } (A) = \text{pôsto-coluna } (A)$$

e denominaremos êste número simplesmente o **pôsto** de A.

Exemplo 19. Êste exemplo será de natureza geral — mais uma discussão que um exemplo. Seja V um espaço vetorial n-dimensional sôbre o corpo F e seja T um operador linear sôbre V. Suponhamos que $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ seja uma base ordenada de V. A matriz de T em relação à base $\mathcal{B}$ é definida como sendo a $n \times n$ matriz A tal que

$$T\alpha_j = \sum_{j=1}^{m} A_{ij}\alpha_i;$$

em outras palavras, A_{ij} é a i-ésima coordenada do vetor $T\alpha_j$ em relação à base ordenada $\mathcal{B}$. Se $\{f_1, \ldots, f_n\}$ é a base dual de $\mathcal{B}$, isto pode ser enunciado simplesmente como

$$A_{ij} = f_i(T\alpha_j).$$

Vejamos o que acontece quando mudamos de base. Suponhamos que

$$\mathcal{B}' = \{\alpha_1', \ldots, \alpha_n'\}$$

seja outra base ordenada de V, com base dual $\{f_1', \ldots, f_n'\}$. Se B é a matriz de T em relação à base ordenada $\mathcal{B}'$, então

$$B_{ij} = (f_i' T\alpha_j').$$

Seja U o operador linear inversível tal que $U\alpha_j = \alpha'_j$. Então, a transposta de U é dada por $U^t f_i' = f_i$. É fácil verificar que por ser U inversível, U^t também o é e $(U^t)^{-1} = (U^{-1})^t$. Assim $f_i' = (U^{-1})^t f_i$, $i = 1, \ldots, n$. Portanto

$$\begin{aligned} B_{ij} &= [(U^{-1})^t f_i]\,(T\alpha_j') \\ &= f_i(U^{-1}T\alpha_j') \\ &= f_i(U^{-1}TU\alpha_j). \end{aligned}$$

O que significa isto? Bem, $f_i(U^{-1}TU\alpha_j)$ é o elemento i, j da matriz de $U^{-1}TU$ em relação à base ordenada $\mathcal{B}$. Nossos cálculos acima mostram que êste escalar é também o elemento i, j da matriz de T em relação à base ordenada $\mathcal{B}'$. Em outras palavras,

$$\begin{aligned} [T]_{\mathcal{B}'} &= [U^{-1}TU]_{\mathcal{B}} \\ &= [U^{-1}]_{\mathcal{B}}[T]_{\mathcal{B}}[U]_{\mathcal{B}} \\ &= [U]_{\mathcal{B}}^{-1}[T]_{\mathcal{B}}[U]_{\mathcal{B}} \end{aligned}$$

e esta é exatamente a fórmula de mudança de base que deduzimos anteriormente.

Exercícios

1. Seja F um corpo e seja f o funcional linear sôbre F^2 definido por $f(x_1, x_2) = ax_1 + bx_2$. Para cada um dos operadores lineares T seguintes, sendo $g = T^t f$, determinar $g(x_1, x_2)$.

   (a) $T(x_1, x_2) = (x_1, 0)$;
   (b) $T(x_1, x_2) = (-x_2, x_1)$;
   (c) $T(x_1, x_2) = (x_1 - x_2, x_1 + x_2)$.

2. Seja V o espaço vetorial das funções polinomiais sôbre o corpo dos números reais. Sejam a e b números reais fixos e seja f o funcional linear sôbre V definido por

$$f(p) = \int_a^b p(x)\,dx.$$

Se D é o operador derivação sôbre V, o que é $D^t f$?

3. Seja V o espaço das $n \times n$ matrizes sôbre um corpo F e seja B uma $n \times n$ matriz fixa. Se T é o operador linear sôbre V definido por $T(A) = AB - BA$ e se f é a função traço, o que é $T^t f$?

4. Seja V um espaço vetorial de dimensão finita sôbre o corpo F e seja T um operador linear sôbre V. Seja c um escalar e suponhamos que exista um vetor não-nulo α em V tal que $T\alpha = c\alpha$. Demonstrar que existe um funcional linear não-nulo f sôbre V tal que $T^t f = cf$.

5. Seja A uma $m \times n$ matriz com elementos *reais*. Demonstrar que $A = 0$ se, e sòmente se, traço $(A^t A) = 0$.

6. Seja n um inteiro positivo e seja V o espaço das funções polinomiais sôbre o corpo dos números reais, de grau menor ou igual a n, isto é, funções da forma

$$f(x) = c_0 + c_1 x + \ldots + c_n x^n.$$

Seja D o operador derivação sôbre V. Determinar uma base do núcleo do operador transposto D^t.

7. Seja V um espaço vetorial de dimensão finita sôbre o corpo F. Mostrar que $T \rightarrow T^t$ é um isomorfismo de $L(V, V)$ em $L(V^*, V^*)$.

8. Seja V o espaço vetorial das $n \times n$ matrizes sôbre o corpo F.

(a) Se B é uma $n \times n$ matriz fixa, definamos uma função f_B sôbre V por $f_B(A) =$ traço $(B^t A)$. Mostrar que f_B é um funcional linear sôbre V!

(b) Mostrar que todo funcional linear sôbre V é da forma acima, isto é, é f_B para algum B.

(c) Mostrar que $B \rightarrow f_B$ é um isomorfismo de V em V^*.

CAPÍTULO 4

POLINÔMIOS

4.1 Álgebras

O objetivo dêste capítulo é estabelecer algumas das propriedades básicas da álgebra dos polinômios sôbre um corpo. A discussão será facilitada se introduzirmos primeiro o conceito de uma álgebra linear sôbre um corpo.

Definição. *Seja* F *um corpo.* **Uma álgebra linear sôbre o corpo** F *é um espaço vetorial* $\mathfrak{A}$ *sôbre* F *com uma operação adicional, dita* **multiplicação de vetores**, *que associa a cada par de vetores* α, β *em* $\mathfrak{A}$ *um vetor* $\alpha\beta$ *em* $\mathfrak{A}$ *dito o* **produto** *de* α *por* β *de maneira tal que*

(1) *a multiplicação é associativa,*

$$\alpha(\beta\gamma) = (\alpha\beta)\gamma$$

(2) *a multiplicação é distributiva em relação à adição,*

$$\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma \qquad \text{e } (\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$$

(3) *para cada escalar* c *em* F,

$$c(\alpha\beta) = (c\alpha)\beta = \alpha(c\beta).$$

Se existir um elemento 1 *em* $\mathfrak{A}$ *tal que* $1\alpha = \alpha 1 = \alpha$ *para todo* α *em* $\mathfrak{A}$, *denominaremos* $\mathfrak{A}$ *uma* **álgebra linear com elemento unidade sôbre** F *e denominaremos* 1 *o* **elemento unidade** *de* $\mathfrak{A}$. *A álgebra* $\mathfrak{A}$ *é dita* **comutativa** *se* $\alpha\beta = \beta\alpha$ *para todos* α *e* β *em* $\mathfrak{A}$.

Exemplo 1. O conjunto das $n \times n$ matrizes sôbre um corpo, com as operações usuais, é uma álgebra linear com elemento unidade; em particular o próprio corpo é um álgebra com elemento unidade. Esta álgebra não é comutativa se $n \geq 2$. O corpo é (evidentemente) comutativo.

Exemplo 2. O espaço dos operadores lineares sôbre um espaço vetorial, com o composto como o produto, é uma álgebra linear com elemento unidade. Ela é comutativa se e sòmente se o espaço é unidimensional.

O leitor talvez tenha tido alguma experiência com o produto escalar e com o produto vetorial em R^3. Se é o caso, deve verificar que nenhum dêsses produtos é o do tipo descrito na definição de uma álgebra linear. O produto escalar, como o nome indica, associa cada par de vetores um escalar, e assim, certamente, não é o tipo de produto que ora discutimos. O produto vetorial associa de fato um vetor a cada par de vetores em R^3; no entanto, esta não é uma multiplicação associativa.

O restante desta seção será dedicado à construção de uma álgebra que é significativamente diferente das álgebras dos exemplos precedentes. Seja F um corpo e S o conjunto dos inteiros não-negativos. Pelo Exemplo 3 do Capítulo 2, o conjunto de tôdas as funções de S em F é um espaço vetorial sôbre F. Indicaremos êste espaço vetorial por F^∞. Os vetores em F^∞ são portanto seqüências infinitas $f = \{f_0, f_1, f_2, \ldots\}$ de escalares f_1 em F. Se $g = \{g_0, g_1, g_2, \ldots\}$, com g_i em F e A, b são escalares em F, $af + bg$ é a seqüência infinita dada por

$$(4\text{-}1) \qquad af + bg = \{af_0 + bg_0, af_1 + bg_1, af_2 + bg_2, \ldots\}.$$

Definamos um produto em F^∞ associado a cada par de vetores f, g em F^∞ o vetor fg que é dado por

$$(4\text{-}2) \qquad (fg)_n = \sum_{i=0}^{n} f_i g_{n-i}, \qquad n = 0, 1, 2, \ldots.$$

Assim

$$fg = \{f_0g_0, f_0g_1 + f_1g_0, f_0g_2 + f_1g_1 + f_2g_0, \ldots\}$$

e como

$$(gf)_n = \sum_{i=0}^{n} g_i f_{n-i} = \sum_{i=0}^{n} f_i g_{n-i} = (fg)_n$$

para $n = 0, 1, 2, \ldots$, segue que a multiplicação é comutativa, isto é, $fg = gf$. Se h também pertence a F^∞, então

$$[(fg)h]_n = \sum_{i=0}^{n} (fg)_i h_{n-i}$$

$$= \sum_{i=0}^{n} \left(\sum_{j=0}^{i} f_j g_{i-j} \quad h_{n-i} \right)$$

$$= \sum_{i=0}^{n} \sum_{j=0}^{i} f_j g_{i-j} h_{n-i}$$

$$= \sum_{j=0}^{n} f_j \sum_{i=0}^{j} g_i h_{n-i-j}$$

$$= \sum_{j=0}^{n} f_j (gh)_{n-j} = [f(gh)]_n$$

para $n = 0, 1, 2, \ldots$, de modo que

(4-3) $$(fg)h = f(gh).$$

Deixamos a cargo do leitor verificar que a multiplicação definida por (4-2) satisfaz (2) e (3) da definição de uma álgebra linear e que o vetor $1 = \{1, 0, 0, \ldots\}$ funciona como um elemento unidade de F^∞. Então, F^∞, com as operações definidas acima, é uma álgebra linear comutativa com elemento unidade sôbre o corpo F.

O vetor $\{0, 1, 0, \ldots, 0, \ldots\}$ desempenha um papel notável no que segue e indicá-lo-emos consistentemente por x. Em todo êste capítulo, x nunca será usado para indicar um elemento do corpo F. O produto de x por si mesmo n vêzes será indicado por x^n e colocaremos $x^0 = 1$. Então

$$x^2 = \{0, 0, 1, 0, \ldots\}, \qquad x^3 = \{0, 0, 0, 1, 0, \ldots\}$$

e, de maneira geral, para cada inteiro $k \geq 0$, $(x^k)_k = 1$ e $(x^k)_n = 0$ para todos inteiros não-negativos $n \neq k$. Concluindo esta seção, observemos que o conjunto formado por $1, x, x^2, \ldots$ é independente e infinito. Assim a álgebra F^∞ não é de dimensão finita.

4.2 A Álgebra dos Polinômios

Estamos agora em condições de definir um polinômio sôbre o corpo F.

Definição. *Seja* F[x] *o subespaço de* F^∞ *gerado pelos vetores* 1, x, x^2, ... *Um elemento de* F[x] *é dito um* **polinômio sôbre** F. *O leitor deve notar que um polinômio não é mesmo o tipo de objeto que uma função polinomial.*

Como $F[x]$ consiste de tôdas as combinações lineares (finitas) de x e suas potências, um vetor não-nulo f em F^∞ é um polinômio se e sòmente se existe um inteiro $n \geq 0$ tal que $f_n \neq 0$ e tal que $f_k = 0$ para todos os inteiros $k > n$; êste inteiro (quando existe) é evidentemente únioo e é denominado o **grau** de f. Indicamos o grau de um polinômio por gr(f) e não atribuímos nenhum grau ao polinômio nulo. Se f é um polinômio não-nulo de grau n temos que

(4—4) $$f = f_0 x^0 + f_1 x + f_2 x^2 + \ldots + f_n x^n, \quad f_n \neq 0.$$

Os escalares $f_0, f_1, \ldots, f_n$ são às vêzes ditos os **coeficientes** de f e podemos dizer que f é um polinômio com coeficientes em F. Denominaremos polinômios da forma cx^0 **polinômios constantes** e freqüentemente indicaremos xc^0 por c. Um polinômio não-nulo f de grau n tal que $f_n = 1$ é dito um polinômio **unitário**.

Teorema 1. *Sejam* f *e* g *polinômios não-nulos sôbre* F. *Então*

(a) fg *é um polinômio não-nulo*;

(b) $gr(fg) = gr(f) + gr(g)$;

(c) fg *é um polinômio unitário se* f *e* g *são ambos polinômios unitários*;

(d) fg *é um polinômio constante se, e sòmente se,* f *e* g *são polinômios constantes*;

(e) *se* $f + g \neq 0$,

$$gr(f + g) \leq \text{máx } (gr)(f)\ gr(g)).$$

Demonstração. Suponhamos que f tenha grau m e que g tenha grau n. Se k é um inteiro não-negativo.

$$(fg)_{m+n+k} = \sum_{i=0}^{m+n+k} f_i g_{m+n+k-i}.$$

Para que $f_i g_{m+n+k-i} \neq 0$, é necessário que $i < m$ e $m + n + k - i \leq n$. Logo, é necessário que $m + k \leq i \leq m$, o que implica $k = 0$ e $i = m$. Assim,

$$(4\text{—}5) \qquad (fg)_{m+n} = f_m g_n$$

e

$$(4\text{—}6) \qquad (fg)_{m+n+k} = 0, \qquad k > 0.$$

As afirmações (a), (b), (c) decorrem imediatamente de (4—5) e (4—6), enquanto (d) é uma conseqüência de (a) e (b). Deixamos a verificação de (e) a cargo do leitor.

Corolário 1. *O conjunto dos polinômios sôbre um dado corpo* F *é uma álgebra linear comutativa com elemento unidade sôbre* F, *em relação às operações dadas por* (4—1) *e* (4—2).

Demonstração. Como as operações (4—1) e (4—2) são aquelas definidas na álgebra F^∞ e como $F[x]$ é um subespaço de F^∞, basta demonstrar que o produto de dois polinômios é novamente um polinômio. Isto é trivial quando um dos fatôres é nulo e, em caso contrário, decorre de (a).

Corolário 2. *Suponhamos que* f, g *e* h *sejam polinômios sôbre o corpo* F *tais que* f $\neq 0$ *e* fg = fh. *Então* g = h.

Demonstração. Como $fg = fh$, $f(g - h) = 0$, e como $f \neq 0$ decorre imediatamente de (a) que $g - h = 0$.

Certos fatos adicionais decorrem bastante fàcilmente da demonstração do Teorema 1 e mencionaremos alguns dêles.

Suponhamos que

$$f = \sum_{i=0}^{m} f_i x^i \quad \text{e} \quad g = \sum_{j=0}^{n} g_j x^j.$$

Então, de (4—6) obtemos

$$fg = \sum_{s=0}^{m+n} \left(\sum_{r=0}^{s} f_r g_{s-r} \right) x^s. \tag{4—7}$$

O leitor deve verificar que, no caso particular de $f = cx^m$, $g = dx^n$ com c, d em F, (4—7) reduz-se a

$$(cx^m)(dx^n) = cdx^{m+n}. \tag{4—8}$$

Ora, de (4—8) e das leis distributivas em $F[x]$, segue que o produto em (4—7) também é dado por

$$\sum_{i,j} f_i g_j x^i x^j \tag{4—9}$$

onde a soma é estendida a todos os pares i, j de inteiros tais que $0 \leq i \leq m$, e $0 \leq j \leq n$.

Definição. *Seja* $\mathfrak{A}$ *uma álgebra linear com elemento unidade sôbre o corpo* F. *Indicaremos o elemento unidade de* $\mathfrak{A}$ *por* 1 *e faremos a conveção* $\alpha^0 = 1$ *para todo* α *em* $\mathfrak{A}$. *Então, a cada polinômio* $f \sum_{i=0}^{n} f_i x^i$ *sôbre* F *e* α *em* $\mathfrak{A}$, *associamos um elemento* $f(\alpha)$ *em* $\mathfrak{A}$ *pela regra*

$$f(\alpha) = \sum_{i=0}^{n} f_i \alpha^i.$$

Exemplo 3. Seja C o corpo dos números complexos e seja $f = x^2 + 2$.

(a) Se $\mathfrak{A} = C$ e z pertence a C, $f(z) = z^2 + 2$, em particular $f(2) = 6$ e

$$f\left(\frac{1+i}{1-i}\right) = 1.$$

(b) Se $\mathfrak{A}$ é a álgebra das 2×2 matrizes sôbre C e se

$$B = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$$

então $$f(B) = 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}^2 = \begin{bmatrix} 3 & 0 \\ -3 & 6 \end{bmatrix}.$$

(c) Se $\mathfrak{A}$ é a álgebra dos operadores lineares sôbre C^3 e T é o elemento de $\mathfrak{A}$ dado por

$$T(c_1, c_2, c_3) = (i\sqrt{2}\, c_1, c_2, i\sqrt{2}\, c_3)$$

então $f(T)$ é o operador linear sôbre C^3 definido por

$$f(T)(c_1, c_2, c_3) = (0, 3c, 0).$$

(d) Se $\mathfrak{A}$ é a álgebra dos polinômios sôbre C e $g = x^4 + 3i$, então $f(g)$ é o polinômio em $\mathfrak{A}$ dado por

$$f(g) = -7 + 6ix^4 + x^8.$$

O leitor observador poderá notar em relação a êste último exemplo que, se f é um polinômio sôbre um corpo arbitrário e x é o polinômio $\{0, 1, 0, \ldots\}$, então $f = f(x)$, mas aconselhamo-lo a esquecer êste fato.

Teorema 2. *Seja* F *um corpo e* $\mathfrak{A}$ *uma álgebra linear com elemento unidade sôbre* F. *Suponhamos que* f *e* g *sejam polinômios sôbre* F, *que* α *seja um elemento de* $\mathfrak{A}$ *e que* c *pertença a* F. *Então*

(a) $(cf + g)(\alpha) = cf(\alpha) + g(\alpha)$.

(b) $(fg)(\alpha) = f(\alpha)\, g(\alpha)$.

Demonstração. Como (a) é bem fácil de demonstrar, demonstraremos sòmente (b).

Suponhamos que

$$f = \sum_{i=0}^{m} f_i x^i \quad \text{e} \quad g = \sum_{j=0}^{n} g_j x^j.$$

Por (4—8) e (4—9)

$$fg = \sum_{i,j} f_i g_j x^{i+j}$$

e portanto por (a)

$$(fg)(\alpha) = \sum_{i,j} f_i g_j \alpha^{i+j}$$

$$= \left(\sum_{i=0}^{m} f_i\alpha^i\right)\left(\sum_{j=0}^{n} g_j\alpha^j\right)$$

$$= f(\alpha)\, g(\alpha).$$

Exercícios

1. Seja F um subcorpo do corpo dos números complexos e seja A a seguinte 2×2 matriz sôbre F:

$$A = \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix}.$$

Para cada um dos seguintes polinômios f sôbre F, calcular $f(A)$:

(a) $f = x^2 - x + 2$;
(b) $f = x_3 - 1$;
(c) $f = x^2 - 5x + 7$.

2. Seja T o operador linear sôbre R^3 definido por

$$T(x_1, x_2, x_3) = (x_1, x_3, -2x_2 - x_3).$$

Seja f o polinômio sôbre R definido por $f = -x^3 + 2$. Determinar $f(T)$.

3. Seja A uma $n \times n$ matriz diagonal sôbre o corpo F, isto é, uma matriz que satisfaz $A_{ij} = 0$ para $i \neq j$. Seja f o polinômio sôbre F definido por

$$f = (x - A_{11}) \ldots (x - A_{nn}).$$

Qual é a matriz $f(A)$?

4. Se f e g são polinômios independentes sôbre um corpo F e h é um polinômio não-nulo sôbre F, mostrar que fh e gh são independentes.

5. Se F é um corpo, mostrar que o produto de dois elementos não-nulos de F^∞ é não-nulo.

6. Seja S um conjunto de polinômios não-nulos sôbre um corpo F. Se dois quaisquer elementos de F nunca têm o mesmo grau, mostrar que S é um conjunto independente em $F[x]$.

7. Se a e b são elementos de um corpo F e $a \neq 0$, mostrar que os polinômios $1, ax + b, (ax + b)^2, (ax + b)^3, \ldots$ formam uma base de $F[x]$.

8. Se F é um corpo e h é um polinômio sôbre F de grau ≥ 1, mostrar que a aplicação $f \to f(h)$ é uma transformação linear injetora de $F[x]$ em $F[x]$. Mostrar que esta transformação é um isomorfismo de $F[x]$ em $F[x]$ se, e sòmente se, $\mathrm{gr}(h) = 1$.

9. Seja F um subcorpo do corpo dos números complexos e sejam T e D transformações sôbre $F[x]$ definidas por

$$T\left(\sum_{i=0}^{n} c_i x^i\right) = \sum_{i=0}^{n} \frac{c_i}{1 + i}\, x^{i+1},$$

e

$$D\left(\sum_{i=0}^{n} c_i x^i\right) = \sum_{i=1}^{n} i c_i x^{i-1}.$$

(a) Mostrar que T é um operador linear não-singular sôbre $F[x]$. Mostrar também que T não é inversível.

(b) Mostrar que D é um operador linear sôbre $F[x]$ e determinar seu núcleo.

(c) Mostrar que $DT = I$ e $TD \neq I$.

(d) Mostrar que $T[(Tf)g] = (Tf)(Tg) - T[f(Tg)]$ para todos f e g em $F[x]$.

(e) Enunciar e demonstrar uma regra para D semelhante à regra dada para T em (*d*).

(f) Suponhamos que V seja um subespaço não-nulo de $F[x]$ tal que Tf pertença a V para todo f em V. Mostrar que V não é de dimensão finita.

(g) Suponhamos que V seja um subespaço de $F[x]$ que tenha dimensão finita. Demonstrar que existe um inteiro $m \geq 0$ tal que $D^m f = 0$ para todo f em V.

4.3 Interpolação de Lagrange

Em tôda esta seção suporemos que F seja um corpo fixo e que $t_0, t_1, \ldots, t_n$ sejam $n + 1$ elementos *distintos* de F. Seja V o subespaço de $F[x]$ que consiste dos polinômios de grau menor ou igual a n (mais o polinômio nulo) e seja L_i a função de V em F definida para f em V por

$$L_i(f) = f(t_i), \quad 0 \leq i \leq n.$$

Pela parte (a) do Teorema 2, cada L_i é um funcional linear sôbre V e uma das coisas que pretendemos mostrar é que o conjunto formado por $L_0, L_1, \ldots, L_n$ é uma base do espaço V^* dual de V.

Evidentemente, para que isto ocorra, é necessário e suficiente (cf. Teorema 16 do Capítulo 3) que $\{L_0, L_1, \ldots, L_n\}$ seja a dual de uma base $\{P_0, P_1, \ldots, P_n\}$ de V. Existe no máximo uma tal base e, se existe, é caracterizada por

$$L_j(P_i) = P_i(t_j) = \delta_{ij}. \tag{4—10}$$

Os polinômios

$$P_i = \frac{(x - t_0)\ldots(x - t_{i-1})(x - t_{i+1})\ldots(x - t_n)}{(t_i - t_0)\ldots(t_i - t_{i-1})(t_i - t_{i+1})\ldots(t_i - t_n)} = \prod_{j \neq i}\left(\frac{x - t_j}{t_i - t_j}\right) \tag{4—11}$$

são de grau n, logo pertencem a V e, pelo Teorema 2, satisfazem (4—10).

Se $f = \sum_i c_i P_i$, então, para cada j,

$$f(t_j) = \sum_i c_i P_i(t_j) = c_j. \tag{4—12}$$

Como o polinômio nulo tem a propriedade de que $0(t) = 0$ para todo t em F, decorre de (4—12) que os polinômios $P_0, P_1, \ldots, P_n$ são linearmente independentes. Os polinômios $1, x, \ldots, x^n$ formam uma base de V, logo a dimensão de V é $n + 1$. Portanto o conjunto independente $\{P_0, P_1, \ldots, P_n\}$ deve também ser uma base de V. Assim, para todos f em V

$$f = \sum_{i=0}^{n} f(t_i)P_i. \tag{4—13}$$

A expressão (4—13) é denominada a **fórmula de interpolação de Lagrange.** Tomando $f = x^j$ em (4—13) obtemos

$$x_j = \sum_{i=0}^{n} (t_i)^j P_i.$$

Decorre então do Teorema 7 do Capítulo 2 que a matriz

$$\begin{bmatrix} 1 & t_0 & t_0^2 & \ldots & t_0^n \\ 1 & t_1 & t_1^2 & \ldots & t_1^n \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 1 & t_n & t_n^2 & \ldots & t_n^n \end{bmatrix} \tag{4—14}$$

é inversível. A matriz em (4—14) é dita uma **matriz de Vandermonde;** constitui um exercício interessante mostrar diretamente que uma tal matriz é inversível, quando $t_0, t_1 \ldots, t_n$ são $n + 1$ elementos distintos de F.

Se f é um polinômio arbitrário sôbre F, indicaremos por f' em nossa presente discussão a função polinominal de F em F que leva cada t em F em $f(t)$. Por definição (cf. Exemplo 4 do Capítulo 2) tôda função polinomial surge desta maneira; contudo, pode acontecer que $f' = g'$ para dois polinômios f e g tais que $f \neq g$. Felizmente, como veremos, esta situação desagradável ocorre apenas quando F é um corpo com um número finito de elementos distintos. Para descrever de maneira precisa a relação entre polinômios e funções polinomiais, precisamos definir o produto de duas funções polinomiais. Se f e g são polinômios sôbre F, o produto de f' por g' é a função $f'g'$ de F em F dada por

$$(f'g')(t) = f'(t)g'(t), \qquad t \text{ em } F. \tag{4—15}$$

Pela parte (b) do Teorema 2, $(fg)(t) = f(t)g(t)$, logo

$$(fg)'(t) = f'(t)g'(t)$$

para cada t em F. Assim, $f'g' = (fg)'$ e é uma função polinomial. Neste ponto, um fato de verificação imediata, que deixamos a cargo do leitor, é que o espaço vetorial das funções polinomiais sôbre F torna-se uma álgebra linear com elemento unidade sôbre F se a multiplicação é definida por (4—15).

Definição. *Seja* F *um corpo e sejam* $\mathfrak{A}$ *e* $\mathfrak{A}'$ *álgebras lineares sôbre* F. *As álgebras* $\mathfrak{A}$ *e* $\mathfrak{A}'$ *são ditas isomorfas se existe uma aplicação bijetora* $\alpha \to \alpha'$ *de* $\mathfrak{A}$ *em* $\mathfrak{A}'$ *tal que*

$$(1) \qquad (c\alpha + d\beta)' = c\alpha' + d\beta'$$

$$(2) \qquad (\alpha\beta)' = \alpha'\beta'$$

para todos α *e* β *em* $\mathfrak{A}$ *e todos escalares* c *e* d *em* F. *A aplicação* $\alpha \to \alpha'$ *é dita um isomorfismo de* $\mathfrak{A}$ *em* $\mathfrak{A}'$. *Um isomorfismo de* $\mathfrak{A}$ *em* $\mathfrak{A}'$ *é assim um isomorfismo de espaço vetorial de* $\mathfrak{A}$ *em* $\mathfrak{A}'$ *que tem a propriedade adicional* (2) *de "conservar" produtos.*

Exemplo 4. Seja V um espaço vetorial n-dimensional sôbre o corpo F. Pelo Teorema 13 do Capítulo 3 e observações subseqüentes, cada base ordenada $\mathfrak{B}$ de V determina um isomorfismo $T \to$ $\to [T]_{\mathfrak{B}}$ da álgebra dos operadores lineares sôbre V na álgebra das $n \times n$ matrizes sôbre F. Suponhamos agora que U seja um operador fixo sôbre V e que nos seja dado um polinômio

$$f = \sum_{i=0}^{n} c_i x^i$$

com coeficientes c_i em F. Então

$$f(U) = \sum_{i=0}^{n} c_i U^i$$

e como $T \to [T]_{\mathfrak{B}}$ é uma aplicação linear

$$[f(U)]_{\mathfrak{B}} = \sum_{i=0}^{n} c_i [U^i]_{\mathfrak{B}}.$$

Ora, do fato adicional

$$[T_1 T_2]_{\mathfrak{B}} = [T_1]_{\mathfrak{B}} [T_2]_{\mathfrak{B}}$$

para quaisquer T_1, T_2 em $L(V, V)$ decorre que

$$[U^i]_{\mathfrak{B}} = ([U]_{\mathfrak{B}})^i, \qquad 2 \leq i \leq n.$$

Como esta relação também é válida para $i = 0, 1$ obtemos o resultado

$$(4\text{—}16) \qquad [f(U)]_{\mathcal{B}} = f([U]_{\mathcal{B}}).$$

Em palavras, se U é um operador linear sôbre V, a matriz de um polinômio em U, em relação a uma dada base, é o mesmo polinômio na matriz de U.

Teorema 3. *Se* F *é um corpo contendo um número infinito de elementos distintos, a aplicação* $\mathrm{f} \to \mathrm{f}'$ *é um isomorfismo da álgebra dos polinômios sôbre* F *na álgebra das funções polinomiais sôbre* F.

Demonstração. Pela definição, a aplicação é sobrejetora e se f, g pertencem a $F[x]$, é evidente que

$$(cf + dg)' = df' + dg'$$

para todos os escalares c e d. Como já mostramos que $(fg)' = f'g'$ basta mostrar que a aplicação é injetora. Para tanto, é suficiente, pela linearidade, demonstrar que $f' = 0$ implica $f = 0$. Suponhamos então que f seja um polinômio de grau menor ou igual a n tal que $f' = 0$. Sejam $t_0, t_1, \ldots, t_n$ $n + 1$ elementos arbitrários distintos de F. Como $f' = 0$, $f(t_i) = 0$ para $i = 0, 1, \ldots, n$, e é uma conseqüência imediata de (4—13) que $f = 0$.

A partir dos resultados da próxima seção, obteremos uma demonstração totalmente diferente dêste teorema.

Exercícios

1. Usar a fórmula de interpolação de Lagrange para determinar um polinômio f com coeficientes reais tal que f tenha grau ≤ 3 e $f(-1) = -6$, $f(0) = 2$, $f(1) = -2$, $f(2) = 6$.

2. Sejam α, β, γ e δ números reais. Perguntamos quando é possível determinar um polinômio f sôbre R, de *grau não maior que* 2, tal que $f(-1) = \alpha$, $f(1) = \beta$, $f(3) = \gamma$, e $f(0) = \delta$. Demonstrar que isto é possível se, e sòmente se,

$$3\alpha + 6\beta - \gamma - 8\delta = 0.$$

3. Seja F o corpo dos números reais,

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

e $\qquad p = (x - 2)(x - 3)(x - 1).$

(a) Mostrar que $p(A) = 0$.

(b) Sejam P_1, P_2, P_3 os polinômios de Lagrange para $t_1 = 2$, $t_2 = 3$, $t_3 = 1$. Calcular $E_i = P_i(A)$, $i = 1, 2, 3$.

(c) Mostrar $E_1 + E_2 + E_3 = I$, $E_iE_j = 0$ se $i \neq j$, $E_i^2 = E_i$.

(d) Mostrar que $A = 2E_1 + 3E_2 + E_3$.

4. Seja $p = (x - 2)(x - 3)(x - 1)$ e seja T um operador arbitrário sôbre R^4 tal que $p(T) = 0$. Sejam P_1, P_2, P_3 os polinômios de Lagrange do Exercício 3 e seja $E_i = P_i(T)$, $i = 1, 2, 3$. Demonstrar que

$$E_1 + E_2 + E_3 = I, \quad E_jE_i = 0 \text{ se } i \neq j,$$
$$E_i^2 = E_i, \text{ e } T = 2E_1 + 3E_2 + E_3.$$

5. Seja n um inteiro positivo e F um corpo. Suponhamos que A seja uma $n \times n$ matriz sôbre F e P seja uma $n \times n$ matriz inversível sôbre F. Se f é um polinômio arbitrário sôbre F, demonstrar que

$$f(P^{-1}AP) = P^{-1}f(A)P.$$

6. Seja F um corpo. Consideremos certos funcionais lineares particulares sôbre F, obtidos pelo "cálculo do valor em t":

$$L(f) = f(t).$$

Tais funcionais não são apenas lineares, mas também têm a propriedade de que $L(fg) = L(f)L(g)$. Demonstrar que, se L é um funcional linear qualquer sôbre $F[x]$ tal que

$$L(fg) = L(f)L(g)$$

para todos f e g, então $L = 0$ ou existe um t em F tal que $L(f) = f(t)$ para todos f.

4.4 Ideais de Polinômios

Nesta seção preocupamo-nos com resultados que dependem fundamentalmente da estrutura multiplicadora da álgebra dos polinômios sôbre um corpo.

Lema. *Suponhamos que* f *e* d *sejam polinômios não-nulos sôbre um corpo* F *tal que gr*(d) $\leq$ *gr*(f). *Então existe um polinômio* g *em* F $[x]$ *tal que*

$$f - dg = 0 \quad \text{ou} \quad \text{gr}(f - dg) < \text{gr}(f)$$

Demonstração. Suponhamos que

$$f = a_mx^m + \sum_{i=0}^{m-1} a_ix^i, \qquad a_m \neq 0$$

e que

$$d = b_nx^n + \sum_{i=0}^{n-1} b_ix^i, \qquad b_n \neq 0.$$

Então $m \geq n$ e

$$f - \frac{a_m}{b_n} x^{m-n}d = 0 \text{ ou } \operatorname{gr}\left[f - \left(\frac{a_m}{b_n}\right) x^{m-n}d\right] < gr(f).$$

Assim, podemos tomar $g = \left(\frac{a_m}{b_n}\right) x^{m-n}$.

Usando êste lema podemos mostrar que o processo familiar de divisão de polinômios com coeficientes reais complexos é possível sôbre um corpo arbitrário.

Teorema 4. *Se* f, d *são polinômios sôbre um corpo* F *e* d *é diferente de* 0, *então existem polinômios* q, r *em* F [x] *tais que*

(i) $\mathrm{f} = \mathrm{dq} + \mathrm{r}$.

(ii) $\mathrm{r} = 0$ *ou* $gr(\mathrm{r}) < gr(\mathrm{d})$.

Os polinômios q,r *que satisfazem* (i) *e* (ii) *são únicos.*

Demonstração. Se f é 0 ou se $\operatorname{gr}(f) < \operatorname{gr}(d)$ podemos tomar $q = 0$ e $r = f$. Caso $f \neq 0$ e $\operatorname{gr}(f) \geq \operatorname{gr}(d)$, o lema anterior mostra que podemos escolher um polinômio g tal que $f - dg = 0$ ou $\operatorname{gr}(f - dg) < \operatorname{gr}(f)$. Se $f - dg \neq 0$ e $\operatorname{gr}(f - dg) \geq \operatorname{gr}(d)$ podemos escolher um polinômio h tal que $(f - dg) - dh = 0$ ou

$$\operatorname{gr}[f - d(g + h)] < \operatorname{gr}(f - dg).$$

Continuando êste processo enquanto fôr necessário, obteremos no final polinômios q, r tais que $r = 0$ ou $\operatorname{gr}(r) < \operatorname{gr}(d)$ e $f = dg + r$. Suponhamos agora que também tenhamos $f = dq' + r'$ onde $r' = 0$ ou $\operatorname{gr}(r') < \operatorname{gr}(d)$. Então $dq + r = dq' + r'$, e $d(q - q') = r' - r$. Se $q - q' \neq 0$ então $d(q - q') \neq 0$ e

$$\operatorname{gr}(d) + \operatorname{gr}(q - q') = \operatorname{gr}(r' - r).$$

Mas como o grau de $r' - r$ é menor que o grau de d, isto é impossível e $q' - q = 0$. Logo $r' - r = 0$.

Definição. *Seja* d *um polinômio não-nulo sôbre o corpo* F. *Se* f *está em* F[x], *o teorema anterior mostra que existe no máximo um polinômio* q *em* F[x] *tal que* f = dq. *Se existe um tal* q *dizemos que* d **divide** f, *que* f *é* **divisível** *por* d, *que* f *é um* **múltiplo** *de* d *e denominaremos* q *o* **quociente** *de* f *por* d. *Escreveremos tambem* q = f/d.

Corolário 1. *Seja* f *um polinômio sôbre o corpo* F *e seja* c *um elemento de* F. *Então* f *é divisível por* x — c *se, e sòmente se,* f(c) = 0.

Demonstração. Pelo teorema, $f = (x - c)q + r$ onde r é um polinômio constante. Pelo Teorema 2,

$$f(c) = 0q(c) + r(c) = r(c).$$

Logo $r = 0$ se, e sòmente se, $f(c) = 0$.

Definição. *Seja* F *um corpo. Diz-se que um elemento* c *em* F *é uma* **raiz** *ou um* **zero** *de um dado polinômio* f *sôbre* F *se* f(c) = 0.

Corolário 2. *Um polinômio* f *de grau* n *sôbre um corpo* F *tem no máximo* n *raízes em* F.

Demonstração. O resultado é òbviamente verdadeiro para polinômios de grau 0 e grau 1. Suponhamos que seja verdadeiro para polinômios de grau $n - 1$. Se a é uma raiz de f, $f = (x - a)q$ onde q tem grau $n - 1$. Como $f(b) = 0$ se, e sòmente se, $a = b$ ou $q(b) = = 0$, decorre de nossa hipótese de indução que f tem no máximo n raízes.

O leitor deve notar que o passo principal na demonstração do Teorema 3 é uma conseqüência imediata dêste corolário.

Definição. *Seja* F *um corpo. Um* **ideal** *em* F[x] *é um subespaço* M *de* F[x] *tal que* fg *pertence a* M *para todo* f *em* F[x] *e todo* g *em* M.

Exemplo 5. Se F é um corpo e d um polinômio sôbre F, o conjunto $M = dF[x]$, de todos os múltiplos df de d com f arbitrário em $F[x]$, é um ideal. De fato, M é não-vazio, pois M contém d. Se f,g pertencem a $F[x]$ e c é um escalar, então

$$c(df) - dg = d(cf - g)$$

pertence a M, portanto M é um subespaço. Finalmente, M também contém $(df)g = d(fg)$. O ideal M é denominado o **ideal principal gerado por** d.

Exemplo 6. Sejam $d_1, \ldots, d_n$ um número finito de polinômios sôbre F. Então a soma M dos subespaços $d_iF[x]$ é um subespaço e também é um ideal. De fato, suponhamos que p pertença a M. Então existem polinômios $f_1, \ldots, f_n$ em $F[x]$ tais que $p = d_1f_1 + \ldots + + d_nf_n$. Se g é um polinômio arbitrário sôbre F, então

$$pg = d_1(f_1g) + \ldots + d_n(f_ng)$$

de modo que pg também pertence a M. Assim, M é um ideal e dizemos que M é o ideal **gerado** pelos polinômios $d_1, \ldots, d_n$.

Exemplo 7. Seja F um subcorpo do corpo dos números complexos e consideremos o ideal

$$M = (x + 2)F[x] + (x^2 + 8x + 16)F[x].$$

Afirmamos que $M = F[x]$. De fato, M contém

$$x^2 + 8x + 16 - x(x + 2) = 6x + 16$$

logo M contém $6x + 16 - 6(x + 2) = 4$. Assim, o polinômio constante 1 pertence a M, bem como todos os seus múltiplos.

Teorema 5. *Se* F *é um corpo e* M *é um ideal não-nulo arbitrário em* F[x], *existe um único polinômio unitário* d *em* F[x] *tal que* M *seja o ideal principal gerado por* d.

Demonstração. Por hipótese, M contém um polinômio não-nulo; entre todos os polinômios não-nulos em M existe um polinômio d de grau mínimo. Podemos supor d unitário, pois caso contrário podemos multiplicar d por um escalar de modo a torná-lo unitário. Ora, se f pertence a M, o Teorema 4 mostra que $f = dq + r$ onde $r = 0$ ou $\text{gr}(r) < \text{gr}(d)$. Como d está em M, dq e $f - dq = r$ também pertencem a M. Como d é um elemento de M de grau mínimo não podemos ter $\text{gr}(r) < \text{gr}(d)$, portanto $r = 0$. Assim, $M = dF[x]$. Se g é um outro polinômio unitário tal que $M = gF[x]$, então existem polinômios não-nulos p,q tais que $d = gp$ e $g = dq$. Assim $d = dpq$ e

$$\text{gr}(d) = \text{gr}(d) + \text{gr}(p) + \text{gr}(q).$$

Logo $\text{gr}(p) = \text{gr}(q) = 0$ e como d,g são unitários, $p = q = 1$. Assim $d = g$.

É interessante observar que na demonstração acima, utilizamos um caso particular de um fato mais geral e bastante útil; a saber, se p é um polinômio não-nulo em um ideal M e se f é um polinômio em M que não é divisível por p, então $f = pq + r$ onde o "resto" r pertence a M, é diferente de 0 e tem grau menor que o de p. Já usamos êste fato no Exemplo 7 para mostrar que o polinômio constante 1 é o gerador unitário do ideal lá considerado. Em princípio, é sempre possível determinar o polinômio unitário que gera um dado ideal não-nulo, pois podemos, em última análise, obter um polinômio no ideal que tenha grau mínimo por meio de um número finito de divisões sucessíveis.

Corolário. *Se* $p_1, \ldots, p_n$ *são polinômios sôbre um corpo* F, *não todos nulos, existe um único polinômio unitário* d *em* F[x] *tal que*

(a) d *está no ideal gerado por* $p_1, \ldots, p_n$;

(b) d *divide cada um dos polinômios* p_i.

Todo polinômio que satisfaz (a) *e* (b) *satisfaz, necessàriamente,*

(c) d *é divisível por todo polinômio que divide cada um dos polinômios* $p_1, \ldots, p_n$.

Demonstração. Seja d o gerador unitário do ideal

$$p_1F[x] + \ldots + p_nF[x].$$

Todo membro dêste ideal é divisível por d; assim, cada um dos polinômios p_i é divisível por d. Suponhamos agora que f seja um polinômio que divida cada um dos polinômios $p_1, \ldots, p_n$. Então, existem polinômios $g_1, \ldots, g_n$ tais que $p_i = fg_i$, $1 \leq i \leq n$. Além disso, como d está no ideal

$$p_1F[x] + \ldots + p_nF[x],$$

existem polinômios $q_1, \ldots, q_n$ em $F[x]$ tais que

$$d = p_1q_1 + \ldots + p_nq_n.$$

Assim
$$d = f[g_1q_1 + \ldots + g_nq_n].$$

Mostramos que d é um polinômio unitário que satisfaz (a), (b) e (c). Se d' é um polinômio arbitrário que satisfaz (a) e (b) decorre de (a) e da definição de d que d' é um múltiplo escalar de d e satisfaz (c). Finalmente, se d' é um polinômio unitário temos $d' = d$.

Definição. Se $p_1, \ldots, p_n$ *são polinômios sôbre um corpo* F, *não todos nulos, o gerador unitário* d *do ideal*

$$p_1F[x] + \ldots + p_nF[x]$$

é denominado o **máximo divisor comum** (m. d. c.) *de* $p_1, \ldots, p_n$ *e é indicado por* $(p_1, \ldots, p_n)$. *Esta terminologia é justificada pelo corolário anterior. Dizemos que os polinômios* $p_1, \ldots, p_n$ *são* **relativamente primos** *se seu máximo divisor comum é* 1 *ou, equivalentemente, se o ideal que êles geram coincide com* F[x].

Exemplo 8. Seja C o corpo dos números complexos. Então

(a) $(x + 2, x^2 + 8x + 16) = 1$ (ver Exemplo 7);

(b) $((x-2)^2(x+i), (x-2)(x^2+1)) = (x-2)(x+i)$.

De fato, o ideal

$$(x-2)^2(x+i)F[x] + (x-2)(x^2+1)F[x]$$

contém

$$(x-2)^2(x+i)-(x-2)(x^2+1)=(x-2)(x+i)(i-2).$$

Logo contém $(x-2)(x+i)$, que é unitário e divide

$$(x-2)^2(x+i) \quad \text{e} \quad (x-2)(x^2+1).$$

Exemplo 9. Seja F o corpo dos números racionais e em $[Fx]$ seja M o ideal gerador por

$$(x-1)(x+2)^2, \ (x+2)^2(x-3) \quad \text{e} \quad (x-3).$$

Então M contém

$$\tfrac{1}{2}(x+2)^2[(x-1)-(x-3)]=(x+2)^2$$

e como

$$(x+2)^2=(x-3)(x+7)-17$$

M contém o polinômio constante 1. Assim, $M = F[x]$ e os polinômios

$$(x-1)(x+2)^2, \ (x+2)^2(x-3) \quad \text{e} \quad (x-3)$$

são relativamente primos.

Exercícios

1. Seja Q o corpo dos números racionais. Determinar quais dos seguintes subconjuntos de $Q[x]$ são ideais. Quando o conjunto fôr um ideal determinar seu gerador unitário.

(a) Todos f de grau par;
(b) Todos f de grau ≥ 5;
(c) Todos f tais que $f(0) = 0$;
(d) Todos f tais que $f(2) = f(4) = 0$;
(e) Todos f na imagem do operador linear T definido por

$$T\left(\sum_{i=0}^{n} c_i x^i\right) = \sum_{i=0}^{n} \frac{c_i}{i+1} x^{i+1}.$$

2. Determinar o m. d. c. de cada um dos seguintes pares de polinômios

(a) $2x^5 - x^3 - 3x^2 - 6x + 4$, $x^4 - x^3 - 2x - 2$;
(b) $3x^4 + 8x^2 - 3$, $x^3 + 2x^2 + 3x + 6$;
(c) $x^4 - 2x^3 - 2x^2 - 2x - 3$, $x^3 + 6x^2 + 7x + 1$.

3. Seja A uma $n \times n$ matriz sôbre um corpo F. Mostrar que o conjunto dos polinômios f em $F[x]$ tais que $f[A] = 0$ é um ideal.

4. Seja F um subcorpo dos números complexos e seja

$$A = \begin{bmatrix} 1 & -2 \\ 0 & 3 \end{bmatrix}.$$

Determinar o gerador unitário do ideal dos polinômios f em $F[x]$ tais que $f(A) = 0$.

5. Seja F um corpo. Mostrar que a interseção de um número arbitrário de ideais em $F[x]$ é um ideal.

6. Seja F um corpo. Mostrar que o ideal gerado por um número finito de polinômios $f_1, \ldots, f_n$ em $F[x]$ é a interseção de todos os ideais que contêm $f_1, \ldots, f_n$.

7. Seja K um subcorpo de um corpo F e suponhamos que f, g sejam polinômios em $K[x]$. Seja M_K o ideal gerado por f e g em $K[x]$ e M_F o ideal que êles geram em $F[x]$. Mostrar que M_K e M_F possuem o mesmo gerador unitário.

4.5 A Decomposição de um Polinômio em Fatôres Primos

Nesta seção demonstraremos que cada polinômio sôbre o corpo F pode ser escrito como um produto de polinômios "primos". Esta fatoração nos fornece um instrumento eficiente para determinar o máximo divisor comum de um número finito de polinômios e, em particular, fornece um meio efetivo para decidir se os polinômios são relativamente primos.

Definição. *Seja* F *um corpo. Diz-se que um polinômio* f *em* F[x] *é* **redutível sôbre** F *se existem polinômios* g,h *em* F[x] *de grau* $\geq$ 1 *tais que* f = gh *e, em caso contrário, diz-se que* f *é* **irredutível sôbre** F. *Um polinômio não constante irredutível sôbre* F *é denominado um* **polinômio primo sôbre** F *e dizemos, às vêzes, que é um* **primo em** F[x].

Exemplo 10. O polinômio $x^2 + 1$ é redutível sôbre o corpo C dos números complexos, pois

$$x^2 + 1 = (x + i)(x - i)$$

e os polinômios $x + i$, $x - i$ pertencem a $C[x]$. Por outro lado, $x^2 + 1$ é irredutível sôbre o corpo R dos números reais, pois se

$$x^2 + 1 = (ax + b)(a'x + b')$$

com a, a', b, b' em R, então

$$aa' = 1, \quad ab' + ba' = 0, \quad bb' = 1.$$

Estas relações implicam $a^2 + b^2 = 0$, o que é impossível com números reais a e b, a menos que $a = b = 0$.

Teorema 6. *Sejam* p, f *e* g *polinômios sôbre o corpo* F. *Suponhamos que* p *seja um polinômio primo e que* p *divida o produto* fg. *Então* p *divide* f *ou* p *divide* g.

Demonstração. Não há perda de generalidade se supomos que p é um polinômio primo unitário. O fato de que p é primo diz então simplesmente que os únicos divisores unitários de p são 1 e p. Seja $d = (f,p)$ o m. d. c. de f e p. Ou d $= 1$ ou $d = p$, pois d é um polinômio unitário que divide p. Se $d = p$ então p divide f e já terminamos. Portanto, suponhamos que $d = 1$, isto é, suponhamos que f e p sejam relativamente primos. Demonstraremos que p divide g. Como $(fp) = 1$, existem polinômios f_0 e p_0 tais que $1 = f_0 f + p_0 p$. Multiplicando por g obtemos

$$\begin{aligned} g &= f_0 f g + p_0 p g \\ &= (fg) f_0 + p(p_0 g). \end{aligned}$$

Como p divide fg, divide também $(fg)f_0$ e certamente p divide $p(p_0 g)$. Assim p divide g.

Corolário. *Se* p *é um primo e divide um produto* $f_1, \ldots, f_n$, *então* p *divide um dos polinômios* $f_1, \ldots, f_n$.

Demonstração. A demonstração é por indução. Para $n = 2$, o resultado é simplesmente o enunciado do Teorema 6. Suponhamos que o resultado seja válido para $n = k$ e que p divide o produto $f_1, \ldots, f_{k+1}$ de certos $(k + 1)$ polinômios. Como p divide $(f_1, \ldots \ldots f_k) f_{k+1}$, p divide f_{k+1} ou p divide $f_1 \ldots f_k$. Pela hipótese de indução, se p divide $f_1 \ldots f_k$, então p divide f_j para algum j, $1 \leq j \leq k$. Portanto vemos que em qualquer caso p deve dividir algum f_j, $1 \leq \leq j \leq k + 1$.

Teorema 7. *Se* F *é um corpo, um polinômio não-constante e unitário em* F[x] *pode ser decomposto como um produto de primos unitários em* F[x] *de uma e, a menos da ordem, sòmente uma maneira.*

Demonstração. Suponhamos que f seja um polinômio não-constante e unitário sôbre F. Como polinômios de grau 1 são irredutíveis, nada há a demonstrar se $\text{gr}(f) = 1$. Suponhamos que f tenha grau $n > 1$. Por indução podemos supor o teorema verdadeiro para todos os polinômios não-constantes e unitários de grau menor que n. Se f é irredutível, f já está decomposto como um produto de primos unitários e, em caso contrário, $f = gh$ onde g e h são polinômios não-constantes e unitários de grau menor que n. Assim g e h podem ser decompostos como produtos de primos unitários em $F[x]$, logo f também pode sê-lo. Suponhamos agora que

$$f = p_1 \ldots p_m = q_1, \ldots q_n$$

onde $p_1, \ldots, p_m$ e $q_1, \ldots, q_n$ são primos unitários em $F[x]$. Então

p_m divide o produto $q_1 \ldots q_n$. Pelo corolário acima, p_m divide algum q_i. Como q_i e p_m são ambos primos e unitários isto significa que

$$q_i = p_m. \tag{4—16}$$

De (4—16) vemos que $m = n = 1$ se $m = 1$ ou $n = 1$, pois

$$\operatorname{gr}(f) = \sum_{i=1}^{m} \operatorname{gr}(p_i) = \sum_{j=1}^{n} \operatorname{gr}(q_j).$$

Neste caso nada resta a demonstrar, portanto podemos supor $m > 1$ e $n > 1$. Reordenando os q podemos supor que $p_m = q_n$ e que

$$p_1 \ldots p_{m-1}p_m = q_1 \ldots q_{n-1}p_m.$$

Ora, pelo Corolário 2 do Teorema 1 decorre que

$$p_1 \ldots p_{m-1} = q_1 \ldots q_{n-1}.$$

Como o polinômio $p_1 \ldots p_{m-1}$ tem grau menor que n, nossa hipótese de indução se aplica e mostra que a seqüência $q_1, \ldots, q_{n-1}$ é no máximo uma reordenação da seqüência $p_1, \ldots, p_{m-1}$. Isto, junto com (4—16), mostra que a decomposição de f num produto de primos unitários é única a menos da ordem dos fatôres.

Na decomposição acima de um polinômio f não-constante e unitário, alguns dos fatôres primos e unitários podem repetir-se. Se $p_1, p_2, \ldots, p_r$ são os primos unitários distintos que ocorrem nesta decomposição de f, então

$$f = p_1^{n_1}p_2^{n_2} \ldots p_r^{n_r}, \tag{4—17}$$

sendo o expoente n_i o número de vêzes que o primo p_i ocorre na decomposição. Esta decomposição é certamente única e é denominada a **decomposição primária** de f. Verifica-se fàcilmente que todo divisor unitário de f é da forma

$$p_1^{m_1}p_2^{m_2} \ldots p_r^{m_r}, \qquad 0 \leq m_i \leq n_i. \tag{4—18}$$

De (4—18) decorre que o m. d. c. de um número finito de polinômios não-constantes e unitários $f_1, \ldots, f_s$ é obtido combinando todos aquêles primos unitários que ocorrem simultâneamente nas decomposições de $f_1, \ldots, f_s$. O expoente ao qual cada primo deve ser elevado é o maior número natural para o qual a correspondente potência prima é um fator de cada f_i. Se não existem potências primas (não-triviais) que sejam fatôres de cada f_i, os polinômios são relativamente primos.

Exemplo 11. Suponhamos que F seja um corpo e sejam a, b e c elementos distintos de F. Então os polinômios $x - a$, $x - b$ e $x - c$ são primos unitários distintos em $F[x]$. Se m, n e s são inteiros positivos, $(x - c)^s$ é o m. d. c. dos polinômios

$$(x - b)^n(x - c)^s \quad \text{e} \quad (x - a)^m(x - c)^s$$

enquanto que os três polinômios

$$(x - b)^n(x - c)^s, \quad (x - a)^m(x - c)^s, \quad (x - a)^m(x - b)^n$$

são relativamente primos.

Teorema 8. *Seja* f *um polinômio não-constante e unitário sôbre o corpo* F *e seja*

$$f = p_1^{n_1} \dots p_k^{n_k}$$

a decomposição de f *fatôres primos. Para cada* j, $1 \leq j \leq k$, *seja*

$$f_j = f/p_j^{n_j} = \prod_{i \neq j} p_i^{n_i}$$

Então $f_1, \dots, f_k$ *são relativamente primos.*

Demonstração. Deixamos a demonstração (fácil) dêste fato a cargo do leitor. Enunciamos êste teorema principalmente porque pretendemos usá-lo posteriormente.

Definição. *O corpo* F *é dito* **algèbricamente fechado** *se todo polinômio primo sôbre* F *tem grau* 1.

Dizer que F é algèbricamente fechado significa que todo polinômio não-constante irredutível e unitário sôbre F é da forma $(x - c)$. Já observamos que cada um dêstes polinômios é irredutível para qualquer F. Conseqüentemente, pode-se dar uma definição equivalente de um corpo algèbricamente fechado como um corpo F tal que todo polinômio não-constante f em $F[x]$ possa ser expresso sob a forma

$$f = c(x - c_1)^{n_1} \dots (x - c_k)^{n_k}$$

onde c é um escalar, $c_1, \dots, c_k$ são elementos distintos de F e $n_1, \dots, n_k$ são inteiros positivos. Ainda uma outra formulação é que, se f é um polinômio não-constante sôbre F, então existe um elemento c em F tal que $f(c) = 0$.

O corpo R dos números reais não é algèbricamente fechado porque o polinômio $(x^2 + 1)$ é irredutível sôbre R mas não tem grau 1, ou porque não existe um número real c tal que $c^2 + 1 = 0$. O chamado Teorema Fundamental da Álgebra afirma que o corpo C dos números complexos é algèbricamente fechado. Não demonstraremos êste teorema, apesar de usarmo-lo um pouco mais adiante neste livro. A demonstração é omitida em parte pelas limitações de tempo e em parte porque a demonstração depende de uma propriedade "não algébrica" do sistema dos números reais. Para uma das demonstrações possíveis o leitor interessado poderá consultar o livro de Schreier e Sperner citado na bibliografia.

Exercícios

1. Seja p um polinômio unitário sôbre o corpo F e sejam f e g polinômios sôbre F, relativamente primos. Demonstrar que o m. d. c. de pf e pg é p.

2. Admitindo o Teorema Fundamental da Álgebra, demonstrar o que segue. Se f e g são polinômios sôbre o corpo dos números complexos, então $(f,g) = 1$ se, e sòmente se, f e g não possuem nenhuma raiz comum.

3. Seja D o operador derivação sôbre o espaço dos polinômios sôbre o corpo dos números complexos. Seja f um polinômio unitário sôbre o corpo dos números complexos. Demonstrar que

$$f = (x - c_1) \dots (x - c_k)$$

onde $c_1, \dots, c_k$ são números complexos *distintos* se, e sòmente se, f e Df são relativamente primos. Em outras palavras, f não possui raiz múltipla se, e sòmente se, f e Df não possuem nenhuma raiz comum. (Admitir o Teorema Fundamental da Álgebra.)

4. Demonstrar a seguinte generalização do Exercício 3. Seja F um subcorpo do corpo dos números complexos e seja f um polinômio sôbre F. Então f é um produto de polinômios primos *distintos* sôbre F se, e sòmente se, $(f,Df) = 1$.

5. Seja F um subcorpo do corpo dos números complexos e seja f um polinômio sôbre F de grau no máximo n. Indiquemos por $f^{(k)}$ a k-ésima derivada de f, isto é, $f^{(k)} = D^k f$. Demonstrar a fórmula de Taylor:

$$\begin{aligned} f &= \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(c)(x-c)^k \\ &= f(c) + f^{(1)}(c)(x-c) + \frac{1}{2!} f^{(2)}(c)(x-c)^2 + \dots + \frac{1}{n!} f^{(n)}(c)(x-c)^n \end{aligned}$$

onde c é um elemento arbitrário de F. (*Sugestão*: Demonstrar primeiro para $f = x^r$; depois usar combinações lineares.)

6. Demonstrar a seguinte generalização da fórmula de Taylor. Sejam f, g e h polinômios sôbre um subcorpo dos números complexos, com $gr(f) \leq n$. Então

$$f(g) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(h)(g-h)^k.$$

(Aqui $f(g)$ indica "f de g".)

7. Demonstrar que todo polinômio irredutível sôbre o corpo dos números reais tem grau par ou grau 1. (É necessário usar um teorema de cácIculo.)

Para os exercícios restantes, precisaremos da definição seguinte: Se f, g e p são polinômios sôbre o corpo C com $p \neq 0$, dizemos que f é **congruente a** g **módulo** p se $(f - g)$ é divisível por p. Se f é congruente a g módulo p, escrevemos

$$f \equiv g \bmod p.$$

8. Demonstrar, para todo polinômio não-nulo p, que a congruência módulo p é uma relação de equivalência.

(a) Reflexiva: $f \equiv f \bmod p$.
(b) Simétrica: se $f \equiv g \bmod p$ e $g \equiv f \bmod p$.
(c) Transitiva: se $f \equiv g \bmod p$ e $g \equiv h \bmod p$, então $f \equiv h \bmod p$.

9. Suponhamos que $f \equiv g \bmod p$ e $f_1 \equiv g_1 \bmod p$.

(a) Demonstrar que $f + f_1 \equiv g + g_1 \bmod p$;
(b) Demonstrar que $ff_1 \equiv gg_1 \bmod p$.

10. Usar o Exercício 9 para demonstrar o que segue. Se f, g, h e p são polinômios sôbre o corpo F e $p \neq 0$ e se $f \equiv g \bmod p$, então $h(f) \equiv h(g) \bmod p$.

11. Se p é um polinômio irredutível e $fg \equiv 0 \bmod p$, demonstrar que ou $f \equiv 0 \bmod p$ ou $g \equiv 0 \bmod p$. Dar um exemplo que mostre que isto é falso se p não é irredutível.

CAPÍTULO 5

DETERMINANTES

5.1 Anéis Comutativos

Neste capítulo demonstraremos os fatos essenciais sôbre determinantes de matrizes quadradas. Faremos isto não apenas para matrizes sôbre um corpo, mas também para matrizes cujos elementos sejam "escalares" de um tipo mais geral. Existem duas razões para esta generalidade. Primeiro, em certos pontos do próximo capítulo, será necessário tratarmos de determinantes de matrizes cujos elementos são polinômios. Segundo, no tratamento de determinantes que apresentamos, um dos axiomas da definição de um corpo não desempenha nenhum papel, a saber, o axioma que garante a existência de um inverso multiplicativo para todo elemento não-nulo. Por estas razões, é conveniente desenvolver a teoria dos determinantes para matrizes cujos elementos sejam pertencentes a um anel comutativo com elemento unidade.

Definição. *Um* **anel** *é um conjunto* K, *munido de duas operações* $(x, y) \rightarrow x + y$ *e* $(x, y) \rightarrow xy$ *que satisfazem*

(1) K *é um grupo comutativo em relação à operação* $(x, y) \rightarrow$ $\rightarrow x + y$ (K *é um grupo comutativo em relação a adição*);

(2) $(xy)z = x(yz)$ (*a multiplicação é associativa*);

(3) $x(y + z) = xy + xz$; $(y + z)\,x = yz + zx$ (*valem as duas leis distributivas*).

Se $xy = yx$ *para todos* x *e* y *em* K, *dizemos que o anel* K *é* **comutativo.** *Se existe um elemento* 1 *em* K *tal que* $1x = x1 = x$ *para todo* x, *dizemos que* K *é um* **anel com elemento unidade** *e* 1 *é denominado o* **elemento unidade** *de* K.

Estamos interessados em anéis comutativos com elemento unidade. Tal anel pode ser descrito ràpidamente como um conjunto *K*, munido de duas operações que satisfazem todos os axiomas da definição de um corpo dado no Capítulo 1, com exceção talvez do axioma (8) e da condição $1 \neq 0$. Assim, um corpo é um anel comutativo com elemento unidade diferente de zero tal que a cada x não-nulo corresponde um elemento x^{-1} tal que $xx^{-1} = 1$. O conjunto dos inteiros, com as operações usuais, é um anel comutativo com elemento unidade mas não é um corpo. Outro anel comutativo com elemento unidade é o conjunto dos polinômios sôbre um corpo, com a adição e a multiplicação que definimos para polinômios.

Se K é um anel comutativo com elemento unidade, definimos uma $m \times n$ matriz sôbre K como sendo uma função A do conjunto dos pares (i, j) de inteiros, $1 \leq i \leq m$, $1 \leq j \leq n$, em K. Como sempre, representamos tal matriz por uma tabela retangular com m linhas e n colunas. A soma e o produto de matrizes sôbre K são definidas como para matrizes sôbre um corpo

$$(A + B)_{ij} = A_{ij} + B_{ij}$$
$$(AB)_{ij} = \sum_k A_{ik}B_{kj}$$

sendo que a soma é definida quando A e B têm o mesmo número de linhas e o mesmo número de colunas e o produto é definido quando o número de colunas de A é igual ao número de linhas de B. As propriedades algébricas básicas destas operações são ainda válidas. Por exemplo

$$A(B + C) = AB + AC, \qquad (AB)C = A(BC), \qquad \text{etc.}$$

Como no caso de corpos, referir-nos-emos aos elementos de K como escalares. Podemos então definir combinações lineares das linhas ou colunas de uma matriz como fizemos anteriormente. A grosso modo, tudo o que fizemos para matrizes sôbre um corpo é válido para matrizes sôbre K, excetuando-se os resultados que dependiam da possibilidade de "dividir" em K.

5.2 Funções Determinantes

Seja K um anel comutativo com elemento unidade. Desejamos associar a cada $n \times n$ matriz (quadrada) sôbre K um escalar (elemento de K) que será conhecido como o determinante da matriz. É possível definir o determinante de uma matriz quadrada A simplesmente escrevendo uma fórmula para êste determinante em função dos elementos de A. Pode-se então deduzir as diversas propriedades dos determinantes a partir desta fórmula. Contudo, tal fórmula é bastante complicada e a fim de ganhar alguma vantagem técnica, procederemos como segue. Definiremos uma "função determinante" sôbre $n \times n$ matrizes como uma função que associa a cada $n \times n$ matriz sôbre K um escalar, sendo que essa função possui propriedades especiais. É linear como uma função de cada uma das linhas da matriz; seu valor é 0 sôbre qualquer matriz que tenha duas linhas iguais e seu valor sôbre a $n \times n$ matriz unidade é 1. Demonstraremos que existe uma tal função e depois que é única, isto é, que existe exatamente uma tal função. Ao demonstrarmos a unicidade será obtida uma fórmula explícita para o determinante, junto com muitas de suas úteis propriedades.

Esta seção será dedicada à definição da "função determinante" e à demonstração de que existe pelo menos uma tal função.

Definição. *Seja* K *um anel comutativo com elemento unidade,* n *um inteiro positivo e seja* D *uma função que associa a cada* n $\times$ n *matriz* A *sôbre* K *um escalar* D(A) *em* K. *Dizemos que* D *é* **n-linear** *se para cada* i, $1 \leq i \leq n$, D *é uma função linear da* i*-ésima linha quando as outras* (n — 1) *linhas são mantidas fixas.*

Esta definição exige algum esclarecimento. Se D é uma função das $n \times n$ matrizes sôbre K em K e se $\alpha_1, \ldots, \alpha_n$, são as linhas da matriz A, escrevamos também

$$D(A) = D(\alpha_1, \ldots, \alpha_n)$$

isto é, consideremos D também como uma função das linhas de A. A afirmação de que D é n-linear significa então

$$\text{(5-1)} \quad D(\alpha_1, \ldots, c\alpha_i + \alpha_i', \ldots, \alpha_n) = cD(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_n) \\ + D(\alpha_1, \ldots, \alpha_i', \ldots, \alpha_n).$$

Se fixamos tôdas as linhas exceto a linha i e consideramos K como uma funcão da i-ésima linha, em geral é conveniente indicar $D(A)$ por $D(\alpha_i)$. Assim, podemos abreviar (5—1) para

$$D(c\alpha_i + \alpha_i') = cD(\alpha_i) + D(\alpha_i')$$

desde que esteja claro o que isto significa.

Exemplo 1. Sejam $k_1, \ldots, k_n$ inteiros positivos, $1 \leq k_i \leq n$ e seja a um elemento de K. Para cada $n \times n$ matriz A sôbre K definamos

$$\text{(5—2)} \qquad D(A) = aA(1, k_1) \ldots A(n, k_n).$$

Então, a função D definida por (5—2) é n-linear. De fato, considerando B como uma função da i-ésima linha de A, com as outras linhas fixas, podemos escrever

$$D(\alpha_i) = A(i, k_i)b$$

onde b é algum elemento fixo de K. Seja $\alpha_i' = (A_{i1}', \ldots, A_{in}')$. Então temos

$$\begin{aligned} D(c\alpha_i + \alpha_i') &= [cA(i, k_i) + A'(i, k_i)]b \\ &= cD(\alpha_i) + D(\alpha_i'). \end{aligned}$$

Assim, D é uma função linear de cada uma das linhas de A.

Uma particular função n-linear dêste tipo é

$$D(A) = A_{11}A_{22} \ldots A_{nn}.$$

Em outras palavras, o "produto dos elementos diagonais" é uma função n-linear sôbre $n \times n$ matrizes sôbre K.

Exemplo 2. Determinemos tôdas as funções bilineares (2-lineares) sôbre 2×2 matrizes sôbre K. Seja D uma dessas funções. Indicando as linhas da 2×2 matriz unidade por ϵ_1, ϵ_2, temos

$$D(A) = D(A_{11}\epsilon_1 + A_{12}\epsilon_2, A_{21}\epsilon_1 + A_{22}\epsilon_2).$$

Usando o fato de que D é bilinear, (5—1) implica

$$\begin{aligned} D(A) &= A_{11}D(\epsilon_1, A_{21}\epsilon_1 + A_{22}\epsilon_2) + A_{12}D(\epsilon_2, A_{21}\epsilon_1 + A_{22}\epsilon_2) \\ &= A_{11}A_{21}D(\epsilon_1, \epsilon_1) + A_{11}A_{22}D(\epsilon_1, \epsilon_2) \\ &\qquad + A_{12}A_{21}D(\epsilon_2, \epsilon_1) + A_{12}A_{22}D(\epsilon_2, \epsilon_2). \end{aligned}$$

Assim, D é completamente determinada pelos quatro escalares

$$D(\epsilon_1, \epsilon_1), \quad D(\epsilon_1, \epsilon_2), \quad D(\epsilon_2, \epsilon_1), \quad \text{e} \quad D(\epsilon_2, \epsilon_2).$$

O leitor deverá achar fácil verificar o seguinte: se a, b, c, d são quatro escalares arbitrários em K e se definimos

$$D(A) = A_{11}A_{21}a + A_{11}A_{22}b + A_{12}A_{21}c + A_{12}A_{22}d,$$

então D é uma função bilinear sôbre as 2×2 matrizes sôbre K e

$$\begin{aligned} D(\epsilon_1, \epsilon_1) &= a, & D(\epsilon_1, \epsilon_2) &= b, \\ D(\epsilon_2, \epsilon_1) &= c, & D(\epsilon_2, \epsilon_2) &= d. \end{aligned}$$

Lema. *Uma combinação linear de funções* n-*lineares é* n-*linear.*

Demonstração. Basta demonstrar que uma combinação linear de duas funções n-lineares é n-linear. Sejam D e E funções n-lineares. Se a e b pertencem a K, a combinação linear $aD + bE$ é evidentemente definida por

$$(aD + bE)(A) = aD(A) + bE(A).$$

Logo, fixando tôdas as linhas exceto a linha i

$$\begin{aligned}(aD + bE)(c\alpha_i + \alpha_i') &= aD(c\alpha_i + \alpha_i') + bE(c\alpha_i + \alpha_i')\\ &= acD(\alpha_i) + aD(\alpha_i') + bcE(\alpha)_i + bE(\alpha_i')\\ &= c(aD + bE)(\alpha_i) + (aD + bE)(\alpha_i').\end{aligned}$$

Se K é um corpo e V é o conjunto das $n \times n$ matrizes sôbre K, o lema acima diz o seguinte: o conjunto das funções n-lineares sôbre V é um subespaço do espaço de tôdas as funções de V em K.

Exemplo 3. Seja D a função definida sôbre as 2×2 matrizes sôbre K por

(5—3) $$D(A) = A_{11}A_{22} - A_{12}A_{21}.$$

Ora, D é a soma de duas funções do tipo descrito no Exemplo 1:

$$\begin{aligned}D &= D_1 + D_2\\ D_1(A) &= A_{11}A_{22}\\ D_2(A) &= -A_{12}A_{21}.\end{aligned}$$

Pelo lema acima, D é uma função bilinear. O leitor que tenha tido alguma experiência com determinantes não achará êste fato surpreendente, pois reconhecerá (5—3) como a definição usual do determinante de uma 2×2 matriz. É claro que a função D que acabamos de definir não é uma função bilinear típica. Ela possui muitas propriedades particulares. Vejamos algumas dessas propriedades Primeiro, se I é a 2×2 matriz unidade, então $D(I) = 1$, isto é, $D(\epsilon_2, \epsilon_1) = 1$. Segundo, se as duas linhas de A são iguais, então

$$D(A) = A_{11}A_{12} - A_{12}A_{11} = 0.$$

Terceiro, se A' é a matriz obtida de uma 2×2 matriz A permutando suas linhas, então $D(A') = -D(A)$; de fato,

$$\begin{aligned}D(A') &= A'_{11}A'_{22} - A'_{12}A'_{21}\\ &= A_{21}A_{12} - A_{22}A_{11}\\ &= -D(A).\end{aligned}$$

Definição. *Seja* D *uma função* n-*linear. Dizemos que* D *é* **alternada** (*ou* **alternante**) *se as duas condições seguintes estão satisfeitas:*

(1) $D(A) = 0$ *sempre que duas linhas de* A *são iguais.*

(2) *Se* A' *é uma matriz obtida de* A *permutando duas linhas de* A, *então* $D(A') = -D(A)$.

Demonstraremos abaixo que tôda função n-linear D que satisfaz (1) automàticamente satisfaz (2). Colocamos as duas propriedades na definição de função n-linear alternada por conveniência. O leitor provàvelmente notará que se D satisfaz (2) e A é uma matriz com duas linhas iguais então $D(A) = -D(A)$. Somos tentados a concluir que D também satisfaz a condição (1). Isto é verdade, por exemplo, se K é um corpo no qual $1 + 1 \neq 0$, mas em geral (1) não é uma conseqüência de (2).

Definição. *Seja* K *um anel comutativo com elemento unidade e seja* n *um inteiro positivo. Suponhamos que* D *seja uma função das* n $\times$ n *matrizes sôbre* K *em* K. *Dizemos que* D *é uma* **função determinante** *se* D *é* n-*linear, alternada e* $D(I) = 1$.

Como afirmamos anteriormente, mostraremos no final que existe exatamente uma função determinante sôbre $n \times n$ matrizes sôbre K. Isto pode ser visto fàcilmente para 1×1 matrizes $A = [a]$ sôbre K. A função D dada por $D(A) = a$ é uma função determinante e é evidentemente a única função determinante sôbre 1×1 matrizes. Estamos em condições de eliminar também o caso $n = 2$. Demonstramos, no Exemplo 3, que a função

$$D(A) = A_{11}A_{22} - A_{12}A_{21}$$

é uma função determinante. Além disso, a fórmula apresentada no Exemplo 2 mostra que D é a única função determinante sôbre 2×2 matrizes, pois demonstramos que para qualquer função bilinear D

$$\begin{aligned} D(A) = A_{11}A_{21}D(\epsilon_1, \epsilon_1) + A_{11}A_{22}D(\epsilon_1, \epsilon_2) \\ + A_{12}A_{21}D(\epsilon_2, \epsilon_1) + A_{12}A_{22}D(\epsilon_2, \epsilon_2). \end{aligned}$$

Se D é alternada, então

$$D(\epsilon_1, \epsilon_1) = D(\epsilon_2, \epsilon_2) = 0$$

e

$$D(\epsilon_2, \epsilon_1) = -D(\epsilon_1, \epsilon_2) = -D(I).$$

Se D satisfaz também $D(I) = 1$, então

$$D(A) = A_{11}A_{22} - A_{12}A_{21}.$$

Exemplo 4. Seja F um corpo e seja D uma função trilinear alternada, arbitrária, sôbre as 3×3 matrizes sôbre o anel de polinômios $F[x]$.

Seja

$$A = \begin{bmatrix} x & 0 & -x^2 \\ 0 & 1 & 0 \\ 1 & 0 & x^3 \end{bmatrix}.$$

Se indicarmos as linhas da 3×3 matriz unidade por $\epsilon_1, \epsilon_2, \epsilon_3$, então

$$D(A) = D(x\epsilon_1 - x^2\epsilon_3, \epsilon_2, \epsilon_1 + x^3\epsilon_3).$$

Como D é uma função linear de cada linha

$$\begin{aligned} D(A) &= xD(\epsilon_1, \epsilon_2, \epsilon_1 + x^3\epsilon_3) - x^2D(\epsilon_3, \epsilon_2, \epsilon_1 + x^3\epsilon_3) \\ &= xD(\epsilon_1, \epsilon_2, \epsilon_1) + x^4D(\epsilon_1, \epsilon_2, \epsilon_3) - x^2D(\epsilon_3, \epsilon_2, \epsilon_1) - x^5D(\epsilon_3, \epsilon_2, \epsilon_3). \end{aligned}$$

Como D é alternada decorre que

$$D(A) = (x^4 + x^2)D(\epsilon_1, \epsilon_2, \epsilon_3).$$

Lema. *Seja* D *uma função bilinear com a propriedade de que* $D(A) = 0$ *para tôdas* 2×2 *matrizes* A *sôbre* K *que tenham linhas iguais. Então* D *é alternada.*

Demonstração. O que precisamos mostrar é que se A é uma 2×2 matriz e A' é obtida transpondo-se as linhas de A, então $D(A') = -D(A)$. Se as linhas de A são α e β, isto significa que precisamos mostrar que $D(\beta, \alpha) = -D(\alpha, \beta)$. Como D é bilinear

$$D(\alpha + \beta, \alpha + \beta) = D(\alpha, \alpha) + D(\alpha, \beta) + D(\beta, \alpha) + D(\beta, \beta).$$

Por nossa hipótese, $D(\alpha + \beta, \alpha + \beta) = D(\alpha, \alpha) = (\beta, \beta) = 0$. Portanto,

$$0 = D(\alpha, \beta) + D(\beta, \alpha).$$

Lema. *Seja* D *uma função* n*-linear sôbre* $n \times n$ *matrizes sôbre* K. *Suponhamos que* D *tenha a propriedade de que* D(A) = 0 *sempre que duas linhas adjacentes de* A *sejam iguais. Então* D *é alternada.*

Demonstração. Precidamos mostrar que $D(A) = 0$ quando duas quaisquer linhas de A são iguais e que $D(A') = -D(A)$ e A' é obtida transpondo-se (trocando entre si) duas quaisquer linhas de A. Em primeiro lugar suponhamos que A' seja obtida pela transposição de duas linhas adjacentes de A. O leitor deve verificar que o argumento usado na demonstração do lema anterior se aplica ao presente caso e nos fornece $D(A') = -D(A)$.

Seja agora B obtida transpondo-se as linhas i e j de A, onde $i < j$. Podemos obter B a partir de A por uma sucessão de transposições de pares de linhas adajacentes. Começamos transpondo a linha i com a linha $(i + 1)$ e continuamos até que as linhas estejam na ordem

$$\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_j, \alpha_i, \alpha_{j+1}, \ldots, \alpha_n.$$

Isto requer $k = j - i$ transposições de linhas adjacentes. Agora, movemos α_j para a i-ésima posição usando $(k - 1)$ transposições de linhas adajacentes. Desta maneira obtivemos B a partir de A por meio de $k + (k-1) = 2k - 1$ transposições de linhas adjacentes. Assim,

$$D(B) = (-1)^{2k-1}D(A) = -D(A).$$

Suponhamos que A seja uma $n \times n$ matriz arbitrária com duas linhas iguais, digamos $\alpha_i = \alpha_j$, com $i < j$. Se $j = i + 1$, então A tem duas linhas iguais e adajacentes e $D(A) = 0$. Se $j > i + 1$ transpomos α_{i+1} e α_j e a matriz resultante B possui duas linhas iguais e adjacentes, portanto $D(B) = 0$. Por outro lado $D(B) = -D(A)$, portanto $D(A) = 0$.

Definição. *Se* $n > 1$ *e* A *é uma* $n \times n$ *matriz sôbre* K, *indiquemos por* $A(i|j)$ *a* $(n - 1) \times (n - 1)$ *matriz obtida de* A *retirando-se a* i*-ésima linha e a* j*-ésima coluna. Se* D *é uma função* $(n - 1)$*-linear e* A *é uma* $n \times n$ *matriz, colocamos* $D_{ij}(A) = D[A(i|j)]$.

Teorema 1. *Seja* $n > 1$ *e seja* D *uma função* $(n - 1)$*-linear alternada sôbre as* $(n - 1) \times (n - 1)$ *matrizes sôbre* K. *Para cada* j, $1 \leq j \leq n$, *a função* E_j *definida por*

$$E_j(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} D_{ij}(A) \qquad (5\text{—}4)$$

é uma função n*-linear alternada sôbre* $n \times n$ *matrizes* A. *Se* D *é uma função determinante,* E_j *também o é.*

Demonstração. Se A é uma $n \times n$ matriz, $D_{ij}(A)$ é independente da i-ésima linha de A. Como D é $(n - 1)$-linear, é claro que D_{ij} é uma função linear de qualquer linha exceto a linha i. Portanto $A_{ij}D_{ij}(A)$ é uma função n-linear de A. Uma combinação linear de funções n-lineares é n-linear, logo E_j é n-linear. Para demonstrar que E_j é alternada, bastará demonstrar que $E_j(A) = 0$ sempre que A tiver duas linhas iguais e adjacentes. Suponhamos que $\alpha_k = \alpha_{k+1}$. Se $i \neq k$ e $i \neq k + 1$, a matriz $(Ai|j)$ tem duas linhas iguais e então $D_{ij}(A) = 0$. Portanto

$$E_j(A) = (-1)^{k+j}A_{kj}D_{kj}(A) + (-1)^{k+1+j}A_{(k+1)j}D_{(k+1)j}(A).$$

Como $\alpha_k = \alpha_{k+1}$,

$$A_{kj} = A_{(k+1)j} \quad \text{e} \quad A(k|j) = A(k + 1|j).$$

Então é evidente que $E_j(A) = 0$.

Suponhamos agora que D seja uma função determinante. Se

$I^{(n)}$ é a $n \times n$ matriz unidade, então $I^{(n)}(j|j)$ é a $(n-1) \times (n-1)$ matriz unidade $I^{(n-1)}$. Como $I^{(n)}_{ij} = \delta_{ij}$, decorre de (5—4) que

$$E_j(I^{(n)}) = D(I^{(n-1)}). \tag{5—5}$$

Ora, $D(I^{(n-1)}) = 1$, de modo que $E_j(I^{(n)}) = 1$ e E_j é uma função determinante.

Corolário. *Seja* K *um anel comutativo com elemento unidade e seja* n *um inteiro positivo. Existe pelo menos uma função determinante sôbre o conjunto das* n $\times$ n *matrizes sôbre* K.

Demonstração. Demonstramos a existência de um função determinante sôbre 1×1 matrizes sôbre K e mesmo sôbre 2×2 matrizes sôbre K. O Teorema 1 nos diz explìcitamente como construir uma função determinante sôbre $n \times n$ matrizes, a partir de uma função determinante sôbre $(n-1) \times (n-1)$ matrizes. O corolário decorre por indução.

Exemplo 5. Se B é uma 2×2 matriz sôbre K façamos

$$|B| = B_{11}B_{22} - B_{12}B_{21}.$$

Então $|B| = D(B)$, onde D *é a função determinante sôbre* 2×2 matrizes.

Seja

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$

uma 3×3 matriz sôbre K. Definindo E_1, E_2, E_3 como em (5—4), então

$$E_1(A) = A_{11}\begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} - A_{21}\begin{vmatrix} A_{12} & A_{13} \\ A_{32} & A_{33} \end{vmatrix} + A_{31}\begin{vmatrix} A_{12} & A_{13} \\ A_{22} & A_{23} \end{vmatrix} \tag{5-6}$$

$$E_2(A) = -A_{12}\begin{vmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{vmatrix} + A_{22}\begin{vmatrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{vmatrix} - A_{32}\begin{vmatrix} A_{11} & A_{13} \\ A_{21} & A_{23} \end{vmatrix} \tag{5-7}$$

$$E_3(A) = A_{13}\begin{vmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{vmatrix} - A_{23}\begin{vmatrix} A_{11} & A_{12} \\ A_{31} & A_{32} \end{vmatrix} + A_{33}\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} \tag{5-8}$$

Decorre do Teorema 1 que E_1, E_2 e E_3 são funções determinantes. Na realidade, como mostraremos posteriormente, $E_1 = E_2 = E_3$, mas isto ainda não é evidente mesmo neste caso simples. Poder-se-ia.

contudo, verificar diretamente êste fato desenvolvendo-se cada uma das expressões acima. Em vez de fazer isto daremos exemplos particulares.

(a) Seja $K = R[x]$ e

$$A = \begin{bmatrix} x-1 & x^2 & x^3 \\ 0 & x-2 & 1 \\ 0 & 0 & x-3 \end{bmatrix}.$$

Então

$$E_1(A) = (x-1)\begin{vmatrix} x-2 & 1 \\ 0 & x-3 \end{vmatrix} = (x-1)(x-2)(x-3)$$

$$E_2(A) = -x^2\begin{vmatrix} 0 & 1 \\ 0 & x-3 \end{vmatrix} + (x-2)\begin{vmatrix} x-1 & x^3 \\ 0 & x-3 \end{vmatrix}$$

$$= (x-1)(x-2)(x-3)$$

$$E_3(A) = x^3\begin{vmatrix} 0 & x-2 \\ 0 & 0 \end{vmatrix} - \begin{vmatrix} x-1 & x^2 \\ 0 & 0 \end{vmatrix} + (x-3)\begin{vmatrix} x-1 & x^2 \\ 0 & x-2 \end{vmatrix}$$

$$= (x-1)(x-2)(x-3).$$

(b) Seja $K = R$ e

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

Então

$$E_1(A) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$$

$$E_2(A) = -\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = 1$$

$$E_3(A) = -\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = 1.$$

Exercícios

1. Cada uma das expressões seguintes, define uma função D sôbre o conjunto das 3×3 matrizes sôbre o corpo dos números reais. Em quais dêstes casos D é um função trilinear?

(a) $D(A) = A_{11} + A_{22} + A_{33}$;
(b) $D(A) = (A_{11})^2 + 3A_{11}A_{22}$;

(c) $D(A) = A_{11}A_{12}A_{33}$;
(d) $D(A) = A_{13}A_{22}A_{32} + 5A_{12}A_{22}A_{32}$;
(e) $D(A) = 0$;
(f) $D(A) = 1$.

2. Verificar diretamente que as três funções E_1, E_2, E_3 definidas por (5—6), (5—7) e (5—8) são idênticas.

3. Seja K um anel comutativo com elemento unidade. Se A é uma 2×2 matriz sôbre K, a **adjunta clássica** de A é a 2×2 matriz adj A definida por

$$\text{adj } A = \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix}.$$

Se det indica a única função determinante sôbre 2×2 matrizes sôbre K, mostrar que

(a) $(\text{adj } A)A = A(\text{adj } A) = (\det A)I$;
(b) $\det(\text{adj } A) = \det(A)$;
(c) $\text{adj}(A^t) = (\text{adj } A)^t$.

(A^t indica a transposta de A).

4. Seja A uma 2×2 matriz sôbre um corpo F. Mostrar que A é inversível se, e sòmente se, $\det A \neq 0$. Quando A é inversível, dar uma fórmula para A^{-1}.

5. Seja A uma 2×2 matriz sôbre um corpo F e suponhamos que $A^2 = 0$. Mostrar para cada escalar c que $\det(cI - A) = c^2$.

6. Seja K um subcorpo do corpo dos números complexos e n um inteiro positivo. Sejam $j_1, \ldots, j_n$ e $k_1, \ldots, k_n$ inteiros positivos menores ou iguais a n. Para uma $n \times n$ matriz A sôbre K definamos

$$D(A) = A(j_1, k_1)A(j_2, k_2) \ldots A(j_n, k_n).$$

Demonstrar que D é n-linear se, e sòmente se, os inteiros $j_1, \ldots, j_n$ são distintos.

7. Seja K um anel comutativo com elemento unidade. Mostrar que a função determinante sôbre as 2×2 matrizes A sôbre K é alternada e bilinear como uma função das colunas de A.

8. Seja K um anel comutativo com elemento unidade. Definamos uma função D sôbre K pela regra

$$D(A) = A_{11} \det \begin{bmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{bmatrix} - A_{12} \det \begin{bmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{bmatrix} + A_{13} \det \begin{bmatrix} A_{2} & A_{22} \\ A_{31} & A_{32} \end{bmatrix}.$$

Mostrar que D é alternada e trilinear como uma função das colunas de A.

9. Seja K um anel comutativo com elemento unidade e D uma função n-linear alternada sôbre as $n \times n$ matrizes sôbre K. Mostrar que

(a) $D(A) = 0$, se uma das linhas de A é 0.
(b) $D(B) = D(A)$ se B é obtida a partir de A somando-se um múltiplo escalar de uma linha de A a outra.

10. Seja F um corpo, A uma 2×3 matriz sôbre F e (c_1, c_2, c_3) o vetor em F_3 definido por

$$c_1 = \begin{vmatrix} A_{12} & A_{13} \\ A_{22} & A_{23} \end{vmatrix} \qquad c_2 = \begin{vmatrix} A_{13} & A_{11} \\ A_{23} & A_{21} \end{vmatrix} \qquad c_3 = \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix}.$$

Mostrar que

(a) pôsto $(A) = 2$ se, e sòmente se, $(c_1, c_2, c_3) \neq 0$;

(b) se A tem pôsto 2, então (c_1, c_2, c_3) é uma base do espaço-solução do sistema de equações $AX = 0$.

11. Seja K um anel comutativo com elemento unidade e seja D uma função bilinear alternada sôbre as 2×2 matrizes sôbre K. Mostrar que $D(A) = (\det A)D(I)$ para qualquer A. Utilizar agora êste resultado (não são permitidos cálculos com os elementos) para mostrar que $\det(AB) = \det(A)\det(B)$ para quaisquer 2×2 matrizes A e B sôbre K.

12. Seja F um corpo e D uma função sôbre as $n \times n$ matrizes sôbre F (com valores em F). Suponhamos que $D(AB) = D(A)\,D(B)$ para quaisquer A, B. Mostrar que ou $D(A) = 0$ para qualquer A ou $D(I) = 1$. No último caso mostrar que $D(A) \neq 0$ sempre que A é inversível.

13. Seja R o corpo dos números reais e seja D uma função sôbre as 2×2 matrizes sôbre R, com valores em R, tal que $D(AB) = D(A)D(B)$ para quaisquer A, B. Suponhamos também que

$$D\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right) \neq D\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right).$$

Demonstrar o seguinte:

(a) $D(0) = 0$;
(b) $D(A) = 0$ se $A^2 = 0$;
(c) $D(B) = -D(A)$ se B é obtida transpondo-se duas linhas (ou colunas de A;
(d) $D(A)$ se uma linha (ou coluna) de A é 0;
(e) $D(A) = 0$ sempre que A é singular.

14. Seja A uma 2×2 matriz sôbre um corpo F. Então o conjunto das matrizes da forma $f(A)$, onde f é um polinômio sôbre F, é um anel comutativo K com elemento unidade. Se B é uma 2×2 matriz sôbre K, o determinante de B é então uma 2×2 matriz sôbre F, da forma $f(A)$. Suponhamos que I seja a 2×2 matriz unidade sôbre F e que B seja a 2×2 matriz sôbre K

$$B = \begin{bmatrix} A - A_{11}I & -A_{12}I \\ -A_{21}I & A - A_{22}I \end{bmatrix}.$$

Mostrar que $\det B = f(A)$, onde $f = x^2 - (A_{11} + A_{22})x + \det A$, e também que $f(A) = 0$.

5.3 Permutações e a Unicidade de Determinantes

Nesta seção demonstraremos a unicidade da função determinante sôbre as $n \times n$ matrizes sôbre K. A demonstração nos levará, de modo bem natural, a considerar permutações e algumas de suas pròpriedades básicas.

Suponhamos que D seja uma função n-linear alternada sôbre as $n \times n$ matrizes sôbre K. Seja A uma $n \times n$ matriz sôbre K com linhas $\alpha_1, \alpha_2, \ldots, \alpha_n$. Se indicarmos as linhas da $n \times n$ matriz unidade sôbre K por $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$, então

$$(5\text{-}9) \qquad \alpha_i = \sum_{j=1}^{n} A(i,j)\epsilon_j, \qquad 1 \leq i \leq n.$$

Logo

$$\begin{aligned} D(A) &= D(\sum_j A(1,j)\epsilon_j, \alpha_2, \ldots, \alpha_n) \\ &= \sum_j A(1,j)D(\epsilon_j, \alpha_2, \ldots, \alpha_n). \end{aligned}$$

Se agora substituirmos α_2 por $\sum_k A(2, k)\epsilon_k$ veremos que

$$D(\epsilon_j, \alpha_2, \ldots, \alpha_n = \sum_k A(2, k)D(\epsilon_j, \epsilon_k, \ldots, \alpha_n).$$

Assim

$$D(A) = \sum_{j,k} A(1,j)A(2, k)D(\epsilon_j, \epsilon_k, \ldots, \alpha_n).$$

Em $D(\epsilon_j, \epsilon_k, \ldots, \alpha_n)$ substituímos em seguida α_3 por $\sum A(3, {}_l)\epsilon_l$ e assim por diante. Obtemos finalmente uma expressão complicada mas teòricamente importante para $D(A)$, a saber

$$(5\text{-}10) \quad D(A) = \sum_{k_1, k_2, \ldots, k_n} A(1, k_1)A(2, k_2) \ldots A(n, k_n)D(\epsilon_{k_1}, \epsilon_{k_2}, \ldots \epsilon_{k_n}).$$

Em (5-10) a soma é estendida a tôdas as seqüências $(k_1, k_2, \ldots, k_n)$ de inteiros positivos menores ou iguais a n. Isto mostra que D é uma soma finita de funções do tipo descrito por (5-2). Deve-se notar que (5-10) é uma conseqüência apenas da hipótese de que D é n-linear e também que um caso particular de (5-10) foi obtido no Exemplo 2. Como D é alternada,

$$D = (\epsilon_{k_1}, \epsilon_{k_2}, \ldots, \epsilon_{k_n}) = 0$$

sempre que dois dos índices k_i são iguais. Uma seqüência $(k_1, k_2, \ldots$ $\ldots, k_n)$ de inteiros positivos menores ou iguais a n, com a propriedade de não existirem dois k_i iguais, é denominada uma **permutação de grau** n. Portanto, *em (5—10), precisamos somar considerando apenas as seqüências que sejam permutações de grau* n.

Com uma seqüência finita, ou n-upla, *é uma função definida* sôbre os n primeiros inteiros positivos, uma permutação de grau n pode ser definida como uma função bijetora do conjunto $\{1, 2, \ldots, n\}$ em si mesmo. Tal função σ corresponde à n-upla $(\sigma 1, \sigma 2, \ldots, \sigma n)$ e é simplesmente uma regra para ordenar $1, 2, \ldots, n$ de alguma maneira bem definida.

Se D é uma função n-linear alternada e A é uma $n \times n$ matriz sôbre K, então temos

$$D(A) = \sum_{\sigma} A(1, \sigma 1) \ldots A(n, \sigma n) D(\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}) \tag{5-11}$$

onde a soma é estendida a tôdas as permutações σ distintas e de grau n.

A seguir mostraremos que

$$D(\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}) = \pm D(\epsilon_1, \ldots, \epsilon_n) \tag{5-12}$$

onde o sinal $\pm$ depende sòmente da permutação σ. A razão para isto é a que segue. A seqüência $(\sigma 1, \sigma 2, \ldots, \sigma n)$ pode ser obtida da seqüência $(1, 2, \ldots, n)$ por um número finito de transposições de pares de elementos. Por exemplo, se $\sigma 1 \neq 1$, podemos transpor 1 e $\sigma 1$ obtendo $(\sigma 1, \ldots, 1, \ldots)$. Procedendo desta maneira chegaremos à seqüência $(\sigma 1, \ldots, \sigma n)$ após n ou menos transposições de pares. Como D é alternada, o sinal de seu valor muda cada vez que transpomos duas das linhas ϵ_i e ϵ_j. Assim, se passamos de $(1, 2, \ldots, n)$ a $(\sigma 1, \sigma 2, \ldots, \sigma n)$ por meio de m transposições de pares (i, j) teremos

$$D(\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}) = (-1)^m D(\epsilon_1, \ldots, \epsilon_n).$$

Em particular, se D é uma função determinante

$$D(\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}) = (-1)^m \tag{5-13}$$

onde m depende sòmente de σ e não de D. Assim, tôdas as funções determinantes associam o mesmo valor à matriz com linhas $\epsilon_{\sigma 1}, \ldots, \ldots, \epsilon_{\sigma n}$, e êste valor é 1 ou -1.

Um fato básico sôbre permutações é o seguinte: se σ é uma permutação de grau n, pode-se passar da seqüência $(1, 2, \ldots, n)$ à seqüência $(\sigma 1, \sigma 2, \ldots, \sigma n)$ por meio de uma série de transposições de pares e isto pode ser feito de diversas maneiras; contudo, qualquer que seja a maneira pela qual isto é feito, o número de transposições usadas é sempre par ou sempre ímpar. A permutação é então denominada **par** ou **ímpar**, respectivamente. Define-se o sinal de uma permutação por

$$\text{sinal } \sigma = \begin{cases} 1, & \text{se } \sigma \text{ é par} \\ -1, & \text{se } \sigma \text{ é ímpar} \end{cases}$$

com o símbolo "1" indicando aqui o inteiro 1. Mostraremos abaixo que esta propriedade básica das permutações pode ser deduzida do que já sabemos sôbre funções determinantes. Suponhamos por ora que isto seja verdade. Então o inteiro m que aparece em (5—13) é sempre par se σ é uma permutação par e é sempre ímpar se σ é uma permutação ímpar. Para qualquer função D n-linear e alternada temos então

$$D(\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}) = (\text{sinal } \sigma) D(\epsilon_1, \ldots, \epsilon_n)$$

e usando (5—11)

$$D(A) = [\sum_{\sigma} (\text{sinal } \sigma) A(1, \sigma 1) \ldots A(n, \sigma n)]\, D(I). \tag{5-14}$$

É claro que I indica a $n \times n$ matriz unidade.

De (5—14) vemos que existe exatamente uma função determinante sôbre as $n \times n$ matrizes sôbre K. Se indicarmos esta função por det, ela será dada por

$$\det(A) = \sum_{\sigma} (\text{sinal } \sigma) A(1, \sigma 1) \ldots A(n, \sigma n) \tag{5-15}$$

sendo a soma estendida a tôdas as permutações σ distintas e de grau n. Podemos resumir formalmente como segue.

Teorema 2. *Seja* K *um anel comutativo com elemento unidade e seja* n *um inteiro positivo. Existe exatamente uma função determinante sôbre o conjunto das* m $\times$ n *matrizes sôbre* K, *que é a função det definida por* (5-15). *Se* D *é uma função* n-*linear alternada arbitrária sôbre o conjunto das* n $\times$ n *matrizes sôbre* K, *então, para tôda matriz* A *dêste tipo,*

$$D(A) = (\det A) D(I).$$

Êste é o teorema que procurávamos, mas deixamos uma lacuna na demonstração. Essa lacuna é a demonstração de que, para uma dada permutação σ, quando passamos de $(1, 2, \ldots, n)$ para $(\sigma 1, \sigma 2, \ldots, \sigma n)$ transpondo pares, o número de transposições é sempre par ou sempre ímpar. Êste fato básico de combinatória pode ser demonstrado sem nenhuma referência a determinantes; contudo, gostaríamos de salientar como êle decorre da *existência* de uma função determinante sôbre $n \times n$ matrizes.

Tomemos K como sendo o anel dos inteiros. Seja D uma função determinante sôbre as $n \times n$ matrizes sôbre K. Seja σ uma permutação de grau n e suponhamos que passemos de $(1, 2, \ldots, n)$ a $(\sigma 1, \sigma 2, \ldots, \sigma n)$ por meio de m transposições de pares (i, j), $i \neq j$. Como mostramos em (5—13)

$$(-1)^m = D(\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}),$$

isto é, o número $(-1)^m$ tem que ser o valor de D sôbre a matriz de linhas $\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}$. Se

$$D\epsilon_{\sigma 1}(, \ldots, \epsilon_{\sigma n}) = 1,$$

então m tem que ser par. Se

$$D(\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}) = -1,$$

então m tem que ser ímpar.

Como temos uma fórmula explícita para o determinante de uma $n \times n$ matriz e esta fórmula envolve as permutações de grau n, vamos concluir esta seção fazendo mais algumas observações sôbre permutações. Primeiro, notemos que existem exatamente $n! = 1 . 2 \ldots n$ permutações de grau n, pois se σ é uma tal permutação, existem n escolhas possíveis para $\sigma 1$; uma vez feita, existem $(n - 1)$ possibilidades para $\sigma 2$; depois $(n - 2)$ possibilidades para $\sigma 3$ e assim por diante. Logo, existem

$$n(n - 1)(n - 2) \ldots 2 . 1 = n!$$

permutação σ. A fórmula (5—15) para det (A) fornece desta maneira det (A) como uma soma de $n!$ têrmos, um para cada permutação de grau n. Um têrmo genérico é um produto

$$A(1, \sigma 1) \ldots A(n, \sigma n) \text{ de } n \text{ elementos de } A,$$

um elemento de cada linha e um de cada coluna, e é acompanhado de um sinal "+" ou "—" conforme σ seja uma permutação par ou ímpar.

Quando as permutações são consideradas como funções bijetoras do conjunto $\{1, 2, \ldots, n\}$ em si mesmo, é possível definir um produto de permutações. O produto de σ por τ será simplesmente a função composta $\sigma\tau$ definida por

$$(\sigma\tau)(i) = \sigma(\tau(i)).$$

Se ϵ indica a permutação idêntica, $\epsilon(i) = i$, então cada σ possui uma inversa σ^{-1} tal que

$$\sigma\sigma^{-1} = \sigma^{-1}\sigma = \epsilon.$$

Pode-se resumir estas observações dizendo que, em relação à operação de composição, o conjunto das permutações de grau n é um grupo. Êste grupo é usualmente denominado o **grupo simétrico de grau n**.

Do ponto de vista de produtos de permutações, a propriedade fundamental do sinal de uma permutação é que

$$\text{(5-16)} \qquad \text{sinal } (\sigma\tau) = (\text{sinal } \sigma)(\text{sinal } \tau).$$

Em outras palavras, $\sigma\tau$ é uma permutação par se σ e τ são ambas pares ou ambas ímpares enquanto que $\sigma\tau$ é ímpar se uma das duas permutações é ímpar e a outra é par. Pode-se ver isto pela definição do sinal em têrmos de permutações sucessivas de pares (i, j). Poderá ser também instrutivo se ressaltarmos como a igualdade sinal $(\sigma\tau) =$ $=$ (sinal σ) (sinal τ) decorre de uma propriedade fundamental dos determinantes.

Seja K o anel dos inteiros e sejam σ e τ permutações de grau n. Sejam $\epsilon_1, \ldots, \epsilon_n$ as linhas da $n \times n$ matriz unidade sôbre K, seja A a matriz de linhas $\epsilon_{\tau 1}, \ldots, \epsilon_{\tau n}$ e seja B a matriz de linhas $\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}$. A i-ésima linha de A contém exatamente um elemento não-nulo, a saber, o 1 na coluna τi. A partir disto, é fácil ver que $\epsilon_{\sigma\tau i}$ é a i-ésima linha da matriz produto AB. Ora,

$$\det(A) = \text{sinal } \tau, \quad \det(B) = \text{sinal } \sigma \quad \text{e } \det(AB) = \text{sinal } (\sigma\tau).$$

Portanto teremos sinal $(\sigma\tau) =$ (sinal σ) (sinal τ) desde que demonstremos o seguinte:

Teorema 3. *Seja* K *um anel comutativo com elemento unidade e sejam* A *e* B n $\times$ n *matrizes sôbre* K. *Então*

$$\det(AB) = (\det A)(\det B).$$

Demonstração. Seja B uma $n \times n$ matriz fixa sôbre K e para cada $n \times n$ matriz A definamos $D(A) = \det(AB)$. Se indicarmos as linhas de A por $\alpha_1, \ldots, \alpha_n$, então

$$D(\alpha_1, \ldots, \alpha_n) = \det(\alpha_1 B, \ldots, \alpha_n B).$$

Aqui $\alpha_j B$ indica a $1 \times n$ matriz que é o produto da $1 \times n$ matriz α_j pela $n \times n$ matriz B. Como

$$(c\alpha_i + \alpha_i')B = c\alpha_i B + \alpha_i' B$$

e det é n-linear, é fácil ver que D é n-linear. Se $\alpha_i = \alpha_j$, então, $\alpha_i B =$ $= \alpha_j B$, e como det é alternada

$$D(\alpha_1, \ldots, \alpha_n) = 0.$$

Logo, D é alternada. Ora, D é uma função n-linear alternada e pelo Teorema 2

$$D(A) = (\det A)D(I).$$

Mas $D(I) = \det(IB) = \det B$, portanto

$$\det(AB) = D(A) = (\det A)(\det B).$$

O fato de que sinal $(\sigma\tau) =$ (sinal σ) (sinal τ) é apenas um dos muitos corolários do Teorema 3. Consideraremos alguns dêsses corolários na próxima seção.

Exercícios

1. Se K é um anel comutativo com elemento unidade e A é a matriz sôbre K dada por

$$A = \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}$$

mostrar que $\det A = 0$.

2. Demonstrar que o determinante da matriz de Vandermonde

$$\begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix}$$

é $(b - a)(c - a)(c - b)$.

3. Enunciar explìcitamente as seis permutações de grau 3, dizer quais são ímpares e quais são pares e usar isto para dar uma fórmula completa (5—15) para o determinante de uma 3×3 matriz.

4. Sejam γ e τ as permutações de grau 4 definidas por $\sigma 1 = 2$, $\sigma 2 = 3$, $\sigma 3 = 4$, $\sigma 4 = 1$, $\tau 1 = 3$, $\tau 2 = 1$, $\tau 3 = 2$, $\tau 4 = 4$.

(a) σ é par ou ímpar? τ é par ou ímpar?

(b) Determinar $\sigma\tau$ e $\tau\sigma$.

5. Se A é uma $n \times n$ matriz inversível sôbre um corpo, mostrar que $\det A \neq 0$.

6. Seja A uma 2×2 matriz sôbre um corpo. Demonstrar que $\det(I + A) = 1 + \det A$ se, e sòmente se, traço $(A) = 0$.

7. Uma $n \times n$ matriz A é denominada **triangular** se $A_{ij} = 0$ sempre que $i > j$ ou se $A_{ij} = 0$ sempre que $i < j$. Demonstrar que o determinante de uma matriz triangular é o produto $A_{11}A_{22} \ldots A_{nn}$ de seus elementos diagonais.

8. Seja A uma 3×3 matriz sôbre o corpo dos números complexos. Formemos a matriz $xI - A$ cujos elementos são polinômios, sendo o elemento i, j desta matriz o polinômio $\delta_{ij}x - A_{ij}$. Se $f = \det(xI - A)$, mostrar que f é um polinômio unitário de grau 3. Escrevendo

$$f = (x - c_1)(x - c_2)(x - c_3)$$

com números complexos c_1, c_2, e c_3, demonstrar que

$$c_1 + c_2 + c_3 = \text{traço}\,(A) \text{ e } c_1c_2c_3 = \det A.$$

9. Seja n um inteiro positivo e F um corpo. Se σ é uma permutação de grau n, demonstrar que a função

$$T(x_1, \ldots, x_n) = (x_{\sigma 1}, \ldots, x_{\sigma n})$$

é um operador linear inversível sôbre F^n.

10. Seja F um corpo, n um inteiro positivo e S o conjunto das $n \times n$ matrizes sôbre F. Seja V o espaço vetorial de tôdas as funções de S em F. Seja W o conjunto das funções n-lineares alternadas sôbre S. Demonstrar que W é um subespaço de V. Qual é a dimensão de W?

11. Seja T um operador linear sôbre F^n. Definamos

$$D_T(\alpha_1, \ldots, \alpha_n) = \det (T\alpha_1, \ldots, T\alpha_n)$$

(a) Mostrar que D_T é uma função n-linear alternada.

(b) Se

$$c = \det (T\epsilon_1, \ldots, T\epsilon_1)$$

mostrar que para n quaisquer vetores $\alpha_1, \ldots, \alpha_n$ temos

$$\det (T\alpha_1, \ldots, T\alpha_n) = c \det (\alpha_1, \ldots, \alpha_n).$$

(c) Se $\mathcal{B}$ é uma base ordenada arbitrária de F^n e A é a matriz de T em relação à base ordenada $\mathcal{B}$, mostrar que $\det A = c$.

(d) Qual é um nome razoável para o escalar c?

12. Se σ é uma permutação de grau n e A é uma $n \times n$ matriz sôbre o corpo F com vetores-linhas $\alpha_1, \ldots, \alpha_n$, indiquemos por $\sigma(A)$ a $n \times n$ matriz com vetores-linhas $\alpha_{\sigma 1}, \ldots, \alpha_{\sigma n}$.

(a) Demonstrar que $\sigma(AB) = \sigma(A)B$ e em particular que $\sigma(A) = \sigma(I)A$.

(b) Se T o operador linear do Exercício 9, demonstrar que a matriz T em relação à base ordenada canônica é $\sigma(I)$.

(c) $\sigma^{-1}(I)$ é a matriz inversa de $\sigma(I)$?

(d) É verdade que $\sigma(A)$ é semelhante a A?

13. Demonstrar qua a função sinal sôbre permutações é única no seguinte sentido: se f é uma função qualquer que associa a cada permutação de grau n um inteiro e se $f(\sigma\tau) = f(\sigma) f(\tau)$, então f é idênticamente nula, ou f é idênticamente 1 ou f é a função sinal.

5.4 Propriedades Adicionais dos Determinantes

Nesta seção relataremos algumas das propriedades úteis da função determinante sôbre $n \times n$ matrizes. Talvez a primeira coisa que devamos destacar seja a que segue. Em nossa discussão de $\det A$, as linhas de A desempenharam um papel privilegiado. Como não existe nenhuma diferença fundamental entre linhas e colunas, pode-se muito bem esperar que $\det A$ seja uma função n-linear alternada das colunas de A. Isto ocorre e, para demonstrá-lo, basta mostrar que

$$\det (A^t) = \det (A) \tag{5-17}$$

onde A^t indica a transposta de A.

Se σ é uma permutação de grau n

$$A^t(i, \sigma i) = A(\sigma i, i).$$

Da expressão (5-15) tem-se então

$$\det (A^t) = \sum_\sigma (\text{sinal } \sigma) A(\sigma 1, 1) \ldots A(\sigma n, n).$$

Quando $i = \sigma^{-1}j$, $A(\sigma i, i) = A(j, \sigma^{-1}j)$. Assim

$$A(\sigma 1, 1) \ldots A(\sigma n, n) = A(1, \sigma^{-1}1) \ldots A(n, \sigma^{-1}n).$$

Como $\sigma\sigma^{-1}$ é a permutação idêntica,

$$(\text{sinal } \sigma)(\text{sinal } \sigma^{-1}) = 1, \text{ ou seja, sinal } (\sigma^{-1}) = \text{sinal } (\sigma).$$

Além disso, quando σ percorre tôdas as permutações de grau n, σ^{-1} também o faz.

Portanto

$$\begin{aligned}\det (A^t) &= \sum_\sigma (\text{sinal } \sigma^{-1}) A(1, \sigma^{-1}1) \ldots A(n, \sigma^{-1}n)\\ &= \det A,\end{aligned}$$

o que demonstra (5—17).

Em certas ocasiões é preciso calcular certos determinantes particulares. Quando isto é necessário, é freqüentemente útil tirar vantagem do fato seguinte: *se* B *é obtida de* A *somando-se um múltiplo de uma linha de* A *a outra* (*ou um múltiplo de uma coluna a outra*), *então*

(5-18) $$\det B = \det A.$$

Demonstraremos a afirmação relativa às linhas. Seja B obtida de A somando-se $c\alpha_j$ a α_i, onde $i < j$. Como det é linear como uma função da i-ésima linha

$$\begin{aligned}\det B &= \det A + c \det (\alpha_1, \ldots, \alpha_j, \ldots, \alpha_j, \ldots, \alpha_n)\\ &= \det A.\end{aligned}$$

Outro fato útil é o seguinte: consideremos uma $n \times n$ matriz da forma em blocos

$$\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

onde A é uma $r \times r$ matriz, C é uma $s \times s$ matriz, B é $r \times s$ e 0 indica a $s \times r$ matriz nula. Então

(5-12) $$\begin{matrix}\det\\ \det\end{matrix} \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} = (\det A)(\det C).$$

Para demonstrar isto, definamos

$$D(A, B, C) = \det \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}.$$

Se fixarmos A e B, então D é alternada e s-linear como uma função das linhas de C. Assim, pelo Teorema 2

$$D(A, B, C) = (\det C)D(A, B, I)$$

onde I é a $s \times s$ matriz unidade. Subtraindo das linhas de B múltiplos das linhas de I e usando a afirmação (5-18) acima, obtemos

$$D(A, B, I) = D(A, 0, I).$$

Ora, $D(A, 0, I)$ é evidentemente alternada e r-linear como uma função das linhas de A. Assim

$$D(A, 0, I) = (\det A)D(I, 0, I).$$

Mas $D(I, 0, I) = 1$, logo

$$\begin{aligned} D(A, B, C) &= (\det C)D(A, B, I) \\ &= (\det C)D(A, 0, I) \\ &= (\det C)(\det A). \end{aligned}$$

Por um raciocínio do mesmo tipo, ou tomando transpostas

$$\det \begin{bmatrix} A & 0 \\ B & C \end{bmatrix} = (\det A)(\det C). \tag{5-20}$$

Exemplo 6. Suponhamos que K seja o corpo dos números racionais e que desejamos calcular o determinante da 4×4 matriz

$$A = \begin{bmatrix} 1 & -1 & 2 & 3 \\ 2 & 2 & 0 & 2 \\ 4 & 1 & -1 & -1 \\ 1 & 2 & 3 & 0 \end{bmatrix}.$$

Subtraindo das linhas 2, 3 e 4 múltiplos convenientes da linha 1, obtemos a matriz

$$\begin{bmatrix} 1 & -1 & 2 & 3 \\ 0 & 4 & -4 & -4 \\ 0 & 5 & -9 & -13 \\ 0 & 3 & 1 & -3 \end{bmatrix}$$

que, como sabemos por (5—18), terá o mesmo determinante que A.

Subtraindo da linha 3, $\frac{5}{4}$ da linha 2 e depois subtraindo da linha 4, $\frac{3}{4}$ da linha 2, obtemos

$$B = \begin{bmatrix} 1 & -1 & 2 & 3 \\ 0 & 4 & -4 & -4 \\ 0 & 0 & -4 & -8 \\ 0 & 0 & 4 & 0 \end{bmatrix}$$

e novamente $\det B = \det A$. A forma em blocos de B nos diz que

$$\det A = \det B = \begin{vmatrix} 1 & -1 \\ 0 & 4 \end{vmatrix} \begin{vmatrix} -4 & -8 \\ 4 & 0 \end{vmatrix} = 4(32) = 128.$$

Seja agora $n > 1$ e seja A uma $n \times n$ matriz sôbre K. No Teorema 1 mostramos como construir uma função determinante sôbre as $n \times n$ matrizes, dada uma tal função sôbre as $(n-1) \times (n-1)$ matrizes. Agora que demonstramos a unicidade da função determinante, a fórmula (5—4) nos diz o que segue. Se fixamos uma coluna arbitrária de índice j,

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \det A(i|j).$$

O escalar $(-1)^{i+j} \det A(i|j)$ é usualmente denominado o i, j **cofator** de A ou o cofator do elemento i, j de A. A fórmula acima para $\det A$ é então denominada o desenvolvimento de $\det A$ pelos cofatores da j-ésima coluna (ou, às vêzes, o desenvolvimento pelos menores da j-ésima coluna). Se colocarmos

$$C_{ij} = (-1)^{i+j} \det A(i|j)$$

então a fórmula acima dirá que para cada j

$$\det A = \sum_{i=1}^{n} A_{ij} C_{ij}$$

onde o cofator C_{ij} é $(-1)^{i+j}$ vêzes o determinante da $(n-1) \times \times (n-1)$ matriz obtida de A retirando-se a i-ésima linha e a j-ésima coluna de A.

Se $j \neq k$, então

$$\sum_{i=1}^{n} A_{ik} C_{ij} = 0.$$

De fato, substituamos a j-ésima coluna de A por sua k-ésima coluna e chamemos a matriz resultante de B. Então B possui duas colunas iguais, logo det $B = 0$. Como $B(i|j) = A(i|j)$ temos

$$\begin{aligned} 0 &= \det B \\ &= \sum_{i=1}^{n} (-1)^{i+j} B_{ij} \det B(i|j) \\ &= \sum_{i=1}^{n} (-1)^{i+j} A_{ik} \det A(i|j) \\ &= \sum_{i=1}^{n} A_{ik} C_{ij}. \end{aligned}$$

Estas propriedades dos cofatores podem ser resumidas por

$$\text{(5-21)} \qquad \sum_{i=1}^{n} A_{ik} C_{ij} = \delta_{jk} \det A.$$

A $n \times n$ matriz adj A, que é a transposta da matriz dos cofatores de A, é denominada a **adjunta clássica** de A. Assim

$$\text{(5-22)} \qquad (\operatorname{adj} A)_{ij} = C_{ji} = (-1)^{i+j} \det A(j|i).$$

As fórmulas (5-21) podem ser resumidas na equação matricial

$$\text{(5-23)} \qquad (\operatorname{adj} A)A = (\det A)I.$$

Desejamos também ver que $A(\operatorname{adj} A) = (\det A)I$. Como $A^t(i|j) = A(j|i)^t$ temos

$$(-1)^{i+j} \det A^t(i|j) = (-1)^{j+i} \det A(j|i)$$

que diz simplesmente que o i, j cofator de A^t é o j, i cofator de A. Assim

$$\text{(5-24)} \qquad \operatorname{adj}(A^t) = (\operatorname{adj} A)^t.$$

Aplicando (5—23) a A^t obtemos

$$(\operatorname{adj} A^t)A^t = (\det A^t)I = (\det A)I$$

e transpondo

$$A(\operatorname{adj} A^t)^t = (\det A)I.$$

Usando (5—23) obtemos o que queremos:

$$\text{(5-25)} \qquad A(\operatorname{adj} A) = (\det A)I.$$

Da mesma forma que para matrizes sôbre um corpo, uma $n \times n$ matriz A sôbre K é dita **inversível sôbre** K se existe uma $n \times n$ matriz A^{-1} com elementos em K tal que $AA^{-1} = A^{-1}A = I$. Se existe uma tal inversa, ela é única, pois o mesmo argumento usado no Capítulo 1 mostra que quando $BA = AC = I$ temos $B = C$. As fórmulas (5—23) e (5—25) nos dizem o seguinte sôbre a inversibilidade de matrizes sôbre K: se o elemento det A possui um inverso multiplicativo em K, então A é inversível a $A^{-1} = (\det A)^{-1} \operatorname{adj} A$ é a única inversa de A. Recìprocamente, é fácil ver que, se A é inversível sôbre K, o elemento det A é inversível em K, pois se $BA = I$ temos

$$1 = \det I = \det (A'B) = (\det A)(\det B).$$

O que demonstramos é o seguinte:

Teorema 4. *Seja* A *uma* n $\times$ n *matriz sôbre* K. *Então* A *é inversível sôbre* K *se, e sòmente se, det* A *é inversível em* K. *Quando* A *é inversível, a única inversa de* A *é*

$$A^{-1} = (\det A)^{-1} \operatorname{adj} A.$$

Em particular, uma n $\times$ n *matriz sôbre um corpo é inversível se, e sòmente se, seu determinante é diferente de zero.*

Gostaríamos de salientar que êste critério relativo a determinantes para a inversibilidade demonstra que uma $n \times n$ matriz com uma inversa à esquerda ou à direita é inversível. Esta demonstração é completamente independente da demonstração que fizemos no Capítulo 1 para matrizes sôbre um corpo. Gostaríamos também de ressaltar o que a inversibilidade significa para matrizes cujos elementos são polinômios. Se K é o anel de polinômios $F[x]$, os únicos elementos de K que são inversíveis são os polinômios constantes não-nulos. De fato, se f e g são polinômios e $fg = 1$, temos $\operatorname{gr}(f) + \operatorname{gr}(g) = 0$ de modo que $\operatorname{gr}(f) = \operatorname{gr}(g) = 0$, isto é, f e g são polinômios constantes. Portanto uma $n + n$ matriz sôbre o anel de polinômios $F[x]$ é inversível sôbre $F[x]$ se, e sòmente se, seu determinante é um polinômio constante não-nulo.

Exemplo 7. Seja $K = R[x]$, o anel de polinômios sôbre o corpo dos números reais. Sejam

$$A = \begin{bmatrix} x^2 + x & x + 1 \\ x - 1 & 1 \end{bmatrix} \quad B = \begin{bmatrix} x^2 - 1 & x + 2 \\ x^2 - 2x + 3 & x \end{bmatrix}.$$

Então, por meio de cálculos simples, det $A = x + 1$ e det $B = -6$.

Assim, A não é inversível sôbre K, enquanto que B é inversível sôbre K. Notemos que

$$\operatorname{adj} A = \begin{bmatrix} 1 & -x-1 \\ -x+1 & x^2+x \end{bmatrix} \qquad \operatorname{adj} B = \begin{bmatrix} x & -x-2 \\ -x^2+2x-3 & x^2-1 \end{bmatrix}$$

e $(\operatorname{adj} A)A = (x+1)I$, $(\operatorname{adj} B)B = -6I$. É claro que

$$B^{-1} = -\frac{1}{6}\begin{bmatrix} x & -x-2 \\ -x^2+2x-3 & 1-x^2 \end{bmatrix}$$

Exemplo 8. Seja K o anel dos inteiros e

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

Então $\det A = -2$ e

$$\operatorname{adj} A = \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}.$$

Assim, A não é inversível como uma matriz sôbre o anel dos inteiros; no entanto, podemos também considerar A como uma matriz sôbre o corpo dos números racionais. Se o fazemos, então A é inversível e

$$A^{-1} = -\frac{1}{2}\begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}.$$

Com relação a matrizes inversíveis, gostaríamos de mencionar mais um fato elementar. Matrizes semelhantes têm o mesmo determinante, isto é, se P é inversível sôbre K e $B = P^{-1}AP$, então $\det B = \det A$. Isto é evidente, pois

$$\det(P^{-1}AP) = (\det P^{-1})(\det A)(\det P) = \det A.$$

Esta observação simples torna possível definir o determinante de um operador linear sôbre um espaço vetorial de dimensão finita. Se T é um operador linear sôbre V, definimos o determinante de T como sendo o determinante de qualquer $n \times n$ matriz que represente T em relação a alguma base ordenada de V. Como tôdas essas matrizes são semelhantes, elas possuem o mesmo determinante e nossa definição faz sentido.

Gostaríamos agora de discutir a regra de Cramer para a resolução de sistemas de equações lineares. Suponhamos que A seja uma $n \times n$ matriz sôbre o corpo F e que desejamos resolver o sistema de equações lineares $AX = Y$ para uma dada n-upla $(y_1, \ldots, y_n)$. Se $AX = Y$, então

$$(\text{adj } A)AX = (\text{adj } A)Y$$

e portanto

$$(\det A)X = (\text{adj } A)Y.$$

Assim

$$\begin{aligned}(\det A)x_j &= \sum_{i=1}^{n} (\text{adj } A)_{ji}y_i \\ &= \sum_{i=1}^{n} (-1)^{i+j}y_i \det A(i|j).\end{aligned}$$

Esta última expressão é o determinante de $n \times n$ matriz obtida substituindo a j-ésima coluna de A por Y. Se $\det A = 0$, tudo isto não nos diz nada; contudo, se $\det A \neq 0$ temos o que é conhecido como a regra de Cramer. Seja A uma $n \times n$ matriz sôbre o corpo F tal que $\det A \neq 0$. Se $y_1, \ldots, y_n$ são escalares arbitrários em F, a única solução $X = A^{-1}Y$ do sistema de equações $AX = Y$ é dada por

$$x_j = \frac{\det B_j}{\det A}, \qquad j = 1, \ldots, n$$

onde B_j é a $n \times n$ matriz obtida de A substituindo a j-ésima coluna de A por Y.

Concluindo êste capítulo, gostaríamos de fazer alguns comentários que sirvam para colocar os determinantes naquilo que acreditamos ser a perspectiva apropriada. De vez em quando, é necessário calcular alguns determinantes particulares e esta seção foi parcialmente dedicada a técnicas que irão facilitar êsse trabalho. No entanto, o papel principal dos determinantes neste livro é teórico. Não se discute a beleza de fatos tais como a regra de Cramer. Mas a regra de Cramer é um instrumento ineficiente para resolver sistemas de equações lineares, principalmente porque envolve cálculos em demasia. Portanto, deve-se concentrar no que a regra de Cramer diz, e não em como calcular por meio dela. De fato, refletindo sôbre êste capítulo, esperamos que o leitor coloque mais ênfase na compreensão do que a função determinante é e como ela se comporta do que no modo de calcular determinantes de matrizes particulares.

Exercícios

1. Usar a fórmula da adjunta clássica para calcular as inversas das 3×3 matrizes reais seguintes:

$$\begin{bmatrix} -2 & 3 & -2 \\ 6 & 0 & 3 \\ 4 & 1 & -1 \end{bmatrix} \qquad \begin{bmatrix} \cos\theta & 0 & -\text{sen }\theta \\ 0 & 1 & 0 \\ \text{sen }\theta & 0 & \cos\theta \end{bmatrix}$$

2. Usar a regra de Cramer para resolver os seguintes sistemas de equações lineares sôbre o corpo dos números racionais:

(a)
$$\begin{aligned} x + y + z &= 11 \\ 2x - 6y - z &= 0 \\ 3x + 4y + 2z &= 0 \end{aligned}$$

(b)
$$\begin{aligned} 3x - 2y &= 7 \\ 3y - 2z &= 6 \\ 3z - 2x &= -1 \end{aligned}$$

3. Uma $n \times n$ matriz A sôbre um corpo F é **anti-simétrica** se $A^t = -A$. Se A é uma $n \times n$ matriz anti-simétrica com elementos complexos e se n é ímpar, demonstrar que $\det A = 0$.

4. Uma $n \times n$ matriz A sôbre um corpo F é dita **ortogonal** se $AA^t = I$. Se A é ortogonal, mostrar que $\det A = \pm 1$. Dar um exemplo de uma matriz ortogonal para a qual $\det A = -1$.

5. Uma $n \times n$ matriz A sôbre o corpo dos números complexos é dita **unitária** se $AA^* = I$ (A^* indica a transposta conjugada de A). Se A é unitária, mostrar que $|\det A| = 1$.

6. Sejam T e U operadores lineares sôbre o espaço vetorial V de dimensão finita. Demonstrar que

(a) $\det (TU) = (\det T)(\det U)$;
(b) T é inversível se, e sòmente se, $\det T \neq 0$.

7. Seja A uma $n \times n$ matriz sôbre K, um anel comutativo com elemento unidade. Suponhamos que A seja da forma em blocos

$$A = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & A_k \end{bmatrix}$$

onde A_j é uma $r_j \times r_j$ matriz. Demonstrar que

$$\det A = (\det A_1)(\det A_2) \dots (\det A_k).$$

8. Seja V o espaço vetorial das $n \times n$ matrizes sôbre o corpo F. Seja B um elemento fixo de V e seja T_B o operador linear sôbre V definido por $T_B(A) = AB - BA$. Mostrar que $\det T_B = 0$.

9. Seja A uma $n \times n$ matriz sôbre um corpo, $A \neq 0$. Se r é um inteiro positivo arbitrário entre 1 e n, uma $r \times r$ **submatriz** de A é qualquer $r \times r$ matriz obtida retirando $(n - r)$ linhas e $(n - r)$ colunas de A. O **pôsto-determinante** de A é o maior inteiro positivo r tal que alguma $r \times r$ submatriz de A possua um determinante não-nulo. Demonstrar que o pôsto-determinante de A é igual ao pôsto-linha de A (= pôsto-coluna de A).

10. Seja A uma $n \times n$ matriz sôbre o corpo F. Demonstrar que existem no máximo n escalares c distintos em F tais que $\det (cI - A) = 0$.

11. Sejam A e B $n \times n$ matrizes sôbre o corpo F. Mostrar que se A é inversível existem no máximo n escalares c em F para os quais a matriz $cA + B$ não é inversível.

12. Seja V o espaço vetorial das $n \times n$ matrizes sôbre F, B uma $n \times n$ matriz fixa sôbre F e sejam L_B e R_B os operadores lineares sôbre V definidos por $L_B(A) = BA$ e $R_B(A) = AB$. Mostrar que

(a) $\det L_B = (\det B)^n$;
(b) $\det R_B = (\det B)^n$.

13. Seja V o espaço vetorial das $n \times n$ matrizes sôbre o corpo dos números complexos e seja B uma $n \times n$ matriz sôbre C, fixa. Definamos um operador linear M_B sôbre V por $M_B(A) = BAB^*$, onde $B^* = \overline{B^t}$. Mostrar que

$$\det M_B = |\det B|^{2n}.$$

Seja agora H o conjunto de tôdas as matrizes hermitianas em V. (A é hermitiana se $A = A^*$.) Então H é um espaço vetorial sôbre o corpo dos números *reais*. Mostrar que a função T_B definida por $T_B(A) = BAB^*$ é um operador linear sôbre o espaço vetorial real H e depois mostrar que $T_B = |\det B|^{2n}$. (*Sugestão*: Ao calcular $\det T_B$, mostrar que V possui uma base constituída de matrizes hermitianas e então mostrar que $\det T_B = \det M_B$.)

14. Sejam A, B, C, D $n \times n$ matrizes sôbre o corpo F tais que elas comutam. Mostrar que o determinante da $2n \times 2n$ matriz

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

é $\det (AD - BC)$.

CAPÍTULO 6

DECOMPOSIÇÕES EM SOMAS DIRETAS INVARIANTES

6.1 Decomposições em Somas Diretas

Mencionamos anteriormente que nossa meta principal é estudar transformações lineares sôbre espaços vetoriais de dimensão finita. Até êste ponto já vimos muitos exemplos particulares de transformações lineares e demonstramos alguns teoremas sôbre transfor-formações lineares arbitrárias. No caso de dimensão finita utilizamos bases ordenadas para representar essas transformações por meio de matrizes e essa representação nos ajuda a perceber o seu comportamento. Pesquisamos o espaço vetorial $L(V, W)$ das transformações lineares de um espaço em outro e também a álgebra linear $L(V, V)$ das transformações lineares de um espaço em si mesmo. Nos próximos dois capítulos, estaremos preocupados com operadores lineares. Nosso programa é tomar um operador linear T sôbre um espaço vetorial V de dimensão finita e "desmontá-lo para ver como êle funciona". Nosso método fundamental para o estudo de T será o de decompor o espaço subjacente V numa soma de subespaços, cada um dos quais é invariante sob T. Isto "decomporá" T numa soma de operadores mais simples. Passamos agora a apresentar as idéias fundamentais de que necessitaremos.

Se W_1 e W_2 são subespaços de V, já discutimos a soma $W = W_1 + W_2$, isto é, o subespaço de todos os $\alpha = \alpha_1 + \alpha_2$ com α_i em W_i. Uma situação particularmente agradável ocorre quando W_1 e W_2 são **disjuntos**, isto é, quando a interseção de W_1 e W_2 é o subespaço nulo. De fato, neste caso, um dado vetor α em W pode ser escrito sob a forma $\alpha = \alpha_1 + \alpha_2$, α_i em W_i, de uma única maneira. Isto resulta do fato de que se também temos $\alpha = \beta_1 + \beta_2$ com β_i em W_i, então

$$\alpha_1 + \alpha_2 = \beta_1 + \beta_2$$

de modo que

$$\alpha_1 - \beta_1 = \beta_2 - \alpha_2$$

e como $(\alpha_1 - \beta_1)$ está em W_1 e $(\beta_2 - \alpha_2)$ está em W_2, devemos ter $\alpha_1 - \beta_1 = \beta_2 - \alpha_2 = 0$, isto é, $\alpha_1 = \beta_1$ e $\alpha_2 = \beta_2$. Quando W_1 e W_2 forem disjuntos diremos que a soma $W = W_2 + W_1$ é **direta** ou que W é a **soma direta** de W_1 e W_2 e escreveremos $W = W_1 \oplus W_2$. A importância das somas diretas está no fato de que se $W = W_1 \oplus W_2$, podemos estudar W através dos pares de vetores (α_2, α_1) com α_i em W_i.

Desejamos considerar "somas diretas" de vários subespaços. Para fazer isto precisaremos de um conceito de independência de subespaços, análogo à condição de disjunção no caso de dois subespaços.

Definição. *Sejam* $W_1, \ldots, W_k$ *subespaços do espaço vetorial* V. *Diremos que* $W_1, \ldots, W_k$ *são* **independentes** *se*

$$\alpha_1 + \ldots + \alpha_k = 0, \ \alpha_i \text{ em } W_i$$

implica que cada α_i *é nulo.*

Teorema 1. *Seja* V *um espaço vetorial sôbre o corpo* F. *Sejam* $W_1, \ldots, W_k$ *subespaços de* V *e seja* $W = W_1 + \ldots + W_k$. *As seguintes condições são equivalentes.*

(i) $W_1, \ldots, W_k$ *são independentes.*

(ii) *Cada vetor* α *em* W *pode ser expresso de uma única maneira sob a forma*

$$\alpha = \alpha_1 + \ldots + \alpha_k$$

com α_i *em* W_i, $i = 1, \ldots, k$.

(iii) *Para cada* j, $2 \leq j \leq k$, *o subespaço* W_j *é disjunto da soma* $(W_1 + \ldots + W_{j-1})$.

Demonstração. (i) → (ii). Suponhamos que $W_1, \ldots, W_k$ sejam independentes. Pela definição de W, cada vetor α ou W pode ser expresso sob a forma $\alpha = \alpha_1 + \ldots + \alpha_k$ com α_i em W_i. Suponhamos também que $\alpha = \beta_1 + \ldots + \beta_k$ com β_i em W_i. Então

$$\alpha_1 + \ldots + \alpha_k = \beta_1 + \ldots + \beta_k$$

portanto $$(\alpha_1 - \beta_1) + \ldots + (\alpha_k - \beta_k) = 0$$

e como $(\alpha_i - \beta_i)$ está em W_i, a independência dos W_i implica que $\alpha_i - \beta_i = 0$, $i = 1, \ldots, k$. Isto mostra que os α_i são determinados de modo único por α.

(ii) → (iii). Seja α um vetor na interseção

$$W_j \cap (W_1 + \ldots + W_{j-1}).$$

Então existem vetores $\alpha_1, \ldots, \alpha_{j-1}$ com α_i em W_i tais que $\alpha = \alpha_{1i} + \ldots + \alpha_{j-1}$. Mais como α está em W_j a única expressão para α como uma soma de vetores, um em cada W_i, deve ser

$$\alpha = 0 + \ldots + 0 + \alpha + 0 + \ldots + 0.$$

Conseqüentemente, devemos ter que $\alpha = \alpha_1 = \ldots = \alpha_{j-1} = 0$.

(iii) $\rightarrow$ (i). Seja α_i em W_i, $i = 1, \ldots, k$ e suponhamos que $\alpha_1 + \ldots + \alpha_k = 0$. Desejamos demonstrar que cada α_i é o vetor nulo. Suponhamos o contrário, isto é, que para algum i tenhamos $\alpha_i \neq 0$. Seja j o maior inteiro i entre 1 e k tal que $\alpha_i \neq 0$. Então temos

$$\alpha_1 + \ldots + \alpha_j = 0, \quad \alpha_j \neq 0. \tag{6-1}$$

Agora (6-1) diz que $\alpha_j = -\alpha_1 - \ldots - \alpha_{j-1}$, de modo que α_j está na interseção de W_j com $(W_1 + \ldots + W_{j-1})$. Logo $\alpha_j = 0$. Contradição.

Se uma (e portanto tôdas) das três condições do Teorema 1 valer para $W_1, \ldots, W_k$, diremos que a soma $W = W_1 + \ldots + W_k$ é **direta** ou que W é a **soma direta** de $W_1, \ldots, W_k$ e escreveremos $W = W_1 \oplus \ldots \oplus W_k$.

Teorema 2. *Seja* V *um espaço vetorial de dimensão finita sôbre o corpo* F *e sejam* $W_1, \ldots, W_k$ *subespaços de* V. *As duas afirmações seguintes são equivalentes:*

(i) $V = W_1 \oplus \ldots \oplus W_k$.

(ii) *Se* $\mathcal{B}_i$ *é uma base de* W_i, $i = 1, \ldots, k$, *então a reunião* $\mathcal{B} = \bigcup_{i=1}^{k}$ *é uma base de* V.

Demonstração. (i) $\rightarrow$ (ii). Seja $\mathcal{B}_i = \{\alpha_1^i, \ldots, \alpha_{d_i}^i\}$ uma base de W_i ($d_i = \dim W_i$) e seja $\mathcal{B}$ a reunião das bases $\mathcal{B}_i$. Uma combinação linear típica dos vetores em $\mathcal{B}$ é

$$\sum_{j=1}^{d_1} c_j^1\alpha_j^1 + \ldots + \sum_{j=1}^{d_k} c_j^k\alpha_j^k \tag{6-2}$$

onde os c_j^i, $1 \leq i \leq k$, $1 \leq j \leq d_i$ são $d_1 \ldots d_k$ escalares em F. Portanto uma combinação linear típica dos vetores em $\mathcal{B}$ é

$$\beta_1 + \ldots + \beta_k \tag{6-3}$$

onde
$$\beta_i = \sum_{j=1}^{d_i} c_j^i\alpha_j^i.$$

Como $V = W_1 + \ldots + W_k$ e $\mathcal{B}_i$ é uma base de W_i é claro que todo vetor em V pode ser expresso sob a forma (6-3), isto é, que $\mathcal{B}$ gera

V. Se em (6-3) $\beta_1 + \ldots + \beta_k = 0$, a independência dos espaços W_i nos diz que $\beta_i = 0$ para cada i; mas então, como $\mathcal{B}_i$ é um conjunto independente, temos $c_j^i = 0$ para todos i e j.

Deixamos a cargo do leitor demonstrar que (ii) → (i). É uma inversão bastante simples do raciocínio usado acima.

Uma conseqüência particular do Teorema 2 é que se

$$V = W_1 \oplus \ldots \oplus W_k,$$

então

$$\dim V = \dim W_1 + \ldots + \dim W_k.$$

Uma espécie de recíproca dêste fato é verdadeira, a saber, se a soma das dimensões dos W_i é igual à dimensão de V e se V é a soma dos W_i, então V é a soma direta dos W_i. Deixamos isto para a parte de exercícios.

Exemplo 1. Seja V um espaço vetorial de dimensão finita sôbre o corpo F e seja $\{\alpha_1, \ldots, \alpha_n\}$ uma base arbitrária de V. Se W_i é o subespaço unidimensional gerado por α_i, então $V = W_i \oplus \ldots \ldots \oplus W_n$.

Exemplo 2. Seja n um inteiro positivo, F um subcorpo do corpo dos números complexos e seja V o espaço das $n \times n$ matrizes sôbre F. Seja W_1 o subespaço das matrizes **simétricas**, isto é, matrizes A tais $A^t = A$. Seja W_2 o subespaço das matrizes **anti-simétricas**, isto é, matrizes tais que $A^t = A$. Então $V = W_1 \oplus W_2$. Se A é uma matriz arbitrária em V, a única expressão para A como uma soma de duas matrizes, uma em W_1 e a outra em W_2, é

$$\begin{aligned} A &= A_1 + A_2 \\ A_1 &= \tfrac{1}{2}(A + A^t) \\ A_2 &= \tfrac{1}{2}(A - A^t). \end{aligned}$$

Definição. *Seja* V *um espaço vetorial e* T *um operador linear sôbre* V. *Se* W *é um subespaço de* V, *dizemos que* W *é* **invariante sob** T *se para cada vetor* α *em* W *o vetor* Tα *está em* W, *isto é, se* T(W) *está contido em* W.

Exemplo 3. Se T é um operador linear arbitrário sôbre V, então V é invariante sob T, da mesma forma que o é o subespaço nulo. A imagem de T e o núcleo de T também são invariantes sob T.

Exemplo 4. Seja F um corpo e seja D o operador derivação sôbre o espaço $F[x]$ dos polinômios sôbre F. Seja n um inteiro positivo e seja W o subespaço dos polinômios de grau menor ou igual a

n. Então W é invariante sob D. Esta é apenas outra maneira de dizer que D diminui o grau.

Exemplo 5. Seja T o operador linear sôbre R^2 que é representado em relação à base ordenada canônica pela matriz

$$\begin{bmatrix} 0 & -0 \\ 1 & 0 \end{bmatrix}.$$

Então os únicos subespaços de R^2 que são invariantes sob T são R^2 e o subespaço nulo. De fato, qualquer outro subespaço invariante W teria necessàriamente dimensão 1. Mas se W é o subespaço gerado por algum vetor não-nulo α, a afirmação de que W é invariante sob T significa que $T\alpha = c\alpha$ para algum número real c. Mas isto é impossível com $\alpha \neq 0$, pois pode-se verificar fàcilmente que para qualquer c o operador $(T - cI)$ é inversível, isto é, para todo número real c a matriz

$$\begin{bmatrix} -c & -1 \\ 1 & -c \end{bmatrix}$$

é inversível.

Quando o subespaço W é invariante sob o operador T, então T induz um operador linear T_W sôbre o espaço W se restringimos o seu domínio de definição a W. O operador linear T_W é definido por $T_W(\alpha) = T(\alpha)$, para α em W, mas T_W é um objeto bem diferente de T uma vez que seu domínio é W e não V.

Quando V é de dimensão finita, a invariância de W sob T admite uma interpretação simples por meio de matrizes e é interessante mencioná-la neste ponto. Suponhamos que tomemos uma base ordenada $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ de V tal que $\mathcal{B}' = \{\alpha_1, \ldots, \alpha_r\}$ seja uma base ordenada de $W(r = \dim W)$. Seja $A = [T]_{\mathcal{B}}$. Então

$$T\alpha_j = \sum_{i=1}^{n} A_{ij}\alpha_i.$$

Como W é invariante sob T, o vetor $T\alpha_j$ pertence a W para $j \leq r$. Isto significa que

$$T\alpha_j = \sum_{i=1}^{r} A_{ij}\alpha_i,\ j \leq r. \tag{6-4}$$

Em outras palavras, $A_{ij} = 0$ se $j \leq r$ e $i \leq r$.

Esquemàticamente, A é da forma em blocos

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$$

onde B é uma $r \times r$ matriz, C é uma $r \times (n - r)$ matriz e D é uma $(n - r) \times (n - r)$ matriz. O leitor deverá notar que, de acôrdo com (6—4), a matriz B é exatamente a matriz do operador induzido T_W, em relação à base ordenada $\mathcal{B}'$.

Ao estudarmos T estaremos muito interessados em decomposições em somas diretas $V = W_1 \oplus \ldots \oplus W_k$, onde cada um dos subespaços W_i seja invariante sob T. Dada uma tal decomposição de V, T induz um operador linear T_i, sôbre cada W_i, através da restrição. A ação de T é então a que segue. Se α é um vetor em V, existem vetores bem determinados $\alpha_1, \ldots, \alpha_k$, com α_i em W_i, tais que

$$\alpha = \alpha_1 + \ldots + \alpha_k$$

e então

$$T\alpha = T_1\alpha_1 + \ldots + T_k\alpha_k. \tag{6-5}$$

Descreveremos esta situação dizendo que T é a **soma direta dos** operadores $T_1, \ldots, T_k$. Deve-se lembrar, ao usarmos esta terminologia, que os T_i não são operadores lineares sôbre o espaço V mas sim sôbre os diversos subespaços W_i. O fato de que $V = W_1 \oplus \oplus \ldots \oplus W_k$ nos permite associar a cada α em V uma única k-dupla $(\alpha_1, \ldots, \alpha_k)$ de vetores α_i em W_i (sendo $\alpha = \alpha_1 + \ldots + \alpha_k$) de uma maneira tal possamos efetuar as operações lineares em V trabalhando em cada subespaço W_i. O fato de que cada W_i é invariante sob T nos permite considerar a ação de T como a ação independente dos operadores T_i sôbre os subespaços W_i. Nosso propósito é estudar T determinando decomposições em somas diretas invariantes nas quais os T_i sejam operadores de natureza elementar.

Antes de considerarmos um exemplo, observemos o análogo desta situação para matrizes. Suponhamos que tomemos uma base ordenada $\mathcal{B}_i$ para cada W_i e seja $\mathcal{B}$ a base ordenada de V formada pela reunião das $\mathcal{B}_i$, ordenada como $\mathcal{B}_1, \ldots, \mathcal{B}_k$. (Sabemos, pelo Teorema 2, que $\mathcal{B}$ é uma base de V.) De nossa discussão acima, quanto ao análogo para matrizes para um único subespaço invariante, é fácil ver que se $A = [T]_{\mathcal{B}}$ e $A_i = [T_i]_{\mathcal{B}_i}$, então A é da forma em blocos

$$A = \begin{bmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & A_k \end{bmatrix}. \tag{6-6}$$

Em (6-6), A_i é a $d_i \times d_i$ matriz ($d_i = \dim W_i$) e os símbolos 0 são blocos retangulares de vários tamanhos constituídos de escalares nulos. Parece ser também apropriado descrever (6-6) dizendo que A é a **soma direta** das matrizes $A_1, \ldots, A_k$.

Exemplo 6. Seja V um espaço vetorial de dimensão finita sôbre o corpo F e seja E um operador linear sôbre V. Suponhamos que E seja idempotente, isto é, que $E^2 = E$. Se R é a imagem de E e N é o núcleo de E, afirmamos que $V = R \oplus N$. De fato, seja α um vetor arbitrário em V. Escrevamos

$$\alpha = E\alpha + (\alpha - E\alpha). \tag{6-7}$$

Ora, $E\alpha$ está na imagem R de E e como

$$\begin{aligned} E(\alpha - E\alpha) &= E\alpha - E^2\alpha \\ &= E\alpha - E\alpha \\ &= 0 \end{aligned}$$

vemos que $(\alpha - E\alpha)$ está no núcleo N de E. Assim, $V = R + N$. Além disso, R e N são disjuntos. De fato, suponhamos que α esteja na interseção $R \cap N$. Então $\alpha = E\beta$ para algum β em V, logo $E\alpha = E(E\beta) = E^2\beta = E\beta = \alpha$. Como α está em N, $E\alpha = 0$, logo $\alpha = 0$. Portanto $V = R \oplus N$. Deveríamos notar que nossos cálculos acima mostram que a imagem R consiste exatamente nos vetores α em V tais que $E\alpha = \alpha$. Para todo operador linear T, a imagem de T e o núcleo de T são invariantes sob T e isto, em particular, é verdadeiro para E. O que é o operador E_R que E induz sôbre R? Como salientamos há pouco, E funciona como o operador idêntico sôbre R, isto é, E_R é o operador idêntico sôbre o espaço R. O operador induzido E_N é evidentemente o operador nulo sôbre N. Suponhamos que tomemos uma base ordenada de R e uma base ordenada de N e reunamo-las, formando uma base ordenada $\mathcal{B}$ de V. Então a matriz de E em relação a $\mathcal{B}$ será

$$\begin{bmatrix} 1 & 0 & \ldots & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & \ldots & 0 \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ 0 & 0 & \ldots & 1 & \ldots & 0 \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ 0 & 0 & \ldots & 0 & \ldots & 0 \end{bmatrix}$$

ou, na forma em blocos,

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$$

onde I é a $r \times r$ matriz unidade e r = pôsto(E).

Um operador idempotente é freqüentemente denominado uma **projeção**. Na verdade se R é a imagem de E e N é o núcleo de E dizemos que E é a **projeção sôbre** R **segundo** N. Um exame rápido do operador idempotente E sôbre F^2 representado pela matriz

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

em relação à base ordenada canônica, deverá explicar a terminologia. Como logo veremos, se R e N são subespaços de V tais que $V = R \oplus N$, então existe uma única projeção sôbre R segundo N.

Projeções podem ser usadas para descrever decomposições do espaço V em somas diretas. De fato, suponhamos $V = W_1 \oplus \ldots \oplus W_k$. Para cada j definiremos um operador E_j sôbre V. Seja α em V, digamos $\alpha = \alpha_1 + \ldots + \alpha_k$ com α_i em W_i. Definamos $E_j\alpha = \alpha_j$. Então, E_j é uma regra bem definida. É fácil ver que E_j é linear, que a imagem E_j é W_j e que $E_j^2 = E_j$. O núcleo de E_j é o subespaço

$$(W_1 + \ldots + W_{j-1} + W_{j+1} + \ldots + W_k)$$

pois a afirmação de que $E_j\alpha = 0$ significa simplesmente que $\alpha_j = 0$, isto é, que α é na realidade uma soma de vetores dos espaços W_i com $i \neq j$. Em têrmos das projeções E_j temos

$$\alpha = E_1\alpha + \ldots + E_k\alpha \tag{6-8}$$

para cada α em V. O que (6-8) diz é que

$$I = E_1 + \ldots + E_k. \tag{6-9}$$

Notemos também que se $i \neq j$ então $E_iE_j = 0$, pois a imagem de E_j é o subespaço W_j que está contido no núcleo de E_i. Resumiremos agora nossas conclusões, enunciaremos e demonstraremos uma recíproca.

Teorema 3. *Se* $V = W_1 \oplus \ldots \oplus W_k$, *então existem* k *operadores* $E_1, \ldots, E_k$ *sôbre* V *tais que*

(a) *cada* E_i *é uma projeção* ($E_i^2 = E_i$);

(b) $E_iE_j = 0$, *se* $i \neq j$;

(c) $I = E_1 + \ldots + E_k$;
(d) *a imagem de* E_i *é* W_i.

Recìprocamente, se $E_1, \ldots, E_k$ *são* k *operadores lineares sôbre* V *que satisfazem as condições* (a), (b) *e* (c) *e se indicamos por* W_i *a imagem de* E_i, *então* $V = W_1 \oplus \ldots \oplus W_k$.

Demonstração. Precisamos demonstrar apenas a afirmação recíproca. Suponhamos que $E_1, \ldots, E_k$ sejam operadores lineares sôbre V que satisfaçam as três primeiras condições e seja W_i a imagem de E_1. Então, certamente

$$V = W_1 + \ldots + W_k;$$

pois, pela condição (c) temos

$$\alpha = E_1\alpha + \ldots + E_k\alpha$$

para cada α em V e $E_i\alpha$ está em W_i. Esta expressão para α é única, porque se

$$\alpha = \alpha_1 + \ldots + \alpha_k$$

com α_i em W_i, digamos $\alpha_i = E_i\beta_i$, então, usando (a) e (b), temos

$$\begin{aligned} E_j\alpha &= \sum_{i=1}^{k} E_j\alpha_i \\ &= \sum_{i=1}^{k} E_jE_i\beta_i \\ &= E_j^2\beta_j \\ &= E_j\beta_j \\ &= \alpha_j. \end{aligned}$$

Isto mostra que V é a soma direta dos W_i.

Suponhamos que T seja um operador linear sôbre V e que W_i e E_i sejam como acima. Como se enuncia em têrmos dos E_i a condição de que cada W_i seja invariante sob T?

Teorema 4. *Seja* T *um operador linear sôbre o espaço* V *e sejam* $W_1, \ldots, W_k$ *e* $E_1, \ldots, E_k$ *como no Teorema* 3. *Então, uma condição necessária e suficiente para que cada subespaço* W_i *seja invariante sob* T *é que* T *comute com cada uma das projeções* E_i, *isto é*,

$$TE_i = E_iT, \ i = 1, \ldots, k.$$

Demonstração. Suponhamos que T comute com cada E_i. Seja α em W_j. Então $E_j\alpha = \alpha$ e

$$\begin{aligned} T\alpha &= T(E_j\alpha) \\ &= E_j(T\alpha) \end{aligned}$$

o que mostra que $T\alpha$ está na imagem de E_j, isto é, que W_j é invariante sob T.

Suponhamos agora que cada W_j seja invariante sob T. Mostraremos que $TE_j = E_jT$. Seja α um vetor arbitrário em V. Então

$$\begin{aligned} \alpha &= E_1\alpha + \dots + E_k\alpha \\ T\alpha &= TE_1\alpha + \dots + TE_k\alpha. \end{aligned}$$

Como $E_i\alpha$ está em W_i, que é invariante sob T, devemos ter $T(E_i\alpha) = E_i\beta_i$ para algum vetor β_i. Então

$$\begin{aligned} E_jTE_i\alpha &= E_jE_i\beta_i \\ &= \begin{cases} 0, & \text{se } i \neq j \\ E_j\beta_j, & \text{se } i = j. \end{cases} \end{aligned}$$

Assim
$$\begin{aligned} E_jT\alpha &= E_jTE_1\alpha + \dots + E_jTE_k\alpha \\ &= E_j\beta_j \\ &= TE_j\alpha. \end{aligned}$$

Isto vale para todo α em V, portanto $E_jT = TE_j$.

Exercícios

1. Seja V um espaço vetorial de dimensão finita e seja W_1 um subespaço arbitrário de V. Demonstrar que existe um subespaço W_2 de V tal que $V = W_1 \oplus W_2$.

2. Seja V um espaço vetorial de dimensão finita e sejam $W_1, \dots, W_k$ subespaços de V tais que

$$V = W_1 + \dots + W_k \text{ e } \dim V = \dim W_1 + \dots + \dim W_k.$$

Demonstrar que $V = W_1 \oplus \dots \oplus W_k$.

3. Seja T o operador linear sôbre R^2, cuja matriz em relação à base ordenada canônica é

$$A = \begin{bmatrix} 1 & -1 \\ 2 & 2 \end{bmatrix}.$$

(a) Demonstrar que os únicos subespaços de R^2, que são invariantes sob T, são R^2 e o subespaço nulo.

(b) Se U é o operador linear sôbre C^2, cuja matriz em relação à base ordenada canônica é A, mostrar que U possui subespaços invariantes unidimensionais.

4. Seja V um espaço vetorial sôbre o corpo F, seja T um operador linear sôbre V e seja f um polinômio sôbre o corpo F. Se W é o núcleo do operador $f(T)$, demonstrar que W é invariante sob T.

5. Seja E uma projeção sôbre V e seja T um operador linear sôbre V. Demonstrar que a imagem de E é invariante sob T se, e sòmente se, $ETE = TE$. Demonstrar que a imagem e o núcleo de E são invariantes sob T se, e sòmente se, $ET = TE$.

6. Determinar uma projeção E sôbre R^2 tal que E projeta R^2 sôbre o subespaço gerado por $(1, —1)$ segundo o subespaço gerado por $(1, 2)$.

7. Seja T o operador linear sôbre R^2, cuja matriz em relação à base ordenada canônica é

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$$

Seja W_1 o subespaço de R^2 gerado pelo vetor ϵ_1.

(a) Demonstrar que W_1 é invariante sob T.

(b) Demonstrar que não existe nenhum subespaço W_2 de R^2 que seja invariante sob T e que também tenha a propriedade de que $R^2 = W_1 \oplus W_2$. (Comparar com o Problema 1.)

8. Demonstrar que se E é a projeção sôbre R segundo N, então $(I — E)$ é a projeção sôbre N segundo R.

9. Sejam $E_1, \ldots, E_k$ operadores lineares sôbre o espaço V tais que $E_1 + \ldots + E_k = I$.

(a) Demonstrar que se $E_iE_j = 0$ para $i \neq j$, então $E_i^2 = E_i$ para todo i.

(b) No caso $k = 2$, demonstrar a recíproca de (a). Isto é, se $E_1 + E_2 = I$ e $E_1^2 = E_1$, $E_2^2 = E_2$, então $E_1E_2 = 0$.

10. Seja V um espaço vetorial real e E um operador linear idempotente sôbre V, isto é, uma projeção. Demonstrar que $(I + E)$ é inversível. Determinar $(I + E)^{-1}$.

11. Seja V um espaço vetorial, sejam $W_1, \ldots, W_k$ subespaços de V e seja

$$V_j = W_1 + \ldots + W_{j-1} + W_{j+1} + \ldots + Wk.$$

Suponhamos que $V = W_1 \oplus \ldots \oplus W_k$. Demonstrar que o espaço dual V^* admite a decomposição em soma direta $C^* = V_1^0 \oplus \ldots \oplus V_k^0$.

12. Seja V um espaço vetorial de dimensão finita e seja T um operador linear sôbre V. Suponhamos que $V = W_1 \oplus \ldots \oplus W_k$, onde cada W_i é invariante sob T. Se T_i é o operador induzido sôbre W_i por T, demonstrar que

$$\det(T) = \det(T_1) \ldots \det(T_k).$$

(*Sugestão*: Usar o fato correspondente sôbre somas diretas de matrizes.)

13. Seja T um operador linear sôbre o espaço vetorial V de dimensão finita, seja R a imagem de T e seja N o núcleo de T. Demonstrar que R e N são disjuntos se, e sòmente se, $V = R \oplus N$.

14. Seja F um subcorpo do corpo dos números complexos e seja V um espaço vetorial de dimensão finita sôbre F. Suponhamos que $E^1, \ldots, E_k$ sejam projeções sôbre V e que $E_1 + \ldots. + E_k = I$. Demonstrar que $E_iE_j = 0$ para $i \neq j$. (*Sugestão*: Usar a função traço, descobrir o que é o traço de uma projeção e usar o resultado do Exercício 2.)

15. Seja V um espaço vetorial de dimensão finita.

(a) Se T é um operador linear inversível sôbre V e W_1 e W_2 são subespaços de V tais que $V = W_1 \oplus W_2$ demonstrar que

$$V = T(V) = T(W_1) \oplus T(W_2).$$

(b) Seja T um operador linear sôbre V que tenha a propriedade de que se $V = W^1 \oplus W_2$ então $T(V) = T(W_1) \oplus T(W_2)$. Demonstrar que T é inversível.

6.2 Valores Característicos e Vetores Característicos

Em tôda esta seção, T será um operador linear sôbre um espaço vetorial n-dimensional V sôbre o corpo F.

Definição. *Um* **valor característico** *de T é um escalar* c *em* F *tal que exista um vetor não-nulo* α *em* V *com* $T\alpha = c\alpha$. *Se* c *é um valor característico de* T, *então todo* α *tal que* $T\alpha = c\alpha$ *é denominado um* **vetor característico de** T **associado ao valor característico** c.

Valores característicos são freqüentemente denominados raízes características, raízes latentes, autovalores, valores próprios ou valores espectrais. Neste livro usaremos apenas o nome valores característicos. Ora, o fato de que $T\alpha = c\alpha$, com $\alpha \neq 0$, diz que o subespaço unidimensional de V que é gerado pelo vetor α é invariante sob T. Assim, valores (vetores) característicos surgem de subespaços invariantes não-nulos e de dimensão mínima. Exploraremos êsse "aspecto de invariância" das coisas na Seção 6.3.

Teorema 5. *As seguintes condições são equivalentes:*

(i) c *é um valor característico de* T.

(ii) *O operador* (T — cI) *é singular (não-inversível).*

(iii) *det*(T — cI) = 0.

Demonstração. A afirmação de que existe um vetor não-nulo α tal que $T\alpha = c\alpha$ é simplesmente a afirmação de que existe um vetor não-nulo α tal que $(T - cI)\alpha = 0$. A partir disto é evidente que (i), (ii) e (iii) são afirmações equivalentes.

Se $\mathcal{B}$ é uma base ordenada arbitrária de V e $A = [T]_{\mathcal{B}}$, então $(T - cI)$ é inversível se e sòmente se a matriz $(A - cI)$ é inversível. Portanto, colocamos a definição que segue.

Definição. *Se* A *é uma* n × n *matriz sôbre o corpo* F, *um* **valor característico** *de* A *é um escalar* c *em* F *tal que a matriz* (A — cI) *seja singular (não-inversível).*

Como c é um valor característico de A se, e sòmente se, det $(A - cI) = 0$, ou equivalentemente, se, e sòmente se, $\det(cI - A) = 0$, podemos formar a matriz $(xI - A)$ cujos elementos são polinômios e considerar o polinômio $f = \det(xI - A)$. Evidentemente os valores característicos de A são exatamente os escalares c em F tais que $f(c) = 0$. Por esta razão f é denominado o **polinômio característico** de A. É importante notar que f é um polinômio unitário cujo grau é exatamente n. Isto pode ser visto fàcilmente através da fórmula para o determinante de uma matriz em função de seus elementos.

Lema. *Matrizes semelhantes têm o mesmo polinômio característico.*

Demonstração. Se $B = P^{-1}AP$, então

$$\begin{aligned}\det(xI - B) &= \det(xI - P^{-1}AP)\\ &= \det(P^{-1}(xI - A)P)\\ &= \det P^{-1} \,.\, \det(xI - A) \,.\, \det P\\ &= \det(xI - A).\end{aligned}$$

Êste lema nos permite definir sensatamente o polinômio característico do operador T como o polinômio característico de qualquer $n \times n$ matriz que represente T em relação a alguma base ordenada para V. Da mesma forma que para matrizes, os valores característicos de T serão as raízes do polinômio característico de T. Em particular, isto nos mostra que T não pode ter mais que n valores característicos distintos. É importante ressaltar que T pode não ter nenhum valor característico.

Exemplo 7. Seja T o operador linear sôbre R^2 que é representado em relação à base ordenada canônica pela matriz

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

O polinômio característico de T (ou de A) é

$$\det(xI - A) = \begin{vmatrix} x & 1 \\ -1 & x \end{vmatrix} = x^2 + 1.$$

Como êste polinômio não possui raízes reais, T não possui valores característicos. Isto é realmente uma repetição de um fato que salientamos anteriormente, a saber, que T não possui subespaços invariantes unidimensionais. Se U é o operador linear sôbre C^2 que é representado por A em relação à base ordenada canônica, então U possui dois valores característicos, i e $-i$. Vemos aqui um ponto sutil. Ao discutirmos os valores característicos de uma matriz A, precisamos tomar o cuidado de estipular o corpo envolvido. A matriz A acima não possui nenhum valor característico quando considerada como uma matriz real, mas possui dois valores característicos, i e $-i$, quando considerada como uma 2×2 matriz complexa.

Exemplo 8. Seja A a 3×3 matriz

$$\begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 4 \\ 3 & -6 & -4 \end{bmatrix}.$$

Então o polinômio característico de A é

$$\begin{vmatrix} x-5 & 6 & 6 \\ 1 & x-4 & -2 \\ -3 & 6 & x+4 \end{vmatrix} = x^3 - 5x^2 + 8x - 4 = (x-1)(x-2)^2.$$

Uma propriedade muito interessante e útil do polinômio característico é dada pelo teorema Cayley-Hamilton, que agora demonstraremos.

Teorema 6. (Cayley-Hamilton). *Seja* T *um operador linear sôbre o espaço vetorial* n*-dimensional* V [*ou, seja* A *uma* n $\times$ n *matriz sôbre o corpo* F] *e seja* f *o polinômio característico de* T [*de* A]. *Então* $f(T) = 0$ [$f(A) = 0$].

Demonstração. Seja T um operador linear sôbre o espaço n-dimensional V. Se g é um polinômio sôbre o corpo F e α é um vetor em V, usaremos a notação $g\alpha$ para o vetor $g(T)\alpha$. Se h é um outro polinômio sôbre F, então

$$(g + h)\alpha = g\alpha + h\alpha \quad \text{e} \quad g(h\alpha) = (gh)\alpha.$$

Estas afirmações são simplesmente

$$g(T) + h(T) = (g + h)(T) \quad \text{e} \quad g(T)h(T) = (gh)(T),$$

que demonstraremos no Capítulo 4.

Tomemos uma base ordenada $\{\alpha_1, \ldots, \alpha_n\}$ de V e seja A a matriz que representa T em relação a esta base ordenada. Seja B a matriz $(xI - A)$. Então B é uma $n \times n$ matriz sôbre o anel de polinômios $F[x]$ e, por definição, $f = \det B$ é o polinômio característico de T (ou de A). Como A representa T em relação à base ordenada acima

$$T\alpha_j = \sum_{i=1}^{n} A_{ij}\alpha_i.$$

Com a notação que introduzimos no primeiro parágrafo da demonstração, isto pòde ser reescrito como

$$x\alpha_j = \sum_{i=j}^{n} A_{ij}\alpha_i$$

ou

$$\sum_{i=1}^{n} (\delta_{ij}x - A_{ij})\alpha_i = 0.$$

Como $\delta_{ij}x - A_{ij} = B_{ij}$ temos

$$\text{(6-10)} \qquad \sum_{i=1}^{n} B_{ij}\alpha_i = 0, \quad j = 1, \ldots, n.$$

Desejamos mostrar que $f(T) = 0$. Para que $f(T)$ seja o operador nulo, é necessário e suficiente que $f(T)\alpha_k = 0$ para $k = 1, \ldots, n$. Fixemos k, $1 \leq k \leq n$; mostraremos que $f(T)\alpha_k = 0$, isto é, $f\alpha_k = 0$. Seja $\tilde{B}$ a matriz adjunta clássica de B, $\tilde{B} = \operatorname{adj} B$. De (6-10) temos

$$\sum_{i=1}^{n} \tilde{B}_{jk}B_{ij}\alpha_i = 0, \quad j = 1, \ldots, n$$

e somando sôbre j

$$\begin{aligned} 0 &= \sum_{j=1}^{n} \sum_{i=1}^{n} \tilde{B}_{jk}B_{ij}\alpha_i \\ &= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} B_{ij}\tilde{B}_{jk} \right)\alpha_i. \end{aligned}$$

Ora, $B\tilde{B} = (\det B)I = fI$, e então

$$\sum_{j=1}^{n} B_{ij}\tilde{B}_{jk} = \delta_{ik}f.$$

Assim, temos

$$\begin{aligned} 0 &= \sum_{i=1}^{n} \delta_{ik}f\alpha_i \\ &= f\alpha_k. \end{aligned}$$

No Capítulo 7 faremos uma outra demonstração dêste teorema. Para compreender a demonstração que fizemos acima, insistimos veementemente para que o leitor examine de perto o argumento para o caso $n = 2$. As equações (6-10) serão então

$$\begin{aligned} (x - A_{11})\alpha_1 - A_{21}\alpha_2 &= 0 \\ -A_{12}\alpha_1 + (x - A_{22})\alpha_2 &= 0 \end{aligned}$$

e o polinômio característico

$$f = x^2 - (A_{11} + A_{22})x + (A_{11}A_{22} - A_{12}A_{21}).$$

Pode-se obter $f\alpha_1 = 0$ "eliminando" α_2 destas equações e pode-se obter $f\alpha_2 = 0$ "eliminando" α_1. No caso geral, êste processo de eliminação é o que se consegue através do uso da adjunta clássica de $(xI - A)$.

O teorema de Cayley-Hamilton nos diz algo que não é muito fácil de ver diretamente, a saber, que se T é um operador linear sôbre o espaço n-dimensional V, então existe um polinômio f de grau n tal que $f(T) = 0$. Notemos que é fácil ver que existe um polinômio não-nulo g de grau não maior que n^2 tal que $g(T) = 0$. De fato, o espaço $L(V, V)$ tem dimensão n^2, portanto os operadores $I, T, T^2, \ldots, T^{n^2}$ devem ser linearmente dependentes. O conjunto dos polinômios g em $F[x]$ tais que $g(T) = 0$ é um ideal, como é fácil de verificar. Além disso, êste ideal contém um polinômio unitário de grau n, que é o polinômio característico de T. Assim, existe um único polinômio unitário p que gera o ideal formado por todos g tais que $g(T) = 0$. Êste p é denominado o **polinômio minimal** de T. O nome vem (é claro) do fato de p ser o polinômio unitário de grau mínimo que leva T em 0. O teorema de Cayley-Hamilton pode ser reenunciado: O polinômio característico de T é divisível pelo polinômio minimal de T.

Definimos o polinômio minimal de uma $n \times n$ matriz da maneira correspondente, isto é, como o único gerador unitário do ideal de todos os polinômios g sôbre F tais que $g(A) = 0$. É fácil ver que matrizes semelhantes têm exatamente o mesmo ideal de polinômios "anuladores" g, porque

$$g(P^{-1}AP) = P^{-1}g(A)P$$

e, conseqüentemente, matrizes semelhantes têm o mesmo polinômio minimal. Se o operador é representado em relação a alguma base ordenada pela matriz A, então T e A também têm o mesmo polinômio minimal, pois para todo g o operador $g(T)$ é representado pela matriz $g(A)$.

Exemplo 9. Seja F um subcorpo do corpo dos números complexos e seja T o operador linear sôbre F^3 que é representado em relação à base ordenada canônica pela matriz A do Exemplo 8. Vimos que o polinômio característico de A (e portanto de T) é $f = (x - 1)(x - 2)^2$. Qual é o polinômio minimal de T? Bem, o teorema de Cayley-Hamilton nos diz que o polinômio minimal de T deve dividir f. Logo, é evidente que êste polinômio deve ser

f, $(x - 1)(x - 2)$, $(x - 2)^2$, ou $(x - 1)$. Logo veremos que os dois últimos polinômios não são, na verdade, candidatos porque o polinômio minimal deve ter todos os valores característicos como raízes; mas por ora não sabemos isto. Qual dêstes é o polinômio minimal? Precisamos tomar cada polinômio, ver se êle "anula" T e então tomar o de menor grau que o faz. É claro que $(A - I) \neq 0$ e cálculos breves mostram que $(A - 2I)^2 \neq 0$. Por outro lado

$$(A - I)(A - 2I) = \begin{bmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{bmatrix} \begin{bmatrix} 3 & -6 & -6 \\ -1 & 2 & 2 \\ 3 & -6 & -6 \end{bmatrix} = 0$$

logo, o polinômio minimal é $p = (x - 1)(x - 2)$.

No Exemplo 9, vimos que seria conveniente, para a determinação do polinômio minimal de um dado operador, se soubéssemos que todo valor característico tem de ser uma raiz do polinômio minimal. Demonstraremos isto agora, com o objetivo de nos auxiliar no cálculo de polinômios minimais. No Capítulo 7, veremos que todo divisor irredutível do polinômio característico também é um divisor do polinômio minimal. Portanto, suponhamos que c seja um valor característico de T, digamos $T\alpha = c\alpha$ com $\alpha \neq 0$. Se g é um polinômio arbitrário sôbre o corpo F, então $g(T)\alpha = g(c)\alpha$, por exemplo,

$$T^2\alpha = T(T\alpha) = T(c\alpha) = cT\alpha = c^2\alpha.$$

Assim, para todo polinômio g, o escalar $g(c)$ é um valor característico do operador $g(T)$. Em particular, se $g(T) = 0$, então $g(c)$ é um valor característico do operador nulo, logo é nulo. Apliquemos isto ao caso em que g é o polinômio minimal. Combinando isto com o Teorema de Cayley-Hamilton, vemos que os valores característicos de T são exatamente as raízes do polinômio minimal, ou as raízes do polinômio característico. Assim, êstes dois polinômios têm exatamente as mesmas raízes, apesar de que as raízes possam ocorrer com maior multiplicidade no polinômio característico.

Exercícios

1. Em cada um dos casos seguintes, seja T o operador linear sôbre R^2 que é representado pela matriz A em relação à base ordenada canônica de R^2 e seja U o operador linear sôbre C^2 representado por A em relação à base ordenada canônica. Determinar os polinômios característicos de T e de U, determinar os valores característicos de cada operador e para cada valor característico c determinar uma base do correspondente espaço de vetores característicos.

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad A = \begin{bmatrix} 2 & 3 \\ -1 & 1 \end{bmatrix} \quad A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

2. Seja A a 3×3 matriz do Exemplo 8 e seja T o operador linear sôbre R^3 representado por A em relação à base ordenada canônica. Determinar uma base de R^3 formada por vetores característicos de T.

3. Seja V um espaço vetorial n-dimensional sôbre F. Qual é o polinômio característico do operador idêntico sôbre V? Qual é o polinômio minimal de I? Quais são os polinômios característico e minimal do operador nulo?

4. Seja A a 3×3 matriz real

$$\begin{bmatrix} 1 & 3 & 3 \\ 3 & 1 & 3 \\ -3 & -3 & -5 \end{bmatrix}.$$

Determinar o polinômio característico e o polinômio minimal de A.

5. Seja A uma $n \times n$ matriz triangular sôbre o corpo F. Demonstrar que os valores característicos de A são os elementos diagonais de A, isto é, os escalares A_{ii}.

6. Seja V um espaço vetorial n-dimensional sôbre o corpo F e seja T um operador linear sôbre V. Suponhamos que exista um inteiro positivo k tal que $T^k = 0$. Demonstrar que $T^n = 0$.

7. Sejam a, b e c elementos de um corpo F e seja A a seguinte 3×3 matriz sôbre F:

$$\begin{bmatrix} 0 & 0 & c \\ 1 & 0 & b \\ 0 & 1 & a \end{bmatrix}$$

Demonstrar que o polinômio característico de A é $x^3 - ax^2 - bx - c$, e que êste polinômio é também o polinômio minimal.

8. Sejam A e $n \times n$ matrizes sôbre o corpo F. Demonstrar que se $(I - AB)$ é inversível, então $(I - BA)$ é inversível e

$$(I - BA)^{-1} = I + B(I - AB)^{-1}A.$$

9. Usar o resultado do Exercício 8 para demonstrar que se A e B são $n \times n$ matrizes sôbre o corpo F, então AB e BA tem exatamente os mesmos valores característicos.

10. Seja A uma $n \times n$ matriz sôbre o corpo F e suponhamos que A seja a soma direta

$$\begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & A_k \end{bmatrix}$$

das matrizes $A_1, \dots, A_k$. Demonstrar que o polinômio característico de A é o produto dos polinômios característicos das A_i.

11. Seja V o espaço vetorial de tôdas as funções f de R em R que são contínuas, isto é, o espaço das funções contínuas, definidas sôbre a reta real e tomando valores reais. Seja T o operador linear sôbre V definido por

$$(Tf)(x) = \int_0^x f(t)dt.$$

Demonstrar que T não possui nenhum valor característico.

6.3 Operadores Diagonalizáveis

Definição. *Seja* T *um operador linear sôbre o espaço* V *de dimensão finita. Dizemos que* T *é* **diagonalizável** *se existe uma base de* V *formada por vetores característicos de* T.

A razão para o nome deveria ser evidente; de fato, se existe uma base ordenada $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ de V na qual cada α_i é um vetor característico de T, então a matriz T em relação à base ordenada $\mathcal{B}$ é diagonal. Se $\alpha_i = c_i\alpha_i$, então

$$[T_{\mathcal{B}}; = \begin{bmatrix} c_1 & 0 & \ldots & 0 \\ 0 & c^2 & \ldots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \ldots & c_n \end{bmatrix}.$$

Certamente não pretendemos que escalares $c_1, \ldots, c_n$ sejam distintos; na verdade, êles podem ser todos iguais (quando T é um múltiplo escalar do operador idêntico).

Poder-se-ia também definir T como sendo diagonalizável quando os vetores característicos de T gerassem V. Isto difere apenas superficialmente de nossa definição, uma vez que podemos sempre obter uma base a partir de qualquer conjunto de vetores que seja gerador.

Suponhamos que T seja diagonalizável e sejam $c_1, \ldots, c_k$ os valores característicos *distintos* de T. Para cada i, seja W_i o subespaço dos vetores característicos associados a c_i, isto é, o núcleo de $(T - c_iI)$. A própria definição "diagonalizável" diz que, juntos, os subespaços $W_1, \ldots, W_k$ geram V, isto é, que $V = W_1 + \ldots + W_k$. Esta soma é direta; na realidade, a independência dos subespaços $W_1, \ldots, W_k$ não depende do fàto de T ser diagonalizável.

Teorema 7. *Seja* T *um operador linear arbitrário sôbre o espaço* V *de dimensão finita, sejam* $c_1, \ldots, c_k$ *os valores característicos distintos de* T *e seja* W_i *o núcleo de* $(T - cI_i)$. *Então os subespaçoss* $W_1, \ldots, W_k$ *são independentes.*

Demonstração. Seja α_i em W_i, $i = 1, \ldots, k$ e suponhamos que

$$\alpha_1 + \ldots + \alpha_k = 0. \tag{6-11}$$

Seja j um inteiro entre 1 e k e seja U_j o produto dos operadores $(T - c_iI)$ para $i \neq j$. Pela definição de W_i, temos $(T - c_iI)\alpha_i = 0$ e, como todos os operadores $(T - c_iI)$ comutam, por serem polinômios em T, é claro que $U_j\alpha_i = 0$ para $i \neq j$. Além disso

$$U_j\alpha_j = [\prod_{j \neq i} (c_j - c_i)]\alpha_j. \tag{6-12}$$

Se aplicamos U_j a ambos os membros de (6—11) obtemos $U_j\alpha_j = 0$. Como os c_i são distintos, o produto

$$\prod_{i \neq j} (c_i - c_j)$$

é um escalar não-nulo. Então (6-12) nos mostra que $U_j\alpha_j = 0$ implica $\alpha_j = 0$. Mostramos que cada $\alpha_i = 0$, portanto os subespaços $W_1, \ldots, W_k$ são independentes.

Voltando ao nosso T diagonalizável, temos agora

$$V = W_1 \oplus \ldots \oplus W_k.$$

Certamente cada W_i é invariante sob T, pois T significa simplesmente a multiplicação de W_i pelo escalar c_i. Temos aqui então um exemplo extremamente simples de uma decomposição em soma direta invariante, sendo que os operadores induzidos sôbre os subespaços W_i são múltiplos escalares do operador idêntico.

Examinemos as projeções associadas a esta decomposição em soma direta, como no Teorema 3. Neste caso, a imagem da projeção E_i é o espaço W_i dos vetores característicos associados a c_i. Se α é um vetor arbitrário em V, então

$$\alpha = E_1\alpha + \ldots + E_k\alpha$$

logo

$$\begin{aligned} T\alpha &= TE_1\alpha + \ldots + TE_k\alpha \\ &= c_1E_1\alpha + \ldots + c_kE_k\alpha. \end{aligned}$$

Isto diz que

$$T = c_1E_1 + \ldots + c_kE_k.$$

Teorema 8. *Se* T *é um operador diagonalizável sôbre o espaço* V *de dimensão finita e* $c_1, \ldots, c_k$ *são os valores característicos distintos de* T, *então existem operadores lineares* $E_1, \ldots, E_k$ *sôbre* V *tais que*

(a) $T = c_1E_1 + \ldots + c_kE_k$;

(b) $I = E_1 + \ldots + E_k$;

(c) $E_iE_j = 0, \; i \neq j$;

(d) $E_i^2 = E_i$;

(e) *a imagem de* E_i *é o espaço dos vetores característicos de* T *associados ao valor característico* c_i.

Esta é a versão para operadores da afirmação de que T é representado por uma matriz diagonal em relação a alguma base ordenada. O leitor poderá sentir que parece bem complicada em comparação com a formulação para matrizes ou em relação à simples afirmação de que os vetores característicos de T geram o espaço V. Temos diversas razões para apresentá-la. Primeiro, o enunciado de tais teoremas puramente em têrmos de operadores é útil quando se deseja estudar posteriormente espaços de dimensão infinita. Segundo, êle nos ajudará a compreender mais tarde alguns teoremas de decomposição mais profundos. Terceiro, existe uma recíproca dêste teorema que agora enunciaremos e que nos ajudará a dar uma caracterização dos operadores diagonalizáveis.

Teorema 9. *Seja* T *um operador linear sôbre o espaço* V *de dimensão finita. Suponhamos que existam* k *escalares* $c_1, \ldots, c_k$, *distintos e* k *operadores lineares não-nulos* $E_1, \ldots, E_k$ *sôbre* V *que satisfaçam as condições* (a), (b) *e* (c) *do Teorema* 8. *Então* T *é diagonalizável,* $c_1, \ldots, c_k$ *são exatamente os valores característicos distintos de* T *e os* E_i *tambem satisfazem as condições* (d) *e* (e) *do Teorema* 8.

Demonstração. Como $E_iE_j = 0$ para $i \neq j$, multiplicando ambos os membros de $I = E_1 + \ldots + E_k$ por E_i obtemos imediatamente $E_i^2 = E_i$. Multiplicando $T = c_1E_1 + \ldots + c_kE_k$ por E_i resulta $TE_i = c_iE_i$, o que mostra que todo vetor na imagem de E_i está no núcleo de $(T - c_iI)$. Como supusemos que $E_i \neq 0$, isto demonstra que existe um vetor não-nulo no núcleo de $(T - c_iI)$, isto é, que c_i é um valor característico de T. Além disso, os c_i são os únicos valores característicos de T; de fato, se c é um escalar arbitrário, então

$$T - cI = (c_1 - c)E_1 + \ldots + (c_k - c)E_k$$

portanto se $(T - cI)\alpha = 0$, devemos ter $(c_i - c)E_i\alpha = 0$. Se α não é o vetor nulo então $E_i\alpha \neq 0$ para algum i, de modo que para êste i temos $c_i - c = 0$.

T é certamente diagonalizável, pois mostramos que todo vetor não-nulo na imagem de E_i é um vetor característico de T e o fato de que $I = E_1 + \ldots + E_k$ mostra que êsses vetores característicos geram V. Tudo o que resta a ser demonstrado é que o núcleo de $(T - c_iI)$ é exatamente a imagem de E_i. Mas isto é evidente, porque se $T\alpha = c_i\alpha$, então

$$\sum_{j=1}^{k} (c_j - c_i)E_j\alpha = 0$$

logo

$$(c_j - c_i)E_j\alpha = 0 \text{ para cada } j$$

e então

$$E_j\alpha = 0, \quad j \neq i.$$

Como $\alpha = E_1\alpha + \ldots + E_k\alpha$, e $E_j\alpha = 0$ para $j \neq i$ temos $\alpha = E_i\alpha$, demonstrando que α está na imagem de E_i.

Uma parte do Teorema 9 diz que para um operador diagonalizável T, os escalares $c_1, \ldots, c_k$ e os operadores $E_1, \ldots, E_k$ são determinados de modo único pelas condições (a), (b), (c), mais o fato de que os c_i são distintos e o fato de que os E_i são não-nulos. Uma das características agradáveis da decomposição $T = c_1E_1 + + \ldots + c_kE_k$ é que g é um polinômio arbitrário sôbre o corpo F, então

$$g(T) = g(c_1)E_1 + \ldots + g(c_k)E_k.$$

Deixemos os detalhes da demonstração a cargo do leitor. Para ver como se demonstra, basta calcular T^r para todo inteiro positivo r. Por exemplo

$$\begin{aligned} T^2 &= \sum_{i=1}^{k} c_iE_i \sum_{j=1}^{k} c_jE_j \\ &= \sum_{i=1}^{k} \sum_{j=1}^{k} c_ic_jE_iE_j \\ &= \sum_{i=1}^{k} c_i^2E_i^2 \\ &= \sum_{i=1}^{k} c_i^2E_i. \end{aligned}$$

O leitor deve comparar isto com $g(A)$ sendo A uma matriz diagonal, pois neste caso $g(A)$ é simplesmente a matriz diagonal de elementos diagonais $g(A_{11}), \ldots, g(A_{nn})$.

Gostaríamos de observar em particular o que acontece no caso dos polinômios de Lagrange correspondentes aos escalares $c_1, \ldots, \ldots, c_k$:

$$p_j = \prod_{i \neq j} \frac{(x - c_i)}{(c_j - c_i)}.$$

Temos $p_j(c_i) = \delta_{ij}$, o que significa que

$$\begin{aligned} p_j(T) &= \sum_{i=1}^{k} \delta_{ij}E_i \\ &= E_j. \end{aligned}$$

Assim, as projeções E_j não apenas comutam com T, mas são polinômios em T.

Notemos também que se g é um polinômio, a fórmula

$$g(T) = g(c_1)E_1 + \ldots + g(c_k)E_k$$

nos mostra que $g(T) = 0$ se, e sòmente se, $g(c_i) = 0$ para cada i. A partir disto é evidente que o polinômio minimal de T é

$$p = (x - c_1) \ldots (x - c_k).$$

Teorema 10. *Seja* T *um operador linear sôbre o espaço* V *de dimensão finita sôbre o corpo* F. *Então, uma condição necessária e suficiente para que* T *seja diagonalizável é que o polinômio minimal de* T *seja da forma*

$$p = (x - c_1) \ldots (x - c_k)$$

onde $c_1, \ldots, c_k$ *são escalares distintos em* F.

Demonstração. Acabamos de ver que se T é diagonalizável, então o polinômio minimal para T é da forma dada. Suponhamos agora que T seja um operador linear sôbre V com polinômio minimal

$$p = (x - c_1) \ldots (x - c_k)$$

onde $c_1, \ldots, c_k$ são elementos *distintos* de F. Formemos os polinômios de Lagrange

$$p_j = \prod_{i \neq j} \frac{(x - c_i)}{(c_j - c_i)}.$$

Recordemos do Capítulo 4 que $p_j(c_i) = \delta_{ij}$ e, para todo polinômio g de grau menor ou igual a $(k - 1)$, temos

$$g = g(c_1)p_1 + \ldots + g(c_k)p_k.$$

Tomando g como sendo o polinômio constante 1 e logo em seguida como o polinômio x, temos

$$\begin{aligned} 1 &= p_1 + \ldots + p_k \\ x &= c_1p_1 + \ldots + c_kp_k \end{aligned} \tag{6-13}$$

(O leitor astuto terá notado que a aplicação a x pode não ser válida porque k pode ser igual a 1. Mas se $k = 1$, T é um múltiplo escalar do operador idêntico, sendo portanto diagonalizável.) Seja agora $E_j = p_j(T)$. De (6-13) temos

$$\begin{aligned} I &= E_1 + \ldots + E_k \\ T &= c_1E_1 + \ldots + c_kE_k. \end{aligned} \tag{6-14}$$

Observemos que se $i \neq j$ então $p_i p_j$ é divisível pelo polinômio minimal p, pois $p_i p_j$ contém cada $(x - c_r)$ como um fator. Assim

$$E_i E_j = 0, \quad i \neq j. \tag{6-15}$$

Precisamos notar ainda mais um fato, a saber, que $E_i \neq 0$ para todo i. Isto ocorre porque p é o polinômio minimal de T e então não podemos ter $p_i(T) = 0$, pois p_i tem grau menor que o de p. Êste último comentário, junto com (6-14), (6-15) e o fato de que os c_i são distintos nos permite aplicar o Teorema 9 e concluir que T é diagonalizável.

Corolário. *Seja* A *uma* n $\times$ n *matriz sôbre o corpo* F. *Então A é semelhante sôbre* F *a uma matriz diagonal se, e sòmente se, o polinômio minimal* p *de* A *é da forma*

$$p = (x - c_1) \dots (x - c_k)$$

onde $c_1, \dots, c_k$ *são elementos distintos de* F.

Exemplo 10. Seja F um subcorpo do corpo dos números complexos e seja T o operador linear sôbre F^3 representado em relação à base ordenada canônica pela matriz A dos Exemplos 8 e 9. Já sabemos que o polinômio característico de T é $(x - 1)(x - 2)^2$ e que o polinômio minimal de T é $(x - 1)(x - 2)$. Como o polinômio minimal é um produto de polinômio sôbre F que são lineares, unitários e distintos, sabemos de imediato que T é diagonalizável. Sabemos ainda que existe uma base ordenada de F^3 em relação à qual T é representado pela matriz

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

O '2' aparece duas vêzes na diagonal porque $(x - 2)$ ocorre duas vêzes no polinômio característico. Podemos determinar uma tal base ordenada determinando simplesmente bases dos núcleos de $(T - I)$ e $(T - 2I)$. Mas, com o objetivo de alcançar alguma compreensão da demonstração do Teorema 10, vamos proceder como segue. O polinômios de Lagrange do Teorema 10 são

$$p_1 = \frac{(x - 2)}{(1 - 2)} = 2 - x$$

$$p_2 = \frac{(x - 1)}{(2 - 1)} = x - 1.$$

Portanto, $E_1 = 2I - T$ e $E_2 = T - I$. As matrizes de E_1 e E_2 em relação à base canônica são então

$$2I - A = \begin{bmatrix} -3 & 6 & 6 \\ 1 & -2 & -2 \\ -3 & 6 & 6 \end{bmatrix} \quad A - I = \begin{bmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{bmatrix}.$$

É óbvio, já à primeira vista, que uma base da imagem de E_1 é o vetor $(-3, 1, -3)$, pois as colunas de $2I - A$ geram a imagem de E_1. Anàlogamente, vê-se que um base da imagem de E_2 é

$$\{(4, -1, 3), (-6, 3, -6)\}.$$

Êstes três vetores juntos formam uma base de F^3 em relação à qual a matriz de T será diagonal. Em linguagem matricial, se fizermos

$$P = \begin{bmatrix} -3 & 4 & -6 \\ 1 & -1 & 3 \\ -3 & 3 & -6 \end{bmatrix}$$

então $P^{-1}AP$ será a matriz diagonal de elementos diagonais 1, 2, 2.

Para o operador diagonalizável arbitrário T do Teorema 9, deve-se observar o seguinte: O polinômio característico de T será

$$f = (x - c_1)^{d_1} \dots (x - c_k)^{d_k}$$

onde d_i é a dimensão do subespaço W_i. Isto é evidente a partir da versão matricial do Teorema 9. Assim, *para um operador diagonalizável* T, *a dimensão do espaço dos vetores característicos asssociados ao valor característico* c_i *é a multiplicidade de* c_i *como uma raiz do polinômio característico de* T.

Outro fato que merece ser notado é que, sendo as projeções E_i polinômios em T e $T = c_1E_1 + \dots + c_kE_k$, um operador U comuta com T se, e sòmente se, U comuta com cada um dos E_j. Como observamos na Seção 6.1, a afirmação de que U comuta com cada E_i significa exatamente que cada W_i é invariante sob U. Isto nos conduz a um teorema interessante.

Teorema 11. *Sejam* T *e* U *operadores lineares diagonalizáveis sôbre o espaço* V *de dimensão finita. Se* T *e* U *comutam, então êles são simultâneamente diagonalizáveis, isto é, existe uma base de* V *na qual cada vetor é um vetor característico de* T *e também um vetor característico de* U.

Demonstração. Seja $T = c_1E_1 + \dots + c_kE_k$, onde os c_i são os valores característicos de T e as E_i são projeções tais que $I = E_1 + \dots + E_k$ e $E_iE_j = 0$ se $i \neq j$. Como observamos acima, o fato de que U comuta com T significa que U comuta com cada E_i, ou seja, que a imagem W_i de E_i é invariante sob U. Seja U_i o operador linear sôbre W_i induzido por U, isto é, a restrição de U a W_i. Como U é diagonalizável, cada U_i também o é; de fato, o polinômio minimal de U é da forma

$$g = (x - d_1) \ldots (x - d_r)$$

com os d_i distintos. É evidente que $g(U_i) = 0$ para todo i, pois $g(U) = 0$. Isto significa que o polinômio minimal de U_i necessàriamente divide g; logo, êste polinômio minimal possui raízes distintas, portanto U_i é diagonalizável. Isto significa que em cada W_i podemos encontrar uma base $\mathcal{B}_i$, cujos vetores sejam vetores característicos de U_i e conseqüentemente de U. Como todo vetor em W_i é um vetor característico de T, a reunião das bases $\mathcal{B}_i$ é uma base de V, formada por vetores que são vetores característicos tanto de T como de U.

O leitor não deverá encontrar dificuldades para estender o Teorema 11 para um conjunto finito arbitrário de operadores sôbre V, que comutem e sejam diagonalizáveis. Deverá também notar que a mesma base que diagonaliza simultâneamente um número finito de operadores lineares também diagonaliza tôda combinação linear dos mesmos. Assim, decorre que se temos um conjunto *arbitrário* de operadores diagonalizáveis sôbre V, todos comutando entre si, existe uma base de V que diagonaliza todos êsses operadores; de fato, basta extrair um conjunto linearmente independente maximal da família de operadores comutáveis; esta família será finita e sua diagonalização irá evidentemente diagonalizar tôda a família.

Corolário. *Sejam* A *e* B n $\times$ n *matrizes sôbre o corpo* F, *sendo cada uma semelhante sôbre* F *a uma matriz diagonal. Se* A *e* B *comutam existe uma* n $\times$ n *matriz inversível* P *sôbre* F *tal que tanto* $P^{-1}AP$ *como* $P^{-1}BP$ *são diagonais.*

Exercícios

1. Seja T o operador linear sôbre R^3 que é representado em relação à base ordenada canônica pela matriz

$$\begin{bmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{bmatrix}.$$

Demonstrar que T é diagonalizável mostrando uma base de R^3, formada por vetores característicos de T.

2. Para a matriz A do Exercício 1, determinar matrizes E_1 e E_2 tais que $A = c_1E_1 + c_2E_2$ (c_1 e c_2 são os valores característicos de A), $E_1 + E_2 = I$, $E_1E_2 = 0$.

3. A matriz

$$A = \begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix}$$

é semelhante sôbre o corpo R a alguma matriz diagonal? Esta matriz é semelhante sôbre o corpo C a alguma matriz diagonal?

4. Responder às perguntas do Exercício 3 para a matriz

$$A = \begin{bmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{bmatrix}$$

5. Seja T um operador linear sôbre o espaço vetorial n-dimensional V e suponhamos que T possua n valores característicos *distintos*. Demonstrar que T é diagonalizável.

6. Seja T um operador linear sôbre um espaço vetorial de dimensão finita sôbre o corpo dos números complexos. Demonstrar que T é diagonalizável se, e sòmente se, existe um polinômio g sôbre C que tenha raízes distintas e seja tal que $g(T) = 0$.

7. Seja T um operador linear diagonalizável sôbre o espaço V de dimensão finita e seja W um subespaço qualquer, invariante sob T. Demonstrar que o operador linear T_W que T induz sôbre W é diagonalizável.

8. Determinar uma matriz real inversível P tal que $P^{-1}AP$ e $P^{-1}BP$ sejam ambas diagonais, sendo A e B as matrizes reais

(a) $$A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 3 & -8 \\ 0 & -1 \end{bmatrix}$$

(b) $$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}$$

9. Suponhamos que A seja uma 2×2 matriz com elementos reais que seja simétrica ($A^t = A$). Demonstrar que A é semelhante sôbre R a uma matriz diagonal.

10. Seja T um operador linear sôbre o espaço V de dimensão finita e suponhamos que o polinômio característico de T seja

$$f = (x - c_1)^{d_1} \dots (x - c_k)^{d_k}$$

sendo $c_1, \dots, c_k$ escalares distintos. Demonstrar que T é diagonalizável se, e sòmente se, a dimensão de $(T - c_iI)$ é d_i para $i = 1, \dots, k$.

11. Seja A uma $n \times n$ matriz *diagonal* com polinômio característico $(x - c_1)^{d_1} \dots (x - c_k)^{d_k}$, sendo os c_i distintos. Seja V o espaço das $n \times n$ matrizes B tais que $AB = BA$. Demonstrar que a dimensão de V é $d_1^2 + \dots + d_k^2$.

12. Seja N uma 2×2 matriz complexa tal que $N^2 = 0$. Demonstrar que ou $N = 0$ ou N é semelhante a

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

13. Usar o resultado do Exercício 12 para demonstrar o seguinte: Se A é uma 2×2 matriz com elementos complexos, então A é semelhante sôbre C a uma matriz de um dos tipos seguintes:

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \quad \begin{bmatrix} a & 0 \\ 1 & a \end{bmatrix}.$$

14. Seja F um corpo, n um inteiro positivo e seja V o espaço das $n \times n$ matrizes sôbre F. Se B é uma $n \times n$ matriz fixa sôbre F, seja T_B o operador linear sôbre V definido por $T_B(A) = AB - BA$. Consideremos a família dos operadores lineares T_B obtida fazendo B percorrer o conjunto das matrizes diagonais. Demonstrar que os operadores desta família são simultâneamente diagonalizáveis.

6.4 O Teorema da Decomposição Primária

Estamos tentando estudar um operador linear I sôbre o espaço de V de dimensão finita, pela decomposição de T numa soma direta de operadores que sejam, num certo sentido, elementares. Podemos fazê-lo através dos valores e vetores característicos de T em certos casos particulares, isto é, quando o polinômio de T decompõe-se sôbre o corpo F de escalares num produto de polinômios unitários, distintos e de grau 1. Que podemos fazer com um T arbitrário? Se tentarmos estudar T usando valores característicos, iremos nos confrontar com dois problemas. Primeiro, T poderá não ter nenhum valor característico; isto, na verdade, é uma deficiência do corpo de escalares, a saber, que êle não é algèbricamente fechado. Segundo, mesmo que o polinômio característico se decomponha completamente sôbre F num produto de polinômios de grau 1, podem não existir vetores característicos suficientes para que T gere o espaço V; isto é, evidentemente uma deficiência de T. A segunda situação é ilustrada pelo operador T sôbre F^3 (F um corpo arbitrário) representado em relação à base canônica por

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

O polinômio característico de A é $(x - 2)^2 (x + 1)$ e êste é obviamente o polinômio minimal de A (ou de T). Assim, T não é diagonalizável. Vê-se que isto ocorre porque o núcleo de $(T - 2I)$ tem dimensão 1 apenas. Por outro lado, o núcleo de $(T + I)$ e o núcleo de $(T - 2I)^2$ juntos geram V, sendo o primeiro o subespaço gerado por ϵ_3 e o segundo o subespaço gerado por ϵ_1 e ϵ_2.

Êste será mais ou menos o nosso método geral para o segundo problema. Se (lembrar que isto é uma hipótese) o polinômio minimal de T se decompõe como

$$p = (x - c_1)^{r_1} \dots (x - c_k)^{r_k}$$

sendo $c_1, \dots, c_k$ elementos distintos de F, então mostraremos que o espaço V é a soma direta dos núcleos de $(T - c_iI)^{r_i}$, $i = 1, \dots, k$. O operador diagonalizável é o caso particular dêste em que $r_i = 1$ para cada i. O teorema que demonstraremos é mais geral que isto que descrevemos, uma vez que trabalhará com a decomposição primária do polinômio minimal, quer sejam ou não de grau 1 os primos que comparecem na decomposição. O leitor achará útil pensar no caso particular em que os primos são de grau 1 e, de modo ainda mais particular, pensar na demonstração do Teorema 10 como um caso particular dêste teorema.

Teorema 12 (Teorema da Decomposição Primária). *Seja* T *um operador linear sôbre o espaço vetorial* V *de dimensão finita sôbre o corpo* F. *Seja* p *o polinômio minimal de* T,

$$p = p_1^{r_1} \dots p_k^{r_k}$$

onde os p_i *são polinômios distintos, irredutíveis e unitários sôbre* F *e os* r_i *são inteiros positivos. Seja* W_i *o núcleo de* $p_i(T)^{r_i}$, $i = 1, \dots, k$. *Então*

(a) $V = W_1 \oplus \dots \oplus W_k$

(b) *cada* W_i *é invariante sob* T

(c) *Se* T_i *é o operador induzido sôbre* W_i *por* T, *então o polinômio minimal de* T_i *é* $p_i^{r_i}$.

Demonstração. A idéia da demonstração é a que segue. Se a decomposição (a) em soma direta é válida, como podemos obter as projeções $E_1, \dots, E_k$ associadas a esta decomposição? A projeção E_i será o operador idêntico sôbre W_i e zero sôbre os outros W_j. Vamos determinar um polinômio h_i tal que $h_i(T)$ seja o operador idêntico sôbre W_i e seja nulo sôbre os outros W_j e então que $h_1(T) + \dots + h_k(T) = I$, etc.

Para cada i, seja

$$f_i = p/p_i^{r_i} = \prod_{j \neq i} p_j^{r_j}. \tag{6-16}$$

Como $p_1, \dots, p_k$ são polinômios primos distintos, os polinômios $f_1, \dots, f_k$ são relativamente primos (Teorema 8, Capítulo 4). Assim existem polinômios $g_1, \dots, g_k$ tais que

$$\sum_{i=1}^{n} f_i g_i = 1. \tag{6-17}$$

Notemos também que se $i \neq j$, então $f_i f_j$ é divisível pelo polinômio p pois $f_i f_j$ contém cada $p_m^{r_m}$ como um fator. Vamos mostrar que os polinômios $h_i = f_i g_i$ comportam-se da maneira descrita no primeiro parágrafo da demonstração.

Seja $E_i = h_i(T) = f_i(T)g_i(T)$. Como $h_1 + \ldots + h_k = 1$ e p divide $f_i f_j$ para $i \neq j$, temos

$$\begin{aligned} E_1 + \ldots + E_k &= I \\ E_i E_j &= 0, \text{ se } i \neq j. \end{aligned}$$

Assim, os E_i são projeções que correspondem a alguma decomposição do espaço V em soma direta. Desejamos mostrar que a imagem de E_i é exatamente o subespaço W_i. É evidente que cada vetor na imagem de E_i está em W_i, pois se α está na imagem de E_i, então $\alpha = E_i\alpha$, logo

$$\begin{aligned} p_i(T)^{r_i}\alpha &= p_i(T)^{r_i}E_i\alpha \\ &= p_i(T)^{r_i}f_i(T)g_i(T)\alpha \\ &= 0 \end{aligned}$$

pois $p^{r_i}f_i g_i$ é divisível pelo polinômio minimal p. Reciprocamente, suponhamos que α esteja no núcleo de $p_i(T)^{r_i}$. Se $j \neq i$, então $f_j g_j$ é divisível por $p_i^{r_i}$, logo $f_j(T)g_j(T)\alpha = 0$, isto é, $E_j\alpha = 0$ para $j \neq i$. Mas então é imediato que $E_i\alpha = \alpha$, isto é, que α está na imagem de E_i. Isto completa a demonstração da afirmação (a).

É certamente óbvio que os subespaços W_i são invariantes sob T. Se T_i é o operador induzido sôbre W_i por T, então, evidentemente $p_i(T_i)^{r_i} = 0$, pois, por definição, $p_i(T)^{r_i}$ se anula no subespaço W_i. Isto mostra que o polinômio minimal de T_i divide p^{r_i}. Reciprocamente, seja g um polinômio arbitrário tal que $g(T_i) = 0$. Então $g(T)f_i(T) = 0$. Assim, gf_i é divisível pelo polinômio minimal p de T, isto é, $p_i^{r_i}f_i$ divide gf_i. Vê-se fàcilmente que $p_i^{r_i}$ divide g. Logo, o polinômio minimal de $T_i = p_i^{r_i}$.

Corolário. *Se* $E_1, \ldots, E_k$ *são as projeções associadas à decomposição primária de* T, *então cada* E_i *é um polinômio em* T *e, conseqüentemente, se um operador linear* U *comuta com* T *então* U *comuta com cada um dos* E_i, *isto é, cada subespaço* W_i *é invariante sob* U.

Com a notação do Teorema 12, vamos considerar ràpidamente o caso particular em que o polinômio minimal de T é um produto de polinômios do primeiro grau, isto é, o caso em que cada p_i é da forma $p_i = x - c_i$. Ora, a imagem de E_i é o núcleo W_i de $(T - c_iI^{r_i})$. Coloquemos $D = c_1E_1 + \ldots + c_kE_k$. Pelo Teorema

9, D é um operador diagonalizável que denominaremos a **parte diagonalizável** de T. Consideremos o operador $N = T - D$. Ora,

$$T = TE_1 + \ldots + TE_k$$
$$D = c_1E_1 + \ldots + c_kE_k$$

portanto $$N = (T - c_1I)E_1 + \ldots + (T - c_kI)E_k.$$

A esta altura o leitor já deverá estar suficientemente familiarizado com projeções, portanto verá que

$$N^2 = (T - c_1I)^2E_1 + \ldots + (T - c_kI)^2E_k$$

e que, em geral,

$$N^r = (T - c_1I)^rE_1 + \ldots + (T - c_kI)^rE_k.$$

Quando $r \geq r_i$ para todo i, teremos $N^r = 0$, pois o operador $(T - c_iI)^r$ será então 0 sôbre a imagem de E_i.

Definição. *Seja* N *um operador linear sôbre o espaço vetorial* V. *Dizemos que* N *é* **nilpotente** *se existe algum inteiro positivo* r *tal que* $N^r = 0$.

Teorema 13. *Seja* T *um operador linear sôbre o espaço vetorial de dimensão finita sôbre o corpo* F. *Suponhamos que o polinômio minimal de* T *se decomponha sôbre* F *num produto de polinômios lineares. Então existe um operador diagonalizável* D *sôbre* V *e um operador nilpotente* N *sôbre* V *tais que*

(a) $T = D + N$,

(b) $DN = ND$.

O operador diagonalizável D *e o operador nilpotente* N *são determinados de modo único por (a) e (b) e cada um dêles é um polinômio em* T.

Demonstração. Acabamos de observar que podemos escrever $T = D + N$ onde D é diagonalizável e N é nilpotente e também que D e N não só comutam mas também são polinômios em T. Suponhamos agora que também tenhamos $T = D' + N'$ sendo D' diagonalizável e $D'N' = N'D'$. Vamos demonstrar que $D = D'$ e $N = N'$.

Como D' e N' comutam entre si e $T = D' + N'$, vemos que D' e N' comutam com T. Assim, D' e N' comutam com qualquer polinômio em T, logo êles comutam com D e com N. Agora temos

$$D + N = D' + N'$$

ou $$D - D' = N' - N$$

e todos êstes quatro operadores comutam entre si. Como D e D' são ambos diagonalizáveis e comutam, êles são simultâneamente diagonalizáveis e $D - D'$ é diagonalizável. Como N e N' são ambos nilpotentes e comutam, o operador $(N' - N)$ é nilpotente; com efeito, usando o fato de que N e N' comutam

$$(N' - N)^r = \sum_{j=0}^{r} \binom{r}{j} (N')^{r-j}(-N)^j$$

e então quando r fôr suficientemente grande todos os têrmos nesta expressão de $(N' - N)^r$ serão nulos. (Na realidade, um operador nilpotente sôbre um espaço n-dimensional deve ter sua n-ésima potência nula; no caso acima, se tomamos $r = 2n$, êste número é suficientemente grande. Decorre então que $r = n$ é suficientemente grande, mas isto não é evidente a partir da expressão acima.) Ora, $D - D'$ é um operador diagonalizável que também é nilpotente. Tal operador tem de ser, òbviamente, o operador nulo; de fato, sendo nilpotente, o polinômio minimal dêste operador é da forma x^r para um certo $r \leq m$; mas como o operador é diagonalizável, o polinômio minimal não pode ter uma raiz múltipla, logo $r = 1$ e o polinômio minimal é simplesmente x, o que diz que o operador é nulo. Assim, vemos que $D = D'$ e $N = N'$.

Corolário. *Seja* V *um espaço vetorial de dimensão finita sôbre um corpo algèbricamente fechado* F, *por exemplo, o corpo dos números complexos. Então, todo operador linear* T *sôbre* V *pode ser escrito como uma soma de um operador diagonalizável* D *com um operador nilpotente* N *os quais comutam. Êstes operadores* D *e* N *são únicos e são polinômios em* T.

Por êsses resultados vê-se que o estudo dos operadores lineares sôbre espaços vetoriais sôbre um corpo algèbricamente fechado fica essencialmente reduzido ao estudo dos operadores nilpotentes. Para espaços vetoriais sôbre corpos não algèbricamente fechados, precisamos ainda encontrar algum substitituto para valores característicos. Um fato interessante é que êstes dois problemas podem ser tratados simultâneamente e é isto o que fazemos no próximo capítulo.

Concluindo esta seção, gostaríamos de dar um exemplo que ilustre algumas das idéias do teorema da decomposição primária. Decidimos dá-lo no final da seção porque êle usa equações diferenciais, não sendo assim álgebra linear pura.

Exemplo 11. No Teorema da decomposição primária não é necessário que o espaço vetorial V tenha dimensão finita, nem é necessário, para as partes (a) e (b), que p seja o polinômio minimal de T. Se T é um operador linear sôbre um espaço vetorial arbitrário e *se* existe um polinômio unitário p tal que $p(T) = 0$, então as partes (a) e (b) do Teorema 12 são válidas para T com a demonstração que fizemos.

Seja n um inteiro positivo e seja T o espaço das funções n vêzes contìnuamente diferenciáveis sôbre a reta real que satisfazem a equação diferencial

$$\frac{d^n f}{dt^n} + a_{n-1}\frac{d^{n-1}f}{dt^{n-1}} + \ldots + a_1\frac{df}{dt} + a_0 f = 0 \tag{6-18}$$

onde $a_0, \ldots, a_{n-1}$ são certos escalares fixos. Se C_n indica o espaço das funções n vêzes contìnuamente diferenciáveis, então o espaço T das soluções desta equação diferencial é um subespaço de C_n. Se D indica o operador derivação e p é o polinômio

$$p = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$$

então V é o núcleo do operador $p(D)$, pois (6-18) diz simplesmente que $p(D)f = 0$. Consideremos agora D como um operador linear sôbre o subespaço V. Então $p(D) = 0$.

Se estamos discutindo funções diferenciáveis com valores complexos, então C_n e V são espaços vetoriais complexos e $a_0, \ldots, a_{n-1}$ podem ser quaisquer números complexos. Dscrevamos agora

$$p = (x - c_1)^{r_1} \ldots (x - c_k)^{r_k}$$

onde $c_1, \ldots, c_k$ são números complexos distintos. Se W_j é o núcleo de $(D - c_j I)^{r_j}$, então o Teorema 12 diz que

$$V = W_1 \oplus \ldots \oplus W_k.$$

Em outras palavras, se f satisfaz a equação diferencial (6-18), então f pode ser expressa de modo único sob a forma

$$f = f_1 + \ldots + f_k$$

onde f_j satisfaz a equação diferencial $(D - c_j I)^{r_j} f_j = 0$. Assim, o estudo das soluções da equação (6-18) fica reduzido ao estudo do espaço das soluções de uma equação diferencial da forma

$$(D - cI)^r f = 0. \tag{6-19}$$

Esta redução foi conseguida por meio dos métodos gerais de álgebra linear, isto é, pelo teorema da decomposição primária.

Para descrever o espaço das soluções de (6-19), é necessário saber alguma coisa sôbre equações diferenciais, isto é, é necessário saber alguma coisa a respeito de D além do fato de que D é um operador linear. No entanto, não se precisa saber muito. É bem fácil demonstrar por indução sôbre r que, se f está em C_r, então

$$(D - cI)^r f = e^{ct} D^r (e^{-ct} f)$$

isto é,

$$\frac{df}{dt} - cf(t) = e^{ct} \frac{d}{dt} (e^{-ct} f), \text{ etc.}$$

Assim, $(D - cI)^r f = 0$ se, e sòmente se, $D^r(e^{-ct}f) = 0$. Uma função g tal que $D^r g = 0$, isto é, $d^r g/dt^r = 0$, deve ser uma função polinomial de grau menor ou igual a $(r - 1)$:

$$g(t) = b_0 + b_1 t + \ldots + b_{r-1} t^{r-1}.$$

Assim f satisfaz (6-19) se, sòmente se, f tem a forma

$$f(t) = e^{ct}(b_0 + b_1 t + \ldots + b_{r-1} t^{r-1}).$$

Conseqüentemente, as "funções" e^{ct}, $te^{ct}, \ldots, t^{r-1}e^{ct}$ geram o espaço das soluções de (6-19). Como $1, t, \ldots, t^{r-1}$ são funções linearmente independentes e a função exponencial não possui raízes, estas r funções $t^j e^{ct}$, $0 \leq j \leq r - 1$, formam uma base do espaço das soluções.

Voltando à equação diferencial (6-18), que é

$$p(D)f = 0$$
$$p = (x - c_1)^{r_1} \ldots (x - c_k)^{r_k}$$

vemos que as n funções $t^m e^{c_j t}$, $0 \leq m \leq r_j - 1$, $1 \leq j \leq k$, formam uma base do espaço das soluções de (6-18). Em particular, o espaço das soluções tem dimensão finita igual ao grau do polinômio p.

Exercícios

1. Seja T um operador linear sôbre R^3 que é representado em relação à base ordenada canônica pela matriz

$$\begin{bmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{bmatrix}.$$

Exprimir o polinômio minimal p de T sob a forma $p = p_1 p_2$, sendo p_1 e p_2 unitários e irredutíveis sôbre o corpo dos números reais. Seja W_i o núcleo de $p_i(T)$. Determinar bases $\mathfrak{B}_i$ dos espaços W_1 e W_2. Se T_i é o operador induzido sôbre W_i por T, determinar a matriz de T_i em relação à base $\mathfrak{B}_i$ (acima).

2. Seja T o operador linear sôbre R^3 que é representado pela matriz

$$\begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix}$$

em relação à base ordenada canônica. Mostrar que existe um operador diagonalizável D sôbre R^3 e um operador nilpotente N sôbre R^3 tais que $T = D + N$ e $DN = ND$. Determinar as matrizes de D e N em relação à base canônica. (Basta repetir a demonstração do Teorema 12 para êste caso particular.)

3. Se V é o espaço dos polinômios de grau menor ou igual a n sôbre um corpo F, mostrar que o operador derivação sôbre V é nilpotente.

4. Seja T um operador linear sôbre o espaço V de dimensão finita cujo polinômio característico seja

$$f = (x - c_1)^{d_1} \dots (x - c_k)^{d_k}$$

e cujo polinômio minimal seja

$$p = (x - c_1)^{r_1} \dots (x - c_k)^{r_k}.$$

Seja W_i o núcleo de $(T - c_iI)^{r_i}$.

(a) Demonstrar que W_i é o conjunto dos vetores α em V tais que $(T - c_iI)^m \alpha = 0$ para algum inteiro positivo m (que pode depender de α).

(b) Demonstrar que a dimensão de W_i é d_i. [*Sugestão*: Se T_i é o operador induzido sôbre W_i por T, então $T_i - c_iI$ é nilpotente; assim, o polinômio característico de $T_i - c_iI$ deve ser x^{e_i}, sendo e_i a dimensão de W_i (demonstração?); assim, o polinômio característico de T_i é $(x - c_i)^{e_i}$; agora usar o fato de que o polinômio característico de T é o produto dos polinômios característicos de T_i para mostrar que $e_i = d_i$.]

5. Seja V um espaço vetorial de dimensão finita sôbre o corpo dos números complexos. Seja T um operador linear sôbre V e seja D a parte diagonalizável de T. Demonstrar que se g é um polinômio qualquer com coeficientes complexos, então a parte diagonalizável de $g(T)$ é $g(D)$.

6. Seja V um espaço vetorial de dimensão finita sôbre o corpo F e seja T um operador linear sôbre T tal que pôsto $(T) = 1$. Demonstrar que ou T é diagonalizável ou T é nilpotente, não ambos.

7. Seja V um espaço vetorial de dimensão finita sôbre F e seja T um operador linear sôbre V. Suponhamos que T comute com todo operador linear diagonalizável sôbre V. Demonstrar que T é um múltiplo escalar do operador idêntico.

8. Seja V o espaço das $n \times n$ matrizes sôbre um corpo F e seja A uma $n \times n$ matriz fixa sôbre F. Definamos um operador linear T sôbre V por $T(B) = AB - BA$. Demonstrar que se A é uma matriz nilpotente, então T é um operador nilpotente.

9. Dar um exemplo de duas 4×4 matrizes nilpotentes que tenham o mesmo polinômio minimal (elas têm, necessàriamente, o mesmo polinômio característico) mas que não sejam semelhantes.

10. Seja T um operador linear sôbre o espaço V de dimensão finita, seja $p = p_1^{r_1} \dots p_k^{r_k}$ o polinômio minimal de T e seja $V = W_1 \oplus \dots \oplus W_k$ a decomposição primária de T, isto é, W_j é o núcleo de $p_j(T)^{r_j}$. Seja W um subespaço qualquer de V que seja invariante sob T. Demonstrar que

$$W = (W \cap W_1) \oplus (W \cap W_2) \oplus \dots \oplus (W \cap W_k).$$

CAPÍTULO 7

AS FORMAS RACIONAL E DE JORDAN

7.1 Subespaços Cíclicos e Anuladores

Uma vez mais V é um espaço vetorial de dimensão finita sôbre o corpo F e T é um operador linear fixo (mas arbitrário) sôbre V. Se α é um vetor qualquer em V, existe um menor subespaço de T que é invariante sob T e contém α. Êste subespaço pode ser definido como a interseção de todos os subespaços T-invariantes que contêm α; contudo, é mais proveitoso no momento considerarmos as coisas de outra maneira. Se W é um subespaço arbitrário de V que seja invariante sob T e contenha α, então W deve conter o vetor $T\alpha$; portanto, W deve conter $T(T\alpha) = T^2\alpha$, $T(T^2\alpha) = T^3\alpha$, etc. Em Em outras palavras, W deve conter $g(T)\alpha$ para todo polinômio g sôbre F. O conjunto dos vetores da forma $g(T)\alpha$, com g em $F[x]$ é evidentemente invariante sob T e é assim o menor subespaço T-invariante que contém.

Definição. *Se* α *é um vetor qualquer em* V, *o* **subespaço T-cíclico gerado por** α *e o subespaço* $Z(\alpha;\ T)$ *dos vetores da forma* $g(T)\alpha$, g *em* $F[x]$. *Se* $Z(\alpha;\ T) = V$, *então* α *é denominado um* **vetor cíclico** *de* T.

Outra maneira de descrever o subespaço $Z(\alpha;T)$ é dizer que $Z(\alpha;\ T)$ é o subespaço gerado pelos vetores $T^k\alpha$, $k \geq 0$, e assim α é um vetor cíclico de T se, e sòmente se, êstes vetores geram V. Prevenimos o leitor de que um operador genérico T não possui vetores cíclicos.

Exemplo 1. Para T arbitrário, o subespaço T-cíclico gerado pelo vetor nulo é o subespaço nulo. O espaço $Z(\alpha;T)$ é unidimensional se, e sòmente se, α é um vetor característico de T. Para o operador idêntico, todo vetor não-nulo gera um subespaço cíclico unidimensional; assim, se dim $V > 1$, o operador idêntico não possui nenhum vetor cíclico. Um exemplo de um operador que possui um vetor cíclico é o operador linear T sôbre F^2 que é representado em relação à base ordenada canônica pela matriz

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Neste caso o vetor cíclico (um vetor cíclico) é ϵ_1; de fato, se $\beta = (a, b)$, então, para $g = a + bx$, temos $\beta = g(T)\epsilon_1$. Para êste mesmo operador T, o subespaço cíclico gerado por ϵ_2 é o espaço unidimensional gerado por ϵ_2, porque ϵ_2 é um vetor característico de T.

Para quaisquer T e α, estaremos interessados em relações lineares

$$c_0\alpha + c_1T\alpha + \ldots + c_kT^k\alpha = 0$$

entre os vetores $T^j\alpha$, isto é, estaremos interessados nos polinômios $g = c_0 + c_1x + \ldots + c_kx^k$ que tenham a propriedade de que $g(T)\alpha = 0$. O conjunto dos g em $F[x]$ tais que $g(T)\alpha = 0$ é evidentemente um ideal em $F[x]$. É também um ideal não-nulo, pois contém o polinômio minimal p do operador $T(p(T)\alpha = 0$ para todo α em $V)$.

Definição. *Se α é um vetor arbitrário em* V, *o* **T-anulador** *de α é o ideal* M(α; T) *em* F[x] *formado pelos polinômios* g *sôbre F tais que* g(T)α = 0. *O único polinômio unitário* p_α *que gera êste ideal também será denominado o* **T-anulador** *de α.*

Como ressaltamos acima, o T-anulador p_α divide o polinômio minimal do operador T. O leitor deverá notar também que $gr(p_\alpha) > 0$ a não ser quando α é o vetor nulo.

Teorema 1. *Seja α um vetor nulo arbitrário em* V *e seja* p_α *o* T-*anulador de α.*

(a) *O grau de* p_α *é igual à dimensão do subespaço cíclico* Z(α; T).

(b) *Se o grau de* p_α *é* k, *então os vetores α, $T\alpha$, $T^2\alpha, \ldots, T^{k-1}\alpha$ formam uma base de* Z(α; T).

(c) *Se* U *é o operador linear sôbre* Z(α; T) *induzido por* T, *então o polinômio minimal de* U *e* p_α.

Demonstração. Seja g um polinômio qualquer sôbre o corpo F. Podemos escrever

$$g = p_\alpha q + r$$

onde $r = 0$ ou $gr(r) < gr(p_\alpha) = k$. O polinômio $p_\alpha q$ está no T-anulador de α, portanto

$$g(T)\alpha = r(T)\alpha.$$

Como $r = 0$ ou $\text{gr}(r) < k$, o vetor $r(T)\alpha$ é uma combinação linear dos vetores $\alpha, T\alpha, \ldots, T^{k-1}\alpha$ e como $g(T)\alpha$ é um vetor típico em $Z(\alpha; T)$ isto mostra que êstes k vetores geram $Z(\alpha; T)$. Êstes vetores são, sem dúvida, linearmente independentes, pois tôda relação linear não-trivial entre êles nos forneceria um polinômio não-nulo g tal que $g(T)\alpha = 0$ e $\text{gr}(g) < \text{gr}(p_\alpha)$, o que é absurdo. Isto demonstra (a) e (b).

Seja U o operador linear sôbre $Z(\alpha; T)$ obtido pela restrição de T àquele subespaço. Se g é um polinômio arbitrário sôbre F, então

$$\begin{aligned} p_\alpha(U)g(T)\alpha &= p_\alpha(T)g(T)\alpha \\ &= g(T)p_\alpha(T)\alpha \\ &= g(T)0 \\ &= 0. \end{aligned}$$

Assim, o operador $p_\alpha(U)$ leva todo vetor de $Z(\alpha;T)$ em 0 e é portanto o operador nulo sôbre $Z(\alpha; T)$. Além disso, se h é um polinômio de grau menor que k, não podemos ter $h(U) = 0$, pois então $h(U)\alpha = h(T)\alpha = 0$, contrariando a definição de p_α. Isto mostra que p_α é o polinômio minimal de U.

Uma conseqüência particular dêste teorema é a seguinte: Se acontecer que α seja um vetor cíclico de T, então o polinômio minimal de T deve ter grau igual à dimensão do espaço V; daqui decorre que o polinômio minimal de T é o polinômio característico de T. Demonstraremos posteriormente que para todo T existe um vetor α em V cujo anulador é o polinômio minimal de T. Decorrerá então que T possui um vetor cíclico se, e sòmente se, os polinômios minimal e característico de T são idênticos. Mas teremos algum trabalho para chegarmos a ver isto.

Nosso plano é estudar um T arbitrário usando operadores que possuam um vetor cíclico. Portanto, consideremos um operador linear U sôbre um espaço W de dimensão k que possua um vetor cíclico α. Pelo Teorema 1, os vetores $\alpha, \ldots, U^{k-1}\alpha$ formam uma base do espaço W e o anulador p_α de α é o polinômio minimal de U (logo, e também o polinômio característico de U). Se fizermos $\alpha_i = U^{i-1}\alpha$, $i = 1, \ldots k$, então a ação de U sôbre a base ordenada $\mathcal{B} = \{\alpha_1, \ldots, \alpha_k\}$ é

$$\begin{aligned} &U\alpha_i = \alpha_{i+1}, \quad i = 1, \ldots, k-1 \\ &U\alpha_k = -c_0\alpha_1 - c_1\alpha_2 - \ldots - c_{k-1}\alpha_k \end{aligned} \tag{7-1}$$

onde $p_\alpha = c_0 + c_1x + \ldots + c_{k-1}x^{k-1} + x^k$. A expressão de $U\alpha_k$ decorre do fato de que $p_\alpha(U)\alpha = 0$, isto é,

$$U^k\alpha + c_{k-1}U^{k-1}\alpha + \ldots + c_1U\alpha + c_0\alpha = 0.$$

Isto diz que a matriz de U em relação à base ordenada $\mathcal{B}$ é

$$(7\text{-}2)\qquad \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 & -c_0 \\ 1 & 0 & 0 & \ldots & 0 & -c_1 \\ 0 & 1 & 0 & \ldots & 0 & -c_2 \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 1 & -c_{k-1} \end{bmatrix}$$

A matriz (7-2) é denominada a **matriz associada** ao polinômio unitário p.

Teorema 2. *Se* U *é um operador linear sôbre o espaço de* W *dimensão finita, então* U *possui um vetor cíclico se, e sòmente se, existe alguma base ordenada de* W *em relação à qual* U *é representado pela matriz associada ao polinômio minimal de* U.

Demonstração. Acabamos de observar que se U possui um vetor cíclico, então existe uma tal base ordenada de W. Reciprocamente, se existe uma base ordenada $\{\alpha_1, \ldots, \alpha_k\}$ de W em relação à qual U é representado pela matriz associada ao seu polinômio minimal, é óbvio que α_1 é um vetor cíclico de U.

Corolário. *Se* A *é a matriz associada a um polinômio unitário* p, *então* p *é tanto o polinômio minimal como o polinômio característico de* A.

Demonstração. Se p é um polinômio unitário de grau k sôbre o corpo F, seja U o operador linear sôbre F^k que é representado por A em relação à base ordenada canônica. Basta agora aplicar nossos resultados acima.

No último comentário — se T é um operador linear arbitrário sôbre o espaço V e α é um vetor qualquer em V, então o operador U que T induz sôbre o subespaço cíclico $Z(\alpha; T)$ possui um vetor cíclico, a saber, α. Assim, $Z(\alpha; T)$ possui uma base ordenada em relação à qual U é representado pela matriz associada a p_α, sendo p_α o T-anulador de α.

Exercícios

1. Seja T um operador linear sôbre F^2. Demonstrar que todo vetor não-nulo que não seja um vetor característico de T é um vetor cíclico de T. Depois, demonstrar que ou T possui um vetor cíclico ou então T é um múltiplo escalar do operador idêntico.

2. Seja T o operador linear sôbre R^3 que é representado em relação à base ordenada canônica pela matriz

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Demonstrar que T não possui vetores cíclicos. Qual é o subespaço T-cíclico gerado pelo vetor $(1, -1, 3)$?

3. Seja T o operador linear sôbre C^3 que é representado em relação à base ordenada canônica pela matriz

$$\begin{bmatrix} 1 & i & 0 \\ -1 & 2 & -i \\ 0 & 1 & 1 \end{bmatrix}.$$

Determinar o T-anulador do vetor $(1, 0, 0)$. Determinar o T-anulador de $(1, 0, i)$.

4. Demonstrar que se T^2 possui um vetor cíclico, então T possui um vetor cíclico. A recíproca é verdadeira?

5. Seja V um espaço vetorial n-dimensional sôbre o corpo F e seja N um operador linear nilpotente sôbre V. Suponhamos que $N^{n-1} \neq 0$ e seja α um vetor arbitrário em V tal que $N^{n-1}\alpha \neq 0$. Demonstrar que α é um vetor cíclico de N. Dizer exatamente qual é a matriz de N em relação à base ordenada $\{\alpha, N\alpha, \ldots, N^{n-1}\alpha\}$?

6. Demonstrar diretamente que se A é a matriz associada ao polinômio unitário p, então p é o polinômio característico de A.

7. Seja V um espaço vetorial n-dimensional e seja T um operador linear sôbre V. Suponhamos que T seja *diagonalizável*.

(a) Se T possui um vetor cíclico, mostrar que T possui n-valores característicos distintos.

(b) Se T possui n valores característicos distintos e se $\{\alpha_1, \ldots, \alpha_n\}$ é uma base formada de vetores característicos de T, mostrar que $\alpha = \alpha_1 + \ldots + \alpha_n$ é um vetor cíclico de T.

8. Seja T um operador linear sôbre o espaço vetorial V de dimensão finita. Suponhamos que T possua um vetor cíclico. Demonstrar que se U é um operador linear arbitrário que comuta com T, então U é um polinômio em T.

7.2 O Teorema da Decomposição Racional

O objetivo primordial desta seção é demonstrar que se T é um operador linear qualquer sôbre V, então existem vetores $\alpha_1, \ldots, \alpha_n$ em V tais que

$$V = Z(\alpha_1; T) \oplus \ldots \oplus Z(\alpha_r; T).$$

Em outras palavras, desejamos demonstrar que V é a soma direta de subespaços T-cíclicos. Isto nos mostrará que T é a soma direta de um número finito de operadores lineares, cada um dos quais possui um vetor cíclico. O resultado disto será a redução de muitos problemas relativos a um operador genérico a problemas semelhantes sôbre um operador que possua um vetor cíclico. O teorema que vamos demonstrar, o teorema da decomposição racional, é um tanto mais forte que a afirmação de que V é a soma direta de subespaços T-cíclicos e êste teorema possui muitos corolários interessantes.

Como seria possível demonstrar que V é a soma direta de subespaços T-cíclicos? Uma maneira de tentar resolver êste problema é a seguinte: Tomar algum vetor não-nulo α_1 em V. Se $V = Z(\alpha_1; T)$, ótimo. Se não, tomar um vetor não-nulo α_2 tal que $Z(\alpha_2; T)$ seja disjunto de $Z(\alpha_1, T)$. Se $V = Z(\alpha_1; T) + Z(\alpha_2; T)$, ótimo. Se não, tomar $\alpha_3 \neq 0$ tal que $Z(\alpha_3; T)$ seja disjunto de $Z(\alpha_1; T) + Z(\alpha_2; T)$, etc. Em outras palavras, tentar chegar a $V = Z(\alpha_1; T) \oplus \ldots \oplus Z(\alpha_r; T)$ selecionando indutivamente os vetores $\alpha_1, \ldots, \alpha_r$. Êste é fundamentalmente o método que utilizaremos; no entanto, teremos de suplantar grandes dificuldades relacionadas com a escolha de $\alpha_1, \ldots, \alpha_r$. Um exemplo de tal dificuldade é o seguinte: Se tomamos um vetor não-nulo arbitrário α_1 em V e se $V \neq Z(\alpha_1; T)$, então pode não existir nenhum vetor não-nulo α_2 tal que $Z(\alpha_2; T)$ disjunto de $Z(\alpha_1; T)$. Assim, teremos de ser argutos quanto à seleção dos vetores α_i. Nesta seleção, nosso princípio orientador será a observação que segue.

Suponhamos que $V = Z(\alpha_1; T) \oplus \ldots \oplus Z(\alpha_r; T)$. Se $1 \leq i \leq r$, sejam

$$W_j = Z(\alpha_1; T) + \ldots + Z(\alpha_j; T)$$

e

$$W_j' = Z(\alpha_{j+1}; T) + \ldots + Z(\alpha_r; T).$$

Então $V = W_j \oplus W_j'$ e tanto W_j como W_j' são invariantes sob T. Em geral, se W é um subespaço invariante sob T, não existe um subespaço W' tal que $V = W \oplus W'$ e W' seja invariante sob T. Mas vemos que, quando tivermos uma decomposição de V em subespaços T-cíclicos, cada W_j satisfará esta propriedade. Assim, se tomarmos $\alpha_1, \ldots, \alpha_r$ indutivamente, deveremos ter que, uma vez $\alpha_1, \ldots, \alpha_r$ selecionados, o subespaço $W_j = Z(\alpha_1; T) + \ldots + Z(\alpha_j; T)$ possui esta propriedade.

Vamos agora nos perguntar o seguinte: Se W é um subespaço invariante sob T, quando é que existe um subespaço T-invariante W' tal que $V = W \oplus W'$? Se W é um subespaço do espaço V de dimensão finita, sempre existe um subespaço W' tal que $V = W \oplus W'$. Usualmente existem muitos dêsses subespaços W' e qualquer um um dêles é denominado um **suplementar** de W. Estamos perguntando quando é que um subespaço T-invariante possui um subespaço complementar que também seja invariante sob T. Vamos agora passar a responder esta pergunta. Uma vez feito isto, obteremos fàcilmente uma decomposição de V em subespaços T-cíclicos.

Suponhamos que $V = W \oplus W'$, onde W e W' são ambos invariantes sob T e vejamos o que é possível dizer a respeito do subespaço W. Cada vetor β em V é da forma $\beta = \alpha + \alpha'$, com α em W e α' em W'. Se f é um polinômio sôbre o corpo de escalares, então

$$f(T)\,\beta = f(T)\alpha + f(T)\alpha'.$$

Como W e W' são invariantes sob T, o vetor $f(T)\alpha$ está em W e $f(T)\alpha'$ está em W'. Em particular, $f(T)\,\beta$ está em W se, e sòmente se, $f(T)\alpha' = 0$, isto é, se, e sòmente se, $f(T)\beta = f(T)\alpha$. Assim, se W possui algum subespaço T-invariante suplementar, então W necessàriamente tem a seguinte propriedade: Se β é um vetor arbitrário em V e f um polinômio tal que $f(T)\beta$ está em W, então existe um vetor α em W tal que $f(T)\alpha = f(T)\beta$. Demonstraremos abaixo que esta condição sôbre W também é suficiente para a existência de um subespaço T-invariante suplementar. Em outras palavras, os subespaços invariantes que possuem um subespaço suplementar invariante serão os subespaços que satisfazem a propriedade que acabamos de descrever. Vamos agora enunciar formalmente o resultado e dividir a demonstração da suficiência numa seqüência de lemas.

Teorema 3. *Seja* T *um operador linear sôbre o espaço vetorial* V *de dimensão finita sôbre o corpo* F *e seja* W *um subespaço de* V *que seja invariante sob* T. *Uma condição necessária e suficiente para que* W *admita um subespaço suplementar* T-*invariante é a seguinte: Se* β *é um vetor em* V *e* f *um polinômio sôbre* F *tal que* $f(T)\,\beta$ *esteja em* W, *existe um vetor* α *em* W *tal que* $f(T)\alpha = f(T)\beta$.

Para o restante desta seção, denominaremos W um **subespaço T-admissível** se

(i) W é invariante sob T;

(ii) se $f(T)\beta$ está em W, existe um α em W tal que $f(T)\alpha = f(T)\beta$.

Isto nos poupará algumas palavras na demonstração do Teorema 3.

Lema 1. *Seja* W *um subespaço* T-*invariante de* V *e seja* β *um vetor qualquer em* V. *Seja* $S(\beta; W)$ *o conjunto dos polinômios* f *sôbre o corpo* F *de escalares tais que* $f(T)\beta$ *esteja em* W. *Então*

(a) $S(\beta; W)$ *é um ideal não-nulo em* F[x];

(b) *se* α *é um vetor arbitrário em* W *então* $S(\beta - \alpha; W) = S(\beta; W)$;

(c) $Z(\beta; T)$ *é disjunto de* W *se e sòmente se* $S(\beta; W)$ *é* T-*anulador de* β, *isto é, se, e sòmente se, o gerador unitário de* $S(\beta; W)$ *é o* T-*anulador de* β.

Demonstração. Usaremos a notação $f\beta$ para $f(T)\beta$. (a) Se f e g estão em $S(\beta; W)$, então $f\beta$ e $g\beta$ estão em W, portanto $(f + g)\beta = f\beta + g\beta$ está em W, isto é, $(f + g)$ está em $S(\beta; W)$. Se f está em $S(\beta; W)$ e h é um polinômio qualquer sôbre F, então $(hf)\beta = h(f\beta)$. Como $f\beta$ está em W e W é invariante sob T, vemos que hf está em $S(\beta; W)$. O conjunto $S(\beta; W)$ é não-vazio e não-nulo, pois contém o polinômio minimal de T. (b) Se α está em W e f é um polinômio arbitrário, então $f(\beta - \alpha) = f\beta - f\alpha$, e como $f\alpha$ está em W vemos que $f(\beta - \alpha)$ está em W se, e sòmente se, $f\beta$ está em W. (c) A interseção de $Z(\beta; T)$ com W é simplesmente o conjunto dos vetores em W da forma $f\beta$, com f em $F[x]$. Dizer que esta interseção é o subespaço nulo é o mesmo que dizer que $f\beta$ está em W se, e sòmente se, $f\beta = 0$ e isto é evidentemente a afirmação de que $S(\beta; W)$ é o T-anulador de β.

Lema 2. *Seja* W *um subespaço próprio de* V *que seja* T-*admissível. Então existe um vetor não-nulo* α_1 *em* V *tal que o subespaço cíclico* $Z(\alpha_1; T)$ *seja disjunto de* W.

Demonstração. Tomemos um vetor β em V que não esteja em W. Seja p o gerador unitário do ideal $S(B; W)$, isto é, $f\beta$ está em W se, e sòmente se, p divide f. Como $p\,\beta$ está em W e W é T-admissível, existe um vetor α em W tal que $p\,\beta = p\alpha$. Seja $\alpha_1 = \beta - \alpha$. Pelo Lema 1, $S(\alpha_1; W) = S(\beta; W)$; isto é, $f\alpha_1$ está em W se, e sòmente se, p divide f. Mas $p\alpha_1 = p(\beta - \alpha) = 0$. Assim, $f\alpha_1$ está em W se, e sòmente se, $f\alpha_1 = 0$, ou seja $Z(\alpha_1; T)$ é disjunto de W. Ora, $\alpha_1 \neq 0$ pois α está em W e β não está em W.

Agora teremos o passo crucial:

Lema 3. *Seja* W *um subespaço* T-*admissível de* V *e seja* α_1 *um vetor em* V *tal que*

(1) $Z(\alpha_1; T)$ *é disjunto de* W;

(2) $Z(\alpha_1; T)$ *possui dimensão máxima entre todos os subespaços* T-*cíclicos disjuntos de* W, *isto é,* $W \cap Z(\beta; T) = \{0\}$ *implica* $\dim Z(\beta; T) \leq \dim Z(\alpha_1; T)$.

Então o subespaço $W \oplus Z(\alpha_1; T)$ *é* T-*admissível.*

Demonstração. Abreviaremos a notação de $Z(\alpha_1; T)$ para $Z\alpha_1$. O que desejamos demonstrar é: Se β é um vetor qualquer em V e f é um polinômio tal que $f\beta$ esteja em $W \oplus Z\alpha_1$, então existe um vetor α em $W \oplus Z\alpha_1$ tal que $f\beta = f\alpha$. Fizemos β em V e consideremos o ideal $S(\beta; W \oplus Z\alpha_1)$ dos f tais que $f\beta$ esteja em $W \oplus Z\alpha_1$. Ora, $S(\beta; W \oplus Z\alpha_1) = pF[x]$ para um certo polinômio unitário p. Suponhamos que seja possível demonstrar a existência de um vetor α em $W \oplus Z\alpha_1$ tal que $p\beta = p\alpha$. Então êste mesmo vetor α servirá para qualquer f em $S(\beta; W \oplus Z\alpha_1)$, isto é, se $f\beta$ está em $W \oplus Z\alpha_1$, então $f\beta = f\alpha$. De fato, se $f\beta$ está em $W \oplus Z\alpha_1$, p divide f e se $f = f_1p$, temos $f\beta = f_1p\beta = f_1p\alpha = f\alpha$.

Pela definição de p, o vetor $p\beta$ está no espaço $W \oplus Z\alpha_1$, ou seja, $p\beta = \gamma + g\alpha_1$ onde γ está em W e g é um polinômio. Temos $g = pq + r$ onde $r = 0$ ou $\text{gr}(r) < \text{gr}(p)$. Vamos demonstrar que $r = 0$, mas antes completaremos a demonstração supondo $r = 0$. Ora, $r = 0$ diz que $p\beta = \gamma + pq\alpha_1$, logo $p(\beta - q\alpha_1) = \gamma$ está em W. Como W é T-admissível, existe um vetor γ_0 em W tal que $p(\beta - q\alpha_1) = p\gamma_0$, isto é, $p\beta = p\gamma_0 + pq\alpha_1$. Se $\alpha = \gamma_0 + q\alpha_1$, então α está em $W \oplus Z\alpha_1$ e $p\beta = p\alpha$.

Voltamos agora ao trabalho que nos resta, a saber, demonstrar que $r = 0$. É neste ponto que entra a hipótese de $Z\alpha_1$ ter dimensão máxima entre todos os subespaços T-cíclicos disjuntos de W. Temos $p\beta = \gamma + pq\alpha_1 + r\alpha_1$, ou seja, $p\beta - pq\alpha_1 = \gamma + r\alpha_1$. Seja $\beta' = \beta - q\alpha_1$. Então

$$S(\beta'; W \oplus Z\alpha_1) = S(\beta; W \oplus Z\alpha_1) = pF[x]$$

e

$$p\beta' = \gamma + r\alpha_1.$$

Consideremos o ideal $S(\beta'; W)$ dos f tais que $f\beta'$ esteja em W. Quando $f\beta'$ está em W temos certamente $f\beta'$ em $W \oplus Z\alpha_1$; conseqüentemente, p divide todos êstes f. Portanto, $S(\beta'; W) = phF[x]$ para algum polinômio h, isto é, o gerador de $S(\beta'; W)$ é divisível por p. Ora,

$$\begin{aligned} ph\beta' &= h(p\beta') \\ &= h(\gamma + r\alpha_1) \\ &= h\gamma + hr\alpha_1. \end{aligned}$$

Pela definição de h, o vetor $ph\beta' = h\gamma + hr\alpha_1$ está em W e como W e $Z\alpha_1$ são disjuntos devemos ter $hr\alpha_1 = 0$. Daqui tiramos duas conclusões.

(i) Como W é T-admissível e $ph\beta'$ está em W, existe um vetor α' em W tal que $ph\beta' = ph\gamma'$. (ii) Como $hr\alpha_1 = 0$, o T-anulador p_1 de α_1 divide hr; em particular, se $r \neq 0$

$$\text{gr}(p_1) \leq \text{gr}(hr) = \text{gr}(h) + \text{gr}(r) < \text{gr}(h) + \text{gr}(p) = \text{gr}(ph).$$

De (i) temos $ph\beta' = ph\gamma'$, em W. Seja $\beta'' = \beta' - \gamma'$; então $S(\beta''; W) = S(\beta'; W) = phF[x]$ e $ph\beta'' = 0$. Assim, o subespaço cíclico $Z(\beta''; T)$ é disjunto de W e ph é o T-anulador de β''. Pela hipótese (2) dêste lema devemos ter então

$$\operatorname{gr}(ph) = \dim Z(\beta''; T) \leq \dim Z(\alpha_1; T) = \operatorname{gr}(p_1).$$

Em (ii) acima, observamos que se $r \neq 0$ temos $\operatorname{gr}(p_1) < \operatorname{gr}(ph)$. Decorre que $r = 0$.

Resumiremos os resultados dos Lemas 2 e 3 e acrescentaremos um fato adicional obtido na demonstração do Lema 3.

Lema 4. *Seja* W *um subespaço próprio de* V *que seja* T-*admissível. Então, existe um vetor não-nulo* α_1 *em* V *tal que* $Z(\alpha_1; T)$ *seja disjunto de* W. *Se* α_1 *foi escolhido de modo que* $Z(\alpha_1; T)$ *tenha dimensão máxima entre todos os subespaços* T-*cíclicos disjuntos de* W, *temos também*

(a) *o subespaço* $W \oplus Z(\alpha_1; T)$ *é* T-*admissível*;

(b) *se* β *é um vetor arbitrário tal que* $Z(\beta; T)$ *seja disjunto de* $W \oplus Z(\alpha_1; T)$, *então o* T-*anulador de* β *divide o* T-*anulador de* α_1.

Demonstração. Temos que demonstrar apenas a afirmação (b). Se $Z(\beta; T)$ é disjunto de $W \oplus Z\alpha_1$, então

$$S(\beta; W + Z\alpha_1) = pF[x]$$

onde p é o T-anulador de β. Assim, $p\beta = 0$. Se p_1 é o T-anulador de α_1, então $p_1\alpha_1 = 0$, portanto

$$p\beta = 0 + p_1\alpha_1.$$

Na demonstração do Lema 3 podemos tomar $\gamma = 0$ e $g = p_1$. Como mostramos naquela demonstração, p deve dividir p_1.

O Teorema 3 decorre fàcilmente do Lema 4. Suponhamos que W seja um subespaço T-admissível de V e que desejamos encontrar um subespaço T-invariante W' com $V = W \oplus W'$. Se $W = V$, tomamos $W' = \{0\}$. Se W é um subespaço próprio de V, tomemos um vetor não-nulo α_1 em V tal que $W \oplus Z\alpha_1$ seja T-admissível. Se $W \oplus Z\alpha_1 = V$, tomando $W' = Z\alpha_1$ a demonstração se completa. Em caso contrário, apliquemos novamente o Lema 4 e determinemos um α_2 não-nulo em V tal que $W \oplus Z\alpha_1 \oplus Z\alpha_2$ seja T-admissível. Se êste espaço esgota V, tomamos $W' = Z\alpha_1 \oplus Z\alpha_2$.

Em caso contrário, prosseguimos. O fato fundamental aqui é êste: por meio de um número finito de aplicações do Lema 4 devemos obter

$$W \oplus Z\alpha_1 \oplus \dots \oplus Z\alpha_r = V$$

pois cada vez que acrescentamos um vetor *não-nulo* α_j, a dimensão do subespaço admissível respectivo aumenta de pelo menos 1.

O problema original desta seção, o de decompor V numa soma direta de subespaços T-cíclicos, também é resolvido pelo Lema 4.

Teorema 4. (Teorema da Decomposição Racional). *Seja* T *um operador linear sôbre o espaço vetorial* V *de dimensão finita* (*dim* V $\geq$ 1). *Então existem* r *vetores não-nulos* $\alpha_1, \ldots, \alpha_r$ *em* V *com os respectivos* T-*anuladores* $p_1, \ldots, p_r$ *tais que*

(a) $V = Z(\alpha_1; T) \oplus \ldots \oplus Z(\alpha_r; T)$;

(b) *se* $1 \leq k \leq r - 1$, *então* p_{k+1} *divide* p_k.

Além disso, o inteiro r *e os anuladores* $p_1, \ldots, p_r$ *são determinados de modo único pelas condições* (*a*) *e* (*b*) *mais o fato de que nenhum* α_k *é nulo.*

Demonstração. Demonstraremos primeiro que existem vetores não-nulos $\alpha_1, \ldots, \alpha_r$ com T-anuladores $p_1, \ldots, p_r$ que satisfazem as condições (a) e (b). Esta demonstração é essencialmente uma repetição da demonstração que acabamos de fazer para o Teorema 3, começando com o subespaço T-admissível $W = \{0\}$. Neste caso, usamos a afirmação (b) do Lema 4 para obter a condição de divisibilidade: p_{k+1} divide p_k. Assim, a demonstração é feita ràpidamente da seguinte maneira: Seja α_1 um vetor não-nulo arbitrário em V que gere um subespaço T-cíclico de dimensão máxima. Como $W = \{0\}$ é trivialmente T-admissível, o Lema 4 afirma que $Z(\alpha_1; T)$ é admissível. Se $V = Z(\alpha_1; T)$, a demonstração está completa. Se $V \neq Z(\alpha_1; T)$, tomemos um α_2 não-nulo tal que $Z(\alpha_2; T)$ tenha dimensão máxima entre os espaços T-cíclicos que são disjuntos de $Z(\alpha_1; T)$. Pelo Lema 4, p_2 divide p_1. Se $V = Z(\alpha_1; T) \oplus Z(\alpha_2; T)$, ótimo. Se não, usamos o Lema 4 para escolher um α_3 conveniente, etc.

Suponhamos que $\alpha_1, \ldots, \alpha_r$ sejam vetores não-nulos satisfazendo (a) e (b). Suponhamos que também existem vetores não-nulos $\beta_1, \ldots, \beta_s$ com T-anuladores $g_1, \ldots, g_s$ tais que

(a) $V = Z(\beta_1; T) \oplus \ldots \oplus Z(\beta_s; T)$;

(b) g_{k+1} divide g_k para $k = 1, \ldots, s - 1$.

Vamos demonstrar que $r = s$ e $p_j = g_j$, $j = 1, \ldots, r$.

Se f é um polinômio sôbre F e W é um subespaço de V, indiquemos por fW o espaço dos vetores $f\alpha$ com α em W. Assim, fV é a imagem do operador $f(T)$. Vamos usar agora dois fatos, cujas demonstrações deixamos como exercícios. Primeiro, se $V = Z(\alpha_1; T) \oplus \ldots \oplus Z(\alpha_r; T)$ e f é um polinômio arbitrário sôbre F, então

$$fV = fZ(\alpha_1; T) \oplus \ldots \oplus fZ(\alpha_r; T) = Z(f\alpha_1; T) \oplus \ldots \oplus Z(f\alpha_r; T).$$

Segundo, se α e β são vetores em V que tenham o mesmo T-anulador, então $f\alpha$ e $f\beta$ têm o mesmo T-anulador; em particular $Z(f\alpha; T)$ e $Z(f\beta; T)$ têm a mesma dimensão.

Voltando a $\alpha_1, \ldots, \alpha_r$; e $\beta_1, \ldots, \beta_s$, demonstremos que $p_1 = g_1$. Temos

$$p_1V = p_1Z(\alpha_1; T) \oplus \ldots \oplus p_1Z(\alpha_r; T)$$
$$p_1V = p_1Z(\beta_1; T) \oplus \ldots \oplus p_1Z(\beta_s; T).$$

Ora, $p_1\alpha_j = 0$ para $j = 1, \ldots, r$, pois p_j divide p_1. Assim, $p_1V = \{0\}$. Portanto, $p_1\beta_j = 0$ para $j = 1, \ldots, s$. Em particular, $p_1\beta_1 = 0$, logo g_1 divide p_1. Invertendo o raciocínio, isto é, considerando g_1V, vemos que p_1 divide g_1. Portanto $p_1 = g_1$.

O fato de que $r = s$ e $p_j = g_j$ podem ser demonstrados por indução sôbre j. Demonstraremos a partir de $p_1 = g_1$ que $p_2 = g_2$ e, a partir disto, a indução deverá ser evidente. A demonstração será feita envolvendo principalmente dimensões. Suponhamos que $r \geq 2$, isto é, que realmente exista um vetor α_2. Então teremos forçosamente $s \geq 2$, pois

$$\dim V = \dim Z(\alpha_1; T) + \ldots + \dim Z(\alpha_r; T)$$
$$\dim V = \dim Z(\beta_1; T) + \ldots + \dim Z(\beta_s; T)$$

e como $\dim Z(\alpha_1; T) + \text{gr}(p_1) = \text{gr}(g_1) = Z(\beta_1; T)$; vemos que

$$\sum_{j \geq 2} \dim Z(\beta_j; T) = \sum_{j \geq 2} \dim Z(\alpha_j; T) > 0.$$

Consideramos agora o subespaço p_2V. Temos duas descrições deste espaço,

$$p_2V = Z(p_2\alpha_1; T) \oplus \ldots \oplus Z(p_2\alpha_r; T)$$
$$p_2V = Z(p_2\beta_1; T) \oplus \ldots \oplus Z(p_2\beta_s; T).$$

Como $p_2\alpha j = 0$ para $j \geq 2$

$$p_2V = Z(p_2\alpha_1; T).$$

Os vetores α_1 e β_1 possuem o mesmo T-anulador, logo $p_2\alpha_1$ e $p_2\beta_1$ possuem o mesmo T-anulador. Portanto

$$\dim (p_2V) = \dim Z(p_2\alpha_1; T) = \dim Z(p_2\beta_1; T)$$

donde segue que $\dim Z(p_2\beta_j; T) = 0$ para $j \geq 2$. Em particular,

$p_2\beta_2 = 0$ e assim g_2 divide p_2. Invertendo o argumento, p_2 divide g_2, e então $p_2 = g_2$.

Gostaríamos de ressaltar que nossa demonstração do Teorema 4 mostra um pouco mais do que o teorema afirma. Suponhamos que existem vetores não-nulos $\alpha_1, \ldots, \alpha_t$ em V que satisfaçam as condições:

(i) $Z(\alpha_1; T)$ tem dimensão máxima entre os subespaços T-cíclicos.

(ii) Para $j \geq 2$, o subespaço $Z(\alpha_j; T)$ tem dimensão máxima entre todos os subespaços T-cíclicos que são disjuntos de $Z(\alpha_1; T) + \ldots + Z(\alpha_{j-1}; T)$.

Então, existem vetores não-nulos $\alpha_{t+1}, \ldots, \alpha_r$ tais que $\alpha_1, \ldots, \alpha_r$ satisfaçam as condições (a) e (b) do Teorema 4.

Gostaríamos também de destacar que se $\alpha_1, \ldots, \alpha_r$ satisfazem as condições do Teorema 4, então o T-anulador de α_1 é o polinômio minimal de T. De fato, como p_j divide p_1 para $j = 1, \ldots, r$, temos $p_1 V = \{0\}$, isto é, $p_1(T) = 0$. Conseqüentemente, p_1 é divisível pelo polinômio minimal de T. Mas p_1 é o T-anulador de α_1, portanto divide o polinômio minimal. Assim, se se pergunta quais vetores α_1 poderiam ser o primeiro vetor no Teorema 4, a resposta é os vetores cujo T-anulador seja o polinômio minimal.

Suponhamos que existem T, r, α_i e p_i como no teorema; indiquemos por T_i o operador linear sôbre $Z(\alpha_i; T)$ obtido pela restrição de T àquele subespaço. Então T_i possui um vetor cíclico, a saber, α_i e, além disso, os polinômios característicos e minimal de T_i são ambos iguais a p_i. Assim, o polinômio característico de T é

$$f = p_1 p_2 \ldots p_r. \tag{7-3}$$

(Nunca demonstramos explìcitamente que se T é a soma direta dos operadores $T_1, \ldots, T_r$, então o polinômio característico de T é o produto dos polinômios característicos dos T_i; no entanto, o leitor deve observar que isto é evidente a partir do resultado correspondente sôbre somas diretas de matrizes e que o último fato decorre da regra da "forma em blocos" para determinantes de matrizes.)

Corolário 1. *Se* T *é um operador linear sôbre o espaço* V *de dimensão finita (ou, se A é uma* n × n *matriz sôbre o corpo* F*), então os fatôres primos do polinômio minimal de* T *(de* A*) são os mesmos que os fatôres primos do polinômio característico de* T *(de* A*).*

Demonstração. De acôrdo com a expressão (7-3) para o polinômio característico, f é um produto de polinômios, cada um dos quais divide o polinômio minimal p_1. Assim, os fatôres primos de f e p_1 são os mesmos, apesar de que êles possam se repetir mais vêzes em f.

Corolário 2. *Seja* T *um operador linear sôbre o espaço* V *de dimensão finita. Então* T *possui um vetor cíclico se, e sòmente se, os polinômios característico e minimal de* T *são idênticos.*

Demonstração. De acôrdo com nossa demonstração do Teorema 4, se α é vetor em V a máxima dimensão que o espaço $Z(\alpha; T)$ pode ter é o grau do polinônio minimal de T; além disso, existe um vetor α tal que $Z(\alpha; T)$ possui esta dimensão. Assim T possui um vetor cíclico se, e sòmente se, os polinômios característico e minimal de T têm o mesmo grau, isto é, se, e sòmente se, êles são idênticos.

Observamos agora o análogo matricial do teorema de decomposição racional. Se temos o operador T e a decomposição em soma direta do Teorema 4, seja $\mathfrak{B}_i$ a "base ordenada cíclica"

$$\{\alpha_i, T\alpha_i, \ldots, T^{k_i-1}\alpha_i\}$$

de $Z(\alpha_i; T)$. Aqui k_i indica a dimensão de $Z(\alpha_i; T)$, ou seja, o grau do anulador p_i. A matriz do operador induzido T_i em relação à base ordenada $\mathfrak{B}_i$ é a matriz associada ao polinômio p_i. Assim, se $\mathfrak{B}$ é a base ordenada de V obtida pela reunião das $\mathfrak{B}_i$, ordenadas como $\mathfrak{B}_1, \ldots, \mathfrak{B}_r$, então a matriz de T em relação à base ordenada $\mathfrak{B}$ será

$$(7\text{-}4) \qquad A = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A_r \end{bmatrix}$$

onde A_i é a $k_i \times k_i$ matriz associada a p_i. Diremos que uma $n \times n$ matriz A, que seja a soma direta (7-4) das matrizes associadas a polinômios não-constantes e unitários $p_1, \ldots, p_r$ tais que p_{i+1} divide p_i para $i = 1, \ldots, r - 1$, está sob a **forma racional**. O teorema da decomposição racional nos diz o seguinte em relação a matrizes:

Teorema 5. *Seja* F *um corpo e seja* B *uma* n $\times$ n *matriz sôbre* F. *Então* B *é semelhante sôbre o corpo* F *a uma e sòmente a uma matriz sob a forma racional.*

Demonstração. Seja T o operador linear sôbre F^n que é representado por B em relação à base ordenada canônica. Como acabamos de observar, existe alguma base ordenada de F^n em relação à qual T é representado por uma matriz A sob a forma racional. Então B é semelhante a esta matriz A. Suponhamos que B seja semelhante sôbre F a uma outra matriz C que esteja sob a forma racional. Isto significa simplesmente que existe alguma base ordenada de F^n em relação à qual o operador T é representado pela matriz C. Se C é a soma direta das matrizes C_i associadas a polinômios unitários $g_1, \ldots, g_s$ tais que g_{i+1} divida g_i para $i = 1, \ldots, s - 1$ então é evidente que teremos vetores não-nulos $\beta_1, \ldots, \beta_s$ em V com T-anuladores $g_1 \ldots, g_s$ tais que

$$V = Z(\beta_1; T) \oplus \ldots \oplus Z(\beta_s; T).$$

Mas então, pela afirmação da unicidade no teorema da decomposição racional, os polinômios g_i são idênticos aos polinômios p_i que definem a matriz A. Assim, $C = A$.

Exemplo 2. Suponhamos que V seja um espaço vetorial bidimensional sôbre o corpo F e T seja um operador linear sôbre V. As possibilidades para a decomposição de T em subespaços cíclicos são bastante limitadas. De fato, se o polinômio minimal de T tem grau 2, êle é igual ao polinômio característico de T e T possui um vetor cíclico. Assim, existe uma base ordenada de V em relação à qual T é representado pela matriz associada ao seu polinômio característico. Se, por outro lado, o polinômio minimal de T tem grau 1, então T é um múltiplo escalar do operador idêntico. Se $T = cI$, então para dois quaisquer vetores linearmente independentes α_1 e α_2 em V, temos

$$V = Z(\alpha_1; T) \oplus Z(\alpha_2; T)$$
$$p_1 = p_2 = x - c.$$

Para matrizes, esta análise diz que tôda 2×2 matriz sôbre o corpo F é semelhante sôbre F a exatamente uma matriz dos tipos

$$\begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix} \quad \begin{bmatrix} 0 & -c_0 \\ 1 & -c_1 \end{bmatrix}.$$

Exemplo 3. Seja T o operador linear sôbre R^3 que é representado pela matriz

$$A = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$$

em relação à base ordenada canônica. Como já calculamos anteriormente, o polinômio característico de T é $f = (x - 1)(x - 2)^2$ e o polinômio minimal de T é $p = (x - 1)(x - 2)$. Assim, sabemos que na decomposição cíclica de T o T-anulador do primeiro vetor α_1 será p. Como estamos operando num espaço tridimensional, só pode existir mais um vetor, α_2. Êste vetor deve gerar um subespaço cíclico de dimensão 1, isto é, deve ser um vetor característico de T. O seu T-anulador p_2 deve ser $(x - 2)$, porque devemos ter $pp_2 = f$. Notemos que isto nos diz imediatamente que a matriz A é semelhante à matriz

$$B = \begin{bmatrix} 0 & -2 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

ou seja, que T é representado por B em relação a uma certa base ordenada. Como podemos encontrar vetores adequados α_1 e α_2? Bem, sabemos que todo vetor que gera um subespaço T-cíclico de dimensão 2 é um α_1 adequado. Portanto experimentamos ϵ_1. Temos

$$T\epsilon_1 = (5, -1, 3)$$

que não é múltiplo escalar de ϵ_1; logo $Z(\alpha_2; T)$ tem dimensão 2. Êste espaço consiste dos vetores $a\,\epsilon_1 + b(T\epsilon_1)$:

$$a(1, 0, 0) + b(5, -1, 3) = (a + 5b, -b, 3b)$$

ou seja, todos os vetores (x_1, x_2, x_3) que satisfazem $x_3 = -3x_2$. O que desejamos agora é um vetor α_2 tal que $T\alpha_2 = 2\alpha_2$ e $Z(\alpha_2; T)$ seja disjunto de $Z(\epsilon_1; T)$. Como α_2 deve ser um vetor característico de T, o espaço $Z(\alpha_2; T)$ será simplesmente o espaço unidimensional gerado por α_2 e então o que exigimos é que α_2 não esteja em $Z(\epsilon_1; T)$. Se $\alpha = (x_1, x_2, x_3)$, pode-se calcular fàcilmente que $T\alpha = 2\alpha$ se, e sòmente se, $x_1 = 2x_2 + 2x_3$. Assim, $\alpha_2 = (2, 1, 0)$ satisfaz $T\alpha_2 = 2\alpha_2$ e gera um subespaço T-cíclico disjunto de $Z(\epsilon_1; T)$. O leitor deverá verificar diretamente que a matriz de T em relação à base ordenada

$$\{(1, 0, 0), (5, -1, 3), (2, 1, 0)\}$$

é a matriz B acima.

Exercícios

1. Seja T o operador linear sôbre F^2 que é representado em relação à base ordenada canônica pela matriz

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Seja $\alpha_1 = (0, 1)$. Mostrar que $F^2 \neq Z(\alpha_1; T)$ e que não existe nenhum vetor não-nulo α_2 em F^2 tal que $Z(\alpha_2; T)$ seja disjunto de $Z(\alpha_1; T)$.

2. Seja T um operador linear sôbre o espaço V de dimensão finita e seja R a imagem de T.

(a) Demonstrar que R possui um subespaço suplementar T-invariante se, e sòmente se, R é disjunto do núcleo N de T.

(b) Se R e N são disjuntos, demonstrar que N é o único subespaço T-invariante que é um suplementar de R.

3. Seja T o operador linear sòbre R^3 que é representado em relação à base ordenada canônica pela matriz

$$\begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Seja W o núcleo de $T - 2I$. Demonstrar que W não possui nenhum subespaço suplementar T-invariante. [*Sugestão*: Seja $\beta = \epsilon_1$ e observemos que $(T - 2I)\beta$ está em W. Demonstrar que não existe nenhum α em W tal que $(T - 2I)\beta = (T - 2I)\alpha$.]

4. Seja T o operador linear sôbre F^4 que é representado em relação à base ordenada canônica pela matriz

$$\begin{bmatrix} c & 0 & 0 & 0 \\ 1 & c & 0 & 0 \\ 0 & 1 & c & 0 \\ 0 & 0 & 1 & c \end{bmatrix}.$$

Seja W o núcleo de $T - cI$.

(a) Demonstrar que W é o subespaço gerado por ϵ_4.

(b) Determinar os geradores unitários dos ideais $S(\epsilon_4; W)$, $S(\epsilon_3; W)$, $S(\epsilon_2; W)$, $S(\epsilon_1; W)$.

5. Seja T um operador linear sôbre o espaço vetorial V sôbre o corpo F. Se f é um polinômio sôbre F e α está em V, seja $f\alpha = f(T)\alpha$. Suponhamos que $\alpha_1, \ldots, \alpha_r$ sejam vetores em V tais que $V = Z(\alpha_1; T) \oplus \ldots \oplus Z(\alpha_r; T)$. Demonstrar que

$$fV = fZ(\alpha_1; T) \oplus \ldots \oplus fZ(\alpha_r; T) = Z(f\alpha; T) \oplus \ldots \oplus Z(f\alpha)_r; T$$

(fV é o conjunto dos $f\alpha$, α em V, etc.).

6. Sejam T, V e F como no Exercício 5. Suponhamos que α e β sejam vetores em V que o mesmo T-anulador. Demonstrar que, para todo polinômio f, os vetores $f\alpha$ e $f\beta$ têm o mesmo T-anulador.

7. Determinar os polinômios minimais e as formas racionais de cada uma das seguintes matrizes reais:

$$\begin{bmatrix} 0 & -1 & -1 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} c & 0 & -1 \\ 0 & c & 1 \\ -1 & 1 & c \end{bmatrix} \qquad \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

8. Seja T o operador linear sôbre R^3 que é representado em relação à base ordenada canônica por

$$\begin{bmatrix} 3 & -4 & -4 \\ -1 & 3 & 2 \\ 2 & -4 & -3 \end{bmatrix}.$$

Determinar vetores não-nulos $\alpha_1, \ldots, \alpha_r$ que satisfaçam as condições do Teorema 4.

9. Seja A a matriz real

$$A = \begin{bmatrix} 1 & 3 & 3 \\ 3 & 1 & 3 \\ -3 & -3 & -5 \end{bmatrix}.$$

Determinar uma 3×3 matriz real inversível P tal que $P^{-1}AP$ esteja sob a forma racional.

10. Seja F um subcorpo do corpo dos números complexos e seja T o operador linear sôbre F^4 que é representado em relação à base ordenada canônica pela matriz

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & a & 2 & 0 \\ 0 & 0 & b & 2 \end{bmatrix}.$$

Determinar o polinômio característico de T. Considerar os casos $a = b = 1$; $a = b = 0$; $a = 0$, $b = 1$. Em cada um dêstes casos, determinar o polinômio minimal de T e vetores não-nulos $\alpha_1, \ldots, \alpha_r$ que satisfaçam as condições do Teorema 4.

11. Demonstrar que se A e B são 3×3 matrizes sôbre o corpo F, uma condição necessária e suficiente para que A e B sejam semelhantes sôbre F é que possuam o mesmo polinômio característico e o mesmo polinômio minimal. Dar um exemplo que mostra que isto é falso para 4×4 matrizes.

12. Seja F um subcorpo do corpo dos números complexos e sejam A e B $n \times n$ matrizes sôbre F. Demonstrar que se A e B são semelhantes sôbre o corpo dos números complexos, então elas são semelhantes sôbre F. (*Sugestão*: Demonstrar que a forma racional de A é a mesma seja A considerada como uma matriz sôbre F ou como uma matriz sôbre C; o mesmo para B.)

13. Seja A uma $n \times n$ matriz com elementos complexos. Demonstrar que se todo valor característico de A é real, então A é semelhante a uma matriz com elementos reais.

14. Seja T um operador linear sôbre o espaço V de dimensão finita. Demonstrar que existe um vetor α em V com a seguinte propriedade: Se f é um polinômio e $f(T)\alpha = 0$, então $f(T) = 0$. (Um tal vetor α é denominado um **vetor separador** para álgebra dos polinômios em T.) Para o caso em que T possui um vetor cíclico, demonstrar diretamente que todo vetor cíclico é um vetor separador para a álgebra dos polinômios em T.

15. Seja F um subcorpo do corpo dos números complexos e seja A uma $n \times n$ matriz sôbre F. Seja p o polinômio minimal de A. Se considerarmos A como uma matriz sôbre C, então A possuirá um polinômio minimal f, quando considerada como uma $n \times n$ matriz sôbre C. Usar um teorema sôbre equações lineares para demonstrar que $p = f$. De que forma êste resultado decorre do teorema da decomposição racional?

16. Seja A uma $n \times n$ matriz com elementos *reais* tal que $A^2 + I = 0$. Demonstrar que n é par e que, se $n = 2k$, então A é semelhante sôbre o corpo dos números reais a uma matriz da forma em blocos

$$\begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}$$

onde I é a $k \times k$ matriz unidade.

17. Seja V um espaço vetorial de dimensão finita e seja T um operador linear sôbre V com polinômio minimal p. Se

$$p = f_1^{d_1} \ldots . f_k^{r_k}$$

é a decomposição de p em fatôres primos, demonstrar que o polinômio característico de T é

$$f = f_k^{d_1} \ldots f_k^{d_k}$$

onde d_i é a nulidade de $f_i(T)^{r_i}$ dividida pelo grau de f_i.

18. Seja T um operador linear sôbre o espaço V de dimensão finita. Demonstrar que T possui um vetor cíclico se, e sòmente se, vale o seguinte: Todo operador linear U que comuta com T é um polinômio em T.

19. Seja V um espaço vetorial de dimensão finita sôbre F e seja T um operador linear sôbre V. Perguntamos quando é que todo vetor não-nulo em V é um vetor cíclico de T. Demonstrar que isto ocorre, se, e sòmente se, o polinômio característico de T é irredutível sôbre F.

20. Seja A uma $n \times n$ matriz com elementos *reais*. Seja T o operador linear sôbre R^n que é representado por A em relação à base ordenada canônica e seja U o operador linear sôbre C^n que é representado por A em relação à base ordenada canônica. Usar o resultado do Exercício 19 para demonstrar o seguinte: Se os únicos subespaços invariantes sob T são R^n e o subespaço nulo, então U é diagonalizável.

7.3 A Forma de Jordan

Suponhamos que N seja um operador linear nilpotente sôbre o espaço V de dimensão finita. Consideremos a decomposição cíclica de N que obtemos por meio do teorema da decomposição racional. Temos um inteiro positivo r e r vetores não-nulos $\alpha_1, \ldots, \alpha_r$ em V com N-anuladores $p_1, \ldots, p_r$, tais que

$$V = Z(\alpha_1; N) \oplus \ldots \oplus Z(\alpha_r; N)$$

e p_{i+1} divide p_i para $i = 1, \ldots, r-1$. Como N é nilpotente, o polinômio característico de N é x^n. O polinômio é x^k para um certo $k < n$. Assim, cada p_i é da forma $p_i = x^{k_i}$ e a condição de divisibilidade diz simplesmente que

$$k_1 \geq k_2 \geq \ldots \geq k_r.$$

Evidentemente, $k_r = k$ e $k_r \geq 1$. A matriz associada a x^{k_i} é a $k_i \times k_i$ matriz

$$(7\text{-}5) \qquad Ai = \begin{bmatrix} 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & \ldots & 1 & 0 \end{bmatrix}.$$

Assim, o Teorema 4 nos fornece uma base ordenada de V em relação à qual a matriz de N é a soma direta das matrizes nilpotentes elementares (7-5), cujas dimensões diminuem à medida que i aumenta. Vê-se, a partir disto, que estão associados a uma $n \times n$ matriz nilpotente um inteiro positivo r e r inteiros positivos $k_1, \ldots, k_r$ tais que $k_1 + \ldots + k_r = n$ e $k_i \geq k_{i+1}$, e êstes inteiros positivos determinam a forma racional da matriz, isto é, determinam a matriz a menos de semelhança.

Eis aqui algo que gostaríamos de ressaltar sôbre o operador nilpotente N acima. O inteiro positivo r é precisamente a nulidade de N; na verdade, uma base do núcleo é formada pelos r vetores

$$N^{k_i - 1}\alpha_i. \tag{7-6}$$

De fato, suponhamos que α esteja no núcleo de N. Podemos escrever α sob a forma

$$\alpha = f_1\alpha_1 + \ldots + f_r\alpha_r$$

onde f_i é polinômio, cujo grau podemos supor menor que k_i. Como $N\alpha = 0$, para cada i temos

$$\begin{aligned} 0 &= N(f_i\alpha_i) \\ &= Nf_i(N)\alpha_i \\ &= (xf_i)\alpha_i. \end{aligned}$$

Assim, xf_i é divisível por x^{k_i} e como $\text{gr}(f_i) > k_i$ isto significa que

$$f_i = c_i x^{k_i - 1}$$

onde c_i é um certo escalar. Mas então

$$\alpha = c_1(x^{k_1 - 1}\alpha_1) + \ldots + c_r(x^{k_r - 1}\alpha_r)$$

o que nos mostra que os vetores (7-6) formam uma base do núcleo de N. O leitor deverá notar que êste fato também é evidente do ponto de vista de matrizes.

O que desejamos fazer agora é combinar nossas conclusões a respeito de operadores ou matrizes nilpotentes com o teorema da decomposição primária do Capítulo 6. A situação é a seguinte: Suponhamos que T seja um operador linear sôbre V e que o polinômio característico de T se decomponha sôbre F como segue:

$$f = (x - c_1)^{d_1} \ldots (x - c_k)^{d_k}$$

onde $c_1, \ldots, c_k$ são elementos distintos em F e $d_i \geq 1$. Então, o polinômio minimal de T será

$$p = (x - c_1)^{r_1} \ldots (x - c_k)^{r_k}$$

onde $1 \leq r_i \leq d_i$. Se W_i é o núcleo de $(T - c_iI)^{r_i}$, então o teorema da decomposição primária nos diz que

$$V = W_1 \oplus \ldots \oplus W_k$$

e que o operador T_i induzido sôbre W_i por T possui polinômio minimal igual a $(x - c_i)^{r_i}$. Seja N_i o operador linear sôbre W_i definido por $N_i = T_i - c_iI$. Então, N_i é nilpotente e seu polinômio minimal é x^{r_i}. Sôbre W_i, T age como N_i mais o escalar c_i vêzes o operador idêntico. Suponhamos que tomemos uma base do subespaço W_i correspondente à decomposição cíclica do operador nilpotente N_i. Então a matriz de T_i em relação a esta base ordenada será a soma direta das matrizes

$$(7\text{-}7) \qquad \begin{bmatrix} c & 0 & \cdots & 0 & 0 \\ 1 & c & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ & & & c & \\ 0 & 0 & \cdots & 1 & c \end{bmatrix}$$

cada uma com $c = c_i$. Além disso, as dimensões destas matrizes diminuem quando se lê da esquerda para a direita. Uma matriz da forma (7-7) é dita uma **matriz elementar de Jordan com valor característico** c. Reunindo tôdas as bases dos W_i obtemos uma base de V. Descrevamos a matriz A de T em relação a esta base ordenada, a matriz A é a soma direta

$$(7\text{-}8) \qquad A = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & A_k \end{bmatrix}$$

das matrizes $A_1, \ldots, A_k$. Cada A_i é da forma

$$A_i = \begin{bmatrix} J_1^{(i)} & 0 & \cdots & 0 \\ 0 & J_2^{(i)} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & J_{n_i}^{(i)} \end{bmatrix}$$

onde cada $J_j^{(i)}$ é uma matriz elementar de Jordan com valor característico c_i. Além disso, dentro de cada A_i, as dimensões das matrizes $J_j^{(i)}$ diminuem à medida que j aumenta. Diremos que uma $n \times n$ matriz A que satisfaz tôdas as condições descritas até agora neste parágrafo (para certos escalares distintos $c_1, \ldots, c_k$) está sob a **forma de Jordan**.

Acabamos de salientar que se T é um operador linear para o qual o polinômio característico se decompõe completamente sôbre o corpo de escalares, então existe uma base ordenada de V em relação à qual T é representado por uma matriz que está sob a forma de Jordan. Gostaríamos de mostrar agora que esta matriz é algo associado de modo único a T, a menos da ordem em que os valores característicos de T são escritos. Em outras palavras, se duas matrizes estão sob a forma de Jordan e se elas são semelhantes, então elas podem diferir apenas quanto à ordem dos escalares c_i.

Podemos ver a unicidade como segue. Suponhamos que exista alguma base ordenada de V em relação à qual T seja representado pela matriz de Jordan A descrita no parágrafo anterior. Se A_i é uma $d_i \times d_i$ matriz, então d_i é evidentemente a multiplicidade de c_i como uma raiz do polinômio característico de A, ou de T. Em outras palavras, o polinômio característico de T é

$$f = (x - c_1)^{d_1} \dots (x - c_k)^{d_k}.$$

Isto mostra que $c_1, \dots, c_k$ e $d_1, \dots, d_k$ são únicos, a menos da ordem em que são escritos. O fato de que A é a soma direta das matrizes A_i nos fornece uma decomposição em soma direta $V = W_1 \oplus \oplus \dots \oplus W_k$ invariante sob T. Observemos agora que W_i deve ser o núcleo de $(T - c_iI)^n$, sendo $n = \dim V$; de fato, $A_i - c_iI$ é òbviamente nilpotente e $A_j - c_iI$ é não-singular para $j \neq i$. Portanto, vemos que os subespaços W_i são únicos. Se T_i é o operador induzido sôbre W_i por T, então a matriz A_i é determinada de um único modo como a forma racional de $(T_i - c_iI)$.

Desejamos agora fazer mais algumas observações sôbre o operador T e a matriz de Jordan A que representa T em relação a uma certa base ordenada. Faremos uma cadeia de observações:

(a) Todo elemento de A que não esteja na diagonal principal ou imediatamente abaixo dela é nulo. Na diagonal de A aparecem os k valores característicos distintos $c_1, \dots, c_k$ de T. Além disso, c_i se repete d_i vêzes, sendo d_i a multiplicidade de c_i como uma raiz do polinômio característico, isto é, $d_i = \dim W_i$.

(b) Para cada i, a matriz A_i é a soma direta de n_i matrizes elementares de Jordan, $J_j^{(i)}$, com valor característico c_i. O número n_i é exatamente a dimensão do espaço dos vetores característicos associados ao valor característico c_i. De fato, n_i é o número de blocos nilpotentes elementares na forma racional de $(T_i - c_iI)$ sendo portanto igual à dimensão do núcleo de $(T - c_iI)$. Em particular, notemos que T é diagonalizável se, e sòmente se, $n_i = d_i$ para todo i.

(c) Para cada i, o primeiro bloco $J_1^{(i)}$ na matriz A_i é uma $r_i \times r_i$ matriz, sendo r_i a multiplicidade de c_i como uma raiz do polinômio *minimal* de T. Isto decorre do fato de que o polinômio minimal do operador nilpotente $(T_i - c_iI)$ é x^{r_i}.

É claro que temos, como sempre, o mesmo resultado para matrizes. Se B é uma $n \times n$ matriz sôbre o corpo F e se o polinômio característico de B se decompõe completamente sôbre F, então B é semelhante sôbre F a uma $n \times n$ matriz A sob a forma de Jordan A é única a menos da ordem dos valores característicos. Dizemos que A é a **forma de Jordan de B.**

Além disso, notemos que se F é um corpo algèbricamente fechado, então as observações acima se aplicam a todo operador linear sôbre um espaço de dimensão finita, ou, a tôda $n \times n$ matriz sôbre F. Assim, por exemplo, tôda $n \times n$ matriz sôbre o corpo dos números complexos é semelhante a uma matriz essencialmente única sob a forma de Jordan.

Exemplo 4. Suponhamos que T seja um operador linear sôbre C^2. O polinômio característico de T é $(x - c_1)(x - c_2)$, sendo c_1 e c_2 números complexos distintos ou então é $(x - c)^2$. No primeiro caso, T é diagonalizável e é representado em relação a alguma base ordenada por

$$\begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix}.$$

No segundo caso, o polinômio minimal de T pode ser $(x - c)$, e então $T = cI$ e pode ser $(x - c)^2$ e então T é representado em relação a alguma base pela matriz

$$\begin{bmatrix} c & 0 \\ 1 & c \end{bmatrix}.$$

Assim, tôda 2×2 matriz sôbre o corpo dos números complexos é semelhante a uma matriz de um dos dois tipos acima exibidos, possívelmente com $c_1 = c_2$.

Exemplo 5. Seja A a 3×3 matriz complexa

$$\begin{bmatrix} 2 & 0 & 0 \\ a & 2 & 0 \\ b & c & -1 \end{bmatrix}.$$

O polinômio característico de A é òbviamente $(x - 2)^2 (x + 1)$. Ou êste polinômio é o polinômio minimal e A é semelhante a

$$\begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

ou então o polinômio minimal é $(x - 2)(x + 1)$, caso em que A é semelhante a

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Ora,

$$(A - 2I)(A + I) = \begin{bmatrix} 0 & 0 & 0 \\ 3a & 0 & 0 \\ ac & 0 & 0 \end{bmatrix}.$$

e assim A é semelhante a uma matriz diagonal se, e sòmente se, $a = 0$.

Exemplo 6. Seja

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & a & 2 \end{bmatrix}.$$

O polinômio característico de A é $(x - 2)^4$. Como A é a soma direta de duas 2×2 matrizes, é evidente que o polinômio minimal de A é $(x - 2)^2$. Ora, se $a = 0$ ou se $a = 1$, então a matriz A está sob a forma de Jordan. Notemos que as duas matrizes que se obtêm para $a = 0$ e $a = 1$ têm o mesmo polinômio característico e o mesmo polinômio minimal, mas não são semelhantes. Elas não são semelhantes porque, para a primeira matriz, o espaço-solução de $(A - 2I)$ tem dimensão 3, enquanto que para a segunda matriz a dimensão é 2.

Exercícios

1. Sejam N_1 e N_2 3×3 matrizes nilpotentes sôbre o corpo F. Demonstrar que N_1 e N_2 são semelhantes, se, e sòmente se, possuem o mesmo polinômio minimal.

2. Usar o resultado do Exercício 1 e a forma de Jordan para demonstrar o seguinte: Sejam A e B $n \times n$ matrizes sôbre o corpo F que possuam o *mesmo* polinômio característico

$$f = (x - c_1)^{d_1} \ldots (x - c_k)^{d_k}$$

e o mesmo polinômio minimal. Se nenhum d_i é maior que 3, então A e B são semelhantes.

3. Se A é uma 5×5 matriz complexa com polinômio característico $f = (x - 2)^3 (x + 7)^2$ e polinômio minimal $p = (x - 2)^2 (x + 7)$, qual é a forma de Jordan de A?

4. Quantas formas de Jordan são possíveis para a 6×6 matriz complexa cujo polinômio característico é $(x + 2)^4 (x - 1)^2$?

5. O operador derivação sôbre o espaço dos polinômios de grau menor ou igual a 3 é representado em relação à base ordenada "natural" pela matriz

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Qual é a forma de Jordan desta matriz? (F é um subcorpo do corpo dos números complexos.)

6. Seja A a matriz complexa

$$\begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 \\ -1 & 0 & 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 & 0 \\ 1 & 1 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}.$$

Determinar a forma de Jordan de A.

7. Se A é uma $n \times n$ matriz sôbre o corpo F com polinômio característico

$$f = (x - c_1)^{d_1} \dots (x - c_k)^{d_k},$$

qual é o traço de A?

8. Classificar, a menos da semelhança, tôdas as 3×3 matrizes complexas A tais que $A^3 = I$.

9. Classificar, a menos da semelhança, tôdas as $n \times n$ matrizes complexas A tais que $A^n = I$.

10. Seja n um inteiro positivo, $n \geq 2$ e seja N uma $n \times n$ matriz sôbre o corpo F tal que $N^n = 0$ mas $N^{n-1} \neq 0$. Demonstrar que N não possui nenhuma raiz quadrada, isto é, que não existe nenhuma $n \times n$ matriz A tal que $N^2 = N$.

11. Sejam N_1 e N_2 6×6 matrizes nilpotentes sôbre o corpo F. Suponhamos que N_1 e N_2 tenham o mesmo polinômio minimal e a mesma nulidade. Demonstrar que N_1 e N_2 são semelhantes. Mostrar que isto não é válido para 7×7 matrizes nilpotentes.

12. Usar o resultado do Exercício 11 e a forma de Jordan para demonstrar o seguinte: Sejam A e B $n \times n$ matrizes sôbre o corpo F que possuam o mesmo polinômio característico

$$f = (x - c_1)^{d_1} \dots (x - c_k)^{d_k}$$

e o mesmo polinômio minimal. Suponhamos também que para cada i, os espaços-soluções de $(A - c_iI)$ e de $(B - c_iI)$ tenham a mesma dimensão. Se nenhum dos d_i é maior que 6, então A e B são semelhantes.

13. Se N é uma $k \times k$ matriz nilpotente elementar, isto é, $N^k = 0$ mas $N^{k-1} \neq 0$, mostrar que N^t é semelhante a N. Usar agora a forma de Jordan para demonstrar que tôda $n \times n$ matriz complexa é semelhante à sua transposta.

14. O que está errado na demonstração que segue? Se A é uma $n \times n$ matriz complexa tal que $A^t = -A$, então A é 0. *Demonstração*: Seja J a forma de Jordan de A. Como $A^t = -A$, $J^t = -J$. Mas J é triangular, logo $J^t = -J$ implica que todo elemento de J é nulo. Como $J = 0$ e A é semelhante a J, vemos que $A = 0$. (Dar um exemplo de uma A não-nula tal que $A^t = -A$.)

15. Se N é uma 3×3 matriz nilpotente sôbre C, demonstrar que $A = I + \frac{1}{2} N - \frac{1}{8} N^2$ satisfaz $A^2 = I + N$, isto é, A é uma raiz quadrada $I + N$. Usar a série binomial $(1 + t)^{1/2}$ para obter uma fórmula semelhante para a raiz quadrada de $I + N$, onde N é uma $n + n$ matriz nilpotente arbitrária sôbre C.

16. Usar o resultado do Exercício 15 para demonstrar que se c é um número complexo não-nulo e N é uma matriz complexa nilpotente, então $(cI + N)$ possui uma raiz quadrada. Usar depois a forma de Jordan para demonstrar que tôda $n \times n$ matriz complexa não-singular possui uma raiz quadrada.

7.4 Resumo; Operadores Semi-Simples

Nos dois últimos capítulos, estivemos tratando de um único operador linear T sôbre um espaço vetorial V de dimensão finita. O programa foi decompor T numa soma direta de operadores lineares de natureza elementar, com o objetivo de obter informações detalhadas sôbre como T 'opera' sôbre o espaço V. Recordemos ràpidamente onde nos encontramos.

Começamos estudando T por meio de valores característicos e vetores característicos. Introduzimos os operadores diagonalizáveis, operadores que podem ser descritos completamente em têrmos de valores e vetores característicos. Observamos então que podia ocorrer que T não tivesse nenhum vetor característico. Mesmo no caso de um corpo de escalares algèbricamente fechado, em que todo operador linear realmente possui pelo menos um vetor característico, notamos que os vetores característicos de T nem sempre geravam o espaço.

Demonstramos então o teorema da decomposição racional, exprimindo um operador linear arbitrário como a soma direta de operadores que tinham um vetor cíclico, sem fazer nenhuma hipótese quanto ao corpo de escalares. Se U é um operador linear que tem um vetor cíclico, existe uma base $\{\alpha_1, \ldots, \alpha_n\}$ tal que

$$U\alpha_j = \alpha_{j+1}, \quad j = 1, \ldots, n-1$$
$$U\alpha_n = -c_o\alpha_1 - c_1\alpha_2 - \ldots - c_{n-1}\alpha_n.$$

A ação de U sôbre esta base é então a de transformar cada α_j no vetor seguinte α_{j+1}, com a exceção de $U\alpha_n$ que é uma combinação linear predeterminada dos vetores da base. Como um operador linear genérico T é a soma direta de um número finito de tais operadores U, obtivemos uma descrição explícita e razoàvelmente elementar da ação de T.

Aplicamos a seguir o teorema da decomposição racional a operadores nilpotentes. Para o caso de um corpo de escalares algèbricamente fechado, combinamos êste resultado com o teorema da decomposição primária obtendo a forma de Jordan. A forma de Jordan fornece uma base $\{\alpha_1, \ldots, \alpha_n\}$ do espaço V tal que, para cada j, ou $T\alpha_j$ é um múltiplo escalar de α_j ou $T\alpha_j = c\alpha_j + \alpha_{j+1}$. Esta base certamente descreve a ação de T de uma maneira explícita e elementar.

A importância da forma racional (ou da forma de Jordan) origina-se do fato de ela existir e não do fato de poder ser determinada em casos particulares. É claro que se se tem um particular operador linear T e se pode determinar a sua forma racional ou de Jordan, deve-se fazê-lo, pois, tendo esta forma, pode-se conseguir vastas quantidades de informações sôbre T. Dois tipos de dificuldades surgem no cálculo dessas formas canônicas. Uma dificuldade é, òbviamente, a extensão dos cálculos. A outra dificuldade é que pode não existir nenhum método de efetuar os cálculos mesmo que se tenham paciência e tempo suficientes. A segunda dificuldade surge ao, digamos, se tentar determinar a forma de Jordan de uma matriz complexa. Simplesmente não existe nenhum método bem definido de se decompor o polinômio característico e assim já se é barrado no início. A forma racional não apresenta esta dificuldade. Em outras palavras, existe um método bem definido para se determinar a forma racional de uma dada $n \times n$ matriz; contudo, tais cálculos são usualmente longos demais. O leitor interessado deverá consultar o livro de A. A. Albert citado na Bibliografia para uma discussão dêste aspecto da forma racional.

Em nosso resumo dos resultados dêstes dois últimos capítulos, ainda não mencionamos um teorema que demonstramos. É o teorema que afirma que se T é um operador linear sôbre um espaço vetorial de dimensão finita sôbre um corpo algèbricamente fechado, então T pode ser expresso de um único modo como a soma de um operador diagonalizável com um operador nilpotente os quais comutam. Êle foi demonstrado a partir do teorema da decomposição primária e algumas informações sôbre operadores diagonalizáveis. Não é um teorema tão profundo como o teorema da decomposição racional ou a existência da forma de Jordan, mas possui aplicações importantes e úteis em certas partes da matemática. Concluindo êste capítulo, vamos demonstrar um teorema análogo, sem supor que o corpo de escalares seja algèbricamente fechado. Começamos por definir os operadores que desempenharão o papel dos operadores diagonalizáveis.

Definição. *Seja* V *um espaço vetorial de dimensão finita sôbre o corpo* F *e seja* T *um operador linear sôbre* V. *Dizemos que* T *é* **semi-simples** *se todo subespaço* T*-invariante possui um subespaço suplementar* T*-invariante.*

O que estamos prestes a demonstrar é que, com certas restrições sôbre o corpo F, todo operador linear T pode ser expresso de um único modo como $T = S + N$, sendo S semi-simples, N nilpotente e $SN = NS$. Primeiro, vamos caracterizar os operadores semi-simples por meio de seus polinômios minimais e esta caracterização nos mostrará que, quando F é algèbricamente fechado, um operador é semi-simples se, e sòmente se, é diagonalizável.

Lema. *Seja* T *um operador linear sôbre o espaço vetorial* V *de dimensão finita e seja* $V = W_1 \oplus \ldots \oplus W_k$ *a decomposição primária de* T. *Em outras palavras, se* n *é o polinômio de* T *e* $p = p_1^{r_1} \ldots$ $\ldots p_k^{r_k}$ *é a decomposição de* n *em fatôres primos, então* W_j *é o núcleo* *de* $p_j(T)^{r_j}$. *Seja* N *um subespaço arbitrário de* V *que seja invariante sob* T. *Então*

$$W = (W \cap W_1) \oplus \ldots \oplus (W \cap W_k).$$

Demonstração. Para a demonstração precisaremos recordar um corolário de nossa demonstração do teorema da decomposição primária na Seção 6.4. Se $E_1, \ldots, E_k$ são as projeções associadas à decomposição $V = W_1 \oplus \ldots \oplus W_k$, então cada E_j é um polinômio em T. Isto é, existem polinômios $h_1, \ldots, h_k$ tais que $E_j = h_j(T)$.

Seja agora W um subespaço invariante sob T. Se α é um vetor qualquer em W, então $\alpha = \alpha_1 + \ldots + \alpha_k$, com α_j em W_j. Ora, $\alpha_j = E_j\alpha = h_j(T)\alpha$, e como W é invariante sob T, cada α_j também está em W. Assim, cada vetor α em W é da forma $\alpha = \alpha_1 + \ldots + \alpha_k$, com α_j na interseção $W \cap W_j$. Esta expressão é única pois $V = W_1 \oplus \ldots \oplus W_k$. Portanto

$$W = (W \cap W_1) \oplus \ldots \oplus (W \cap W_k).$$

Lema. *Seja* T *um operador linear sôbre* V *e suponhamos que o polinômio minimal de* T *seja irredutível sôbre o corpo* F *de escalares Então* T *é semi-simples.*

Demonstração. Seja W um subespaço de V que seja invariante sob T. Precisamos demonstrar que W possui um subespaço suplementar T-invariante. De acôrdo com o Teorema 3, será suficiente demonstrar que se f é um polinômio e β é um vetor em V tais que $f(T)\beta$ esteja em W, então existe um vetor α em W tal que $f(T)\beta = f(T)\alpha$. Portanto, suponhamos que β esteja em V e que f seja um polinômio tal que $f(T)\beta$ esteja em W. Se $f(T)\beta = 0$, fazemos $\alpha = 0$ e então α é um vetor em W tal que $f(T)\beta = f(T)\alpha$. Se $f(T)\beta \neq 0$, o polinômio f não é divisível pelo polinômio minimal p do operador T. Como p é primo, isto significa que f e p são relativamente primos e existem polinômios g e h tais que $fg + ph = 1$. Como $p(T) = 0$ temos $f(T)g(T) = I$. Daqui segue que o vetor B deve estar no subespaço W; de fato.

$$\begin{aligned} \beta &= g(T)f(T)\beta \\ &= g(T)\, f(T)\beta) \end{aligned}$$

enquanto $f(T)\beta$ está em W e W é invariante sob T. Basta tomar $\alpha = \beta$.

Teorema 6. *Seja* T *um operador linear sôbre o espaço vetorial* V *de dimensão finita. Uma condição necessária e suficiente para que* T *seja semi-simples é que o polinômio minimal* p *de* T *seja da forma* $p = p_1 \ldots p_k$, *sendo* $p_1, \ldots, p_k$ *polinômios irredutíveis distintos sôbre o corpo* F *de escalares.*

Demonstração. Suponhamos que T seja semi-simples. Mostraremos que nenhum polinômio irredutível se repete na decomposição do polinômio minimal p em fatôres primos. Suponhamos o contrário. Então existe um polinômio unitário não-constante g tal que g^2 divide p. Seja W o núcleo do operador $g(T)$. Então W é invariante sob T. Ora, $p = g^2h$ para algum polinômio h. Como g não é um polinômio constante, o operador $g(T)h(T)$ não é o operador nulo e existe um vetor β em V tal que $g(T)h(T)\beta = 0$, isto é, $(gh)\beta \neq 0$. Ora, $(gh)\beta$ está no subespaço W, pois $g(gh\beta) = g^2h\beta = p\beta = 0$. Mas não existe nenhum vetor α em W tal que $gh\,\beta = gh\,\alpha$; de fato, se α está em W

$$(gh)\,\alpha = (hg)\,\alpha = h(g\,\alpha) = h(0) = 0.$$

Assim, W não pode ter um subespaço suplementar T-invariante, contradizendo a hipótese de T ser semi-simples.

Suponhamos agora que a decomposição de p em fatôres primos seja $p = p_1 \dots p_k$, sendo $p_1, \dots, p_k$ polinômios unitários (não-constantes), irredutíveis e distintos. Seja W um subespaço de V que seja invariante sob T. Vamos demonstrar que W possui um subespaço suplementar T-invariante. Seja $V = W_1 \oplus \dots \oplus W_k$ a decomposição primária de T, isto é, seja W_j o núcleo de $p_j(T)$. Seja T_j o operador linear induzido sôbre W_j por T, de modo que o polinômio minimal de T_j é o primo p_j. Ora, $W \cap W_j$ é um subespaço de W_j que é invariante sob T_j (ou sob T). Pelo último lema, existe um subespaço V_j de W_j tal que $W_j = (W \cap W_j) \oplus V_j$ e V_j seja invariante sob T_j (e portanto sob T). Então temos

$$\begin{aligned} V &= W_1 \oplus \dots \oplus W_k \\ &= (W \cap W_1) \oplus V_1 \oplus \dots \oplus (W \cap W_k) \oplus V_k \\ &= (W \cap W_1) + \dots + (W \cap W_k) \oplus V_1 \oplus \dots \oplus V_k. \end{aligned}$$

Pelo primeiro lema acima, $W = (W \cap W_1) \oplus \dots \oplus (W \cap W_k)$, de modo que se $W' = V_1 \oplus \dots \oplus V_k$, então $V = W \oplus W'$ e W' é invariante sob T.

Corolário. *Se* T *é um operador linear sôbre um espaço vetorial de dimensão finita sôbre um corpo algèbricamente fechado, então* T *é semi-simples se, e sòmente se,* T *é diagonalizável.*

Demonstração. Se o corpo F de escalares é algèbricamente fechado, os primos unitários sôbre F são os polinômios $x - c$. Neste caso, T é semi-simples se, e sòmente se, o polinômio minimal de T é $p = (x - c_1)\dots(x - c_k)$, sendo $c_1, \dots, c_k$ elementos distintos de F. Êste é exatamente o critério para a diagonalização de T, por nós estabelecido no Capítulo 6.

Gostaríamos de destacar que T é semi-simples se, e sòmente se, existe um polinômio f, que seja um produto de primos distintos, tal que $f(T) = 0$. Isto difere apenas superficialmente da condição de que o polinômio minimal seja um produto de primos distintos.

Voltemos agora ao problema de exprimir um operador linear como a soma de um operador semi-simples e um operador nilpotente que comutem. Para esta parte, restringiremos o corpo de escalares a um subcorpo do corpo dos números complexos. O leitor informado verá que o importante é o corpo F ser um corpo de característica zero, isto é, para cada inteiro positivo n, a soma $1 + \ldots + 1$ (n vêzes) em F não deve ser nula. Para um polinômio f sôbre F, indiquemos por $f^{(k)}$ a k-ésima derivada formal de f. Em outras palavras, $f^{(k)} = D^k f$ onde D é o operador derivação sôbre o espaço dos polinômios. Se g é um outro polinômio, $f(g)$ indica o resultado de se substituir g em f, isto é, o polinômio obtido aplicando f ao elemento g na álgebra linear $F[x]$.

Lema (Fórmula de Taylor). *Seja* F *um subcorpo do corpo dos números complexos e sejam* g *e* h *polinômios sôbre* F. *Se* f *é um polinômio qualquer sôbre* F *com* $gr(f) \leq n$, *então*

$$f(g) = f(h) + f^{(1)}(h)(g-h) + \frac{f^{(2)}(h)}{2!}(g-h)^2 + \ldots + \frac{f^{(n)}(h)}{n!}(g-h)^n.$$

Demonstração. O que estamos demonstrando é uma fórmula de Taylor generalizada. O leitor provàvelmente está acostumado a ver o caso particular em que $h = c$, um polinômio constante, e $g = x$. Nesse caso, a fórmula diz:

$$f = f(x) = f(c) + f^{(1)}(c)(x-c) + \frac{f^{(2)}(c)}{2!}(x-c)^2 + \ldots + \frac{f^{(n)}(c)}{n!}(x-c)^n.$$

A demonstração desta fórmula é simplesmente uma aplicação do teorema binomial

$$(a+b)^k = a^k + ka^{k-1}b + \frac{k(k-1)}{2!}a^{k-2}b^2 + \ldots + b^k.$$

De fato, o leitor notará que, sendo a substituição e a derivação processos lineares, basta demonstrar a fórmula para $f = x^k$. A fórmula para $f = \sum_{k=0}^{n} c_k x^k$ decorre por uma combinação linear. No caso $f = x^k$ com $k \leq n$, a fórmula diz

$$g^k = h^k + kh^{k-1}(g-h) + \frac{k(k-1)}{2!}h^{k-2}(g-h)^2 + \ldots + (g-h)^k$$

que é exatamente o desenvolvimento binomial de

$$g^k = [h + (g - h)]^k.$$

Lema. *Seja* F *um subcorpo do corpo dos números complexos, seja* f *um polinômio sôbre* F *e seja* f' *a derivada de* f. *As seguintes afirmações são equivalentes:*

(i) f *é um produto de polinômios irredutíveis e distintos sôbre* F.
(ii) f *e* f' *são relativamente primos.*
(iii) *Considerado como um polinômio com coeficientes complexos,* f *não possui raízes múltiplas.*

Demonstração. Demonstraremos primeiro que (i) e (ii) são afirmações equivalentes sôbre f. Suponhamos, na decomposição de f em fatôres primos sôbre o corpo F, que algum polinômio (não-constante) primo p se repita. Então $f = p^2h$ para algum h em $F[x]$. Então,

$$f' = p^2h' + 2pp'h$$

e p é também um divisor de f'. Logo, f e f' não são relativamente primos. Concluímos que (ii) implica (i).

Suponhamos agora que $f = p_1 \ldots p_k$, onde $p_1, \ldots, p_k$ são polinômios não-constantes, irredutíveis e distintos sôbre F. Seja $f_j = f/p_j$. Então,

$$f' = p_1'f_1 + p_2'f_2 + \ldots + p_k'f_k.$$

Seja p um polinômio primo que divida f e f'. Então $p = p_i$ para algum i. Ora, p_i divide f_j para $j \neq i$ e como p_i também divide

$$f' = \sum_{j=1}^{k} p_j'f_j$$

vemos que p_i deve dividir $p_i'f_i$. Portanto, p_i divide f_i ou p_i'. Mas p_i não divide f_i uma vez que $p_1, \ldots, p_k$ são distintos. Então, p_i divide p_i'. Isto não é possível, pois o grau de p_i' é um a menos que o grau de p_i. Concluímos que nenhum primo divide f e f', ou seja, que $(f, f') = 1$.

Para ver que a afirmação (iii) é equivalente a (i) e (ii), precisamos observar apenas o seguinte: Suponhamos que f e g sejam polinômios sôbre F, um subcorpo do corpo dos números complexos. Podemos considerar f e g também como polinômios com coeficientes complexos. A afirmação de que f e g são relativamente primos como polinômios sôbre F é equivalente à afirmação de que f e g são relativamente primos como polinômios sôbre o corpo dos números complexos. Deixamos a demonstração dêste resultado como exercício. Usemos êste fato com $g = f'$. Notemos que (iii) é exatamente (i) quando f é considerado como um polinômio sôbre o corpo dos números complexos. Assim, (ii) e (iii) são equivalentes, pelo mesmo argumento utilizado acima.

Podemos agora demonstrar um teorema que tornará mais evidente a relação entre operadores semi-simples e operadores diagonalizáveis.

Teorema 7. *Seja* F *um subcorpo do corpo dos números complexos, seja* V *um espaço vetorial de dimensão finita sôbre* F *e seja* T *um operador linear sôbre* V. *Seja* $\mathfrak{B}$ *uma base ordenada de* V *e seja* A *a matriz de* T *em relação à base ordenada* $\mathfrak{B}$. *Então,* T *é semi-simples se, e sòmente se, a matriz* A *é semelhante, sôbre o corpo dos números complexos, a uma diagonal.*

Demonstração. Seja p o polinômio minimal de T. De acôrdo com o teorema 6, T é semi-simples se, e sòmente se, $p = p_1 \ldots p_k$ onde $p_1, \ldots, p_k$ são polinômios distintos irredutíveis sôbre F. Pelo último lema, temos que T é semi-simples se, e sòmente se, p não possui raízes complexas múltiplas.

Ora, p também é o polinômio minimal da matriz A. Sabemos que A é semelhante sôbre o corpo dos números complexos a uma matriz diagonal se, e sòmente se, o seu polinômio minimal não possui raízes complexas múltiplas. Isto demonstra o teorema, a menos de um detalhe. Deve-se observar que A possui o mesmo polinômio minimal, seja considerada como uma matriz sôbre F ou como uma matriz sôbre o corpo dos números complexos. Em outras palavras, se tomarmos o polinômio unitário de menor grau dentre todos os polinômios com coeficientes complexos que levam A em 0, êste polinômio terá seus coeficientes no subcorpo F. Um polinômio $f = c_0 + c_1x + \ldots + c_nx^n$ tal que $f(A) = 0$ corresponde a uma relação linear

$$c_0I + c_1A + \ldots + c_nA^n = 0.$$

entre as potências de A. Tal relação diz que $c_0, \ldots, c_n$ satisfazem um sistema de r^2 equações lineares homogêneas, onde A é uma $r \times r$ matriz. Os coeficientes dêste sistema de equações se originam dos elementos de A e são, portanto, elementos do subcorpo F. Se tal um sistema de equações possui uma solução não-trivial com $c_0, \ldots, c_n$ números complexos, então êle admite uma solução não trivial com $c_0, \ldots, c_n$ em F. Isto mostra que o polinômio minimal de A, considerada como uma matriz sôbre F, tem o mesmo grau que o polinômio minimal de A, como uma matriz sôbre C. Êstes dois polinômios unitários têm o mesmo grau e o segundo divide o primeiro, logo são idênticos.

Teorema 8. *Seja* F *um subcorpo do corpo dos números complexos, seja* V *um espaço vetorial de dimensão finita sôbre* F *e seja* T *um operador linear sôbre* V. *Existe um operador semi-simples* S *sôbre* V *e um operador nilpotente* N *sôbre* V *tais que*

(i) $T = \mathrm{S} + \mathrm{N}$;

(ii) $\mathrm{SN} = \mathrm{NS}$.

Além disso, o S *semi-simples e o N nilpotente que satisfazem* (*i*) *e* (*ii*) *são únicos e cada um é um polinômio em* T.

Demonstração. Seja $p_1^{r_1} \dots p_k^{r_k}$ a decomposição em fatôres primos do polinômio minimal de T e seja $f = p_1 \dots p_k$. Seja r o maior dos inteiros positivos $r_1, \dots, r_k$. Então, o polinômio f é um produto de primos distintos, f^r é divisível pelo polinômio minimal de T e então

$$f(T)^r = 0.$$

Vamos construir uma seqüência de polinômios: $g_0, g_1, g_2, \dots$ tais que

$$f\left(x - \sum_{j=0}^{n} g_j f^j\right)$$

seja divisível por f^{n-1}, $n = 0, 1, 2, \dots$. Tomando $g_0 = 0$ temos que $f(x - g_0 f^0) = f(x) = f$ é divisível por f. Suponhamos que tenhamos escolhido $g_0, \dots, g_{n-1}$. Seja

$$h = x - \sum_{j=0}^{n-1} g_j f^j$$

de modo que, por hipótese, $f(h)$ é divisível por f^n. Queremos tomar g_n de modo que

$$f(h - g_n f^n)$$

seja divisível por f^{n+1}. Aplicando a fórmula geral de Taylor, obtemos

$$f(h - g_n f^n) = f(h) - g_n f^n f'(h) + f^{n+1} b$$

onde b é algum polinômio. Por hipótese, $f(h) = qf^n$. Assim, vemos que para $f(h - g_n f^n)$ ser divisível por f^{n+1} basta escolher g_n de maneira tal que $(q - g_n f')$ seja divisível por f. Isto pode ser feito, pois f não possui fatôres primos repetidos e então f e f' são relativamente primos. Se a e e são polinômios tais que $af + ef' = 1$ e se fizermos $g_n = eq$, então $q - g_n f'$ será divisível por f.

Agora temos uma seqüência $g_0, g_1, \dots$ tal que f^{n+1} divide $f\left(x - \sum_{j=0}^{n} g_j f^j\right)$. Tomemos $n = r - 1$; como $f(T)^r = 0$, temos

$$f\left(T - \sum_{j=0}^{r-1} g_j(T) f(T)^j\right) = 0.$$

Seja

$$N = \sum_{j=1}^{r-1} g_j(T)f(T)^j = \sum_{j=0}^{r-1} g_j(T)f(T)^j.$$

Como $\sum_{j=1}^{n} g_j f^j$ é divisível por f, vemos que $N^r = 0$ e N é nilpotente. Seja $S = T - N$. Então $f(S) = f(T - N) = 0$. Como f possui fatôres primos distintos, S é semi-simples.

Temos agora $T = S + N$ onde S é semi-simples, N é nilpotente e cada um é um polinômio em T. Para demonstrar a afirmação da unicidade, passaremos do corpo de escalares F ao corpo dos números complexos. Seja $\mathcal{B}$ uma base ordenada do espaço V. Então temos

$$[T]_{\mathcal{B}} = [S]_{\mathcal{B}} + [N]_{\mathcal{B}}$$

sendo $[S]_{\mathcal{B}}$ diagonalizável sôbre o corpo dos números complexos e $[N]_{\mathcal{B}}$ nilpotente. Esta matriz diagonalizável e esta matriz nilpotente que comutam são determinadas de modo único, como demonstraremos no Capítulo 6.

Exercícios

1. Se N é um operador linear nilpotente sôbre V, mostrar que para todo polinômio f, a parte semi-simples de $f(N)$ é um múltiplo escalar do operador idêntico (F é um subcorpo de C).

2. Seja F um subcorpo do corpo dos números complexos, B um espaço vetorial de dimensão finita sôbre F e T um operador linear semi-simples sôbre V. Se f é um polinômio arbitrário sôbre F, demonstrar que $f(T)$ é semi-simples.

3. Seja T um operador linear sôbre um espaço de dimensão finita sôbre um subcorpo de C. Demonstrar que T é semi-simples se, e sòmente se, vale o seguinte: Se f é um polinômio e $f(T)$ é nilpotente, então $f(T) = 0$.

CAPÍTULO 8

ESPAÇOS COM PRODUTO INTERNO

8.1 Produtos Internos

Em todo êste capítulo trataremos apenas de espaços vetoriais reais ou complexos, isto é, de espaços vetoriais sôbre o corpo dos números reais ou sôbre o corpo dos números complexos. Nosso objetivo principal é estudar espaços vetoriais nos quais tenha sentido falar do "comprimento" de um vetor e do "ângulo" entre dois vetores. Faremos isto por meio do estudo de um certo tipo de função definida sôbre pares de vetores e tomando valores escalares, conhecida como um "produto interno". Um exemplo de produto interno é o produto escalar de vetores em R^3. O produto escalar de

$$\alpha = (x_1, x_2, x_3) \quad \text{e} \quad \beta = (y_1, y_2, y_3)$$

em R^3 é o escalar

$$(\alpha,\beta) = x_1y_1 + x_2y_2 + x_3y_3.$$

Geomètricamente, êste produto escalar é o produto do comprimento de α pelo comprimento de β e pelo cosseno do ângulo entre α e β. Assim, é possível definir os conceitos geométricos de "comprimento" e "ângulo" em R^3 em têrmos do produto escalar que é algèbricamente definido. Um produto interno sôbre um espaço vetorial é uma generalização do produto escalar e, em têrmos de tal produto interno, pode-se também definir "comprimento" e "ângulo". Nossos comentários sôbre ângulos restringir-se-ão ao conceito de perpendicularidade (ou ortogonalidade) de dois vetores.

Nesta primeira seção, vamos definir produto interno, considerar alguns exemplos particulares e estabelecer algumas propriedades básicas do produto interno geral. Então, voltar-nos-emos ao trabalho de discutir comprimento e ortogonalidade.

Definição. *Seja* F *o corpo dos números reais ou o corpo dos números complexos e seja* V *um espaço vetorial sôbre* F. *Um* **produto interno** *sôbre* V *é uma função que associa a cada par ordenado de vetores* α, β *em* V *um escalar* (α, β) *em* F *de maneira tal que*

(a) $(\alpha + \beta, \gamma) = (\alpha, \gamma) + (\beta, \gamma)$;

(b) $(c\alpha, \beta) = c(\alpha, \beta)$;

(c) $(\beta, \alpha) = \overline{(\alpha, \beta)}$, *onde a barra indica conjugação complexa*;

(d) $(\alpha, \alpha) > 0$ *se* $\alpha \neq 0$.

Deve-se observar que as condições (a), (b) e (c) implicam o seguinte:

(e) $(\alpha, c\beta + \gamma) = \bar{c}(\alpha, \beta) + (\alpha, \gamma)$.

Outro fato merece ser mencionado. Quando F é o corpo R dos números reais, os complexos conjugados que aparecem em (c) e (e) são supérfluos; no entanto, no caso de F ser complexo êles são necessários para se obter a condição (d). Sem êstes complexos conjugados, teríamos a contradição óbvia:

$$(\alpha,\alpha) > 0 \quad \text{e} \quad (i\alpha, i\alpha) = -1(\alpha,\alpha) > 0.$$

Nos exemplos que seguem, como em todo o capítulo, F é o corpo dos números reais ou o corpo dos números complexos.

Exemplo 1. Sôbre F^n existe um produto interno que denominamos o **produto interno canônico**. É definido sôbre $\alpha = (x_1, \ldots, x_n)$ e $\beta = (y_1, \ldots y_n)$ por

(8-1) $$(\alpha,\beta) = x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n.$$

Quando $F = R$, esta definição torna-se

$$(\alpha,\beta) = x_1y_1 + \ldots + x_ny_n$$

e êste produto interno é freqüentemente denominado o "produto escalar".

Exemplo 2. Para $\alpha = (x_1, x_2)$ e $\beta = (y_1, y_2)$ em R^2, seja

$$(\alpha,\beta) = x_1y_1 - x_1y_2 - x_1y_2 + 4x_2y_2.$$

Como $(\alpha,\alpha) = (x_1 - x_2)^2 + 3x_2^2$, decorre que $(\alpha,\alpha) > 0$ se $\alpha \neq 0$. As condições (a), (b) e (c) da definição são fàcilmente verificadas.

Exemplo 3. Seja V o espaço de tôdas as $n \times n$ matrizes sôbre F. Então V é isomorfo a F^{n^2}, de uma maneira natural. Decorre portanto do Exemplo 1 que

$$(A, B) = \sum_{j,k} A_{jk}\overline{B_{jk}}$$

define um produto interno sôbre V. Além disso, introduzindo a matriz transposta conjugada B^*, onde $B^*kj = \overline{B_{jk}}$, podemos exprimir êste produto interno sôbre V em têrmos da função traço

$$(A, B) = \operatorname{tr}(AB^*) = \operatorname{tr}(B^*A).$$

De fato, pois

$$\begin{aligned}
\operatorname{tr}(AB^*) &= \sum_j (AB^*)_{jj} \\
&= \sum_j \sum_k A_{jk} B^*_{jk-kj} \\
&= \sum_j \sum_k A_{jk} \overline{B_{jk}}.
\end{aligned}$$

Exemplo 4. Seja V o espaço das $n \times 1$ matrizes (— colunas) sôbre F e seja Q uma $n \times n$ matriz inversível sôbre F. Para X, Y em V definamos

$$(X, Y) = Y^*Q^*QX.$$

Estamos identificando uma 1×1 matriz sôbre F com o seu único elemento. Quando Q é a matriz unidade, êste exemplo é essencialmente o mesmo que o Exemplo 1.

Exemplo 5. Seja V o espaço vetorial das funções contínuas definidas sôbre o intervalo unitário, $0 \leq t \leq 1$ e tomando valores complexos. Seja.

$$(f, g) = \int_0^1 f(t)\overline{g(t)}dt.$$

O leitor provàvelmente tem mais familiaridade com o espaço das funções contínuas definidas sôbre o intervalo unitário e tomando valores reais, e para êste espaço, a conjugação complexa sôbre g pode ser omitida.

Exemplo 6. Êste é na realidade tôda uma classe de exemplos. Pode-se construir novos produtos internos a partir de um dado produto interno pelo seguinte método: Sejam V e W espaços vetoriais sôbre o mesmo corpo e suponhamos que (,) seja um produto interno sôbre W. Se T é uma transformação linear não-singular de V em W, então

$$p_T(\alpha,\beta) = (T\alpha, T\beta)$$

define um produto interno p_T sôbre V. O produto interno do Exemplo 4 é um caso particular desta situação. Os que seguem também são casos particulares.

(a) Seja V um espaço vetorial de dimensão finita e seja

$$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$$

uma base ordenada de V. Sejam $\epsilon_1, \ldots, \epsilon_n$ os vetores da base canônica de F^n e seja T a transformação linear de V em F^n tal que $T\alpha_j = \epsilon_j$, $j = 1, \ldots, n$. Em outras palavras, seja T o isomorfismo "natural" de V em F^n determinado por $\mathcal{B}$. Se tomarmos o produto interno canônico sôbre F^n, então

$$p_T(\sum_j x_j\alpha_j, \sum_k y_k\alpha_k) = \sum_{j=1}^{n} x_j\bar{y}_j.$$

Assim, para tôda base $\mathcal{B}$ de V, existe um produto interno sôbre V com a propriedade de que $(\alpha_j, \alpha_k) = \delta_{jk}$; na verdade, é fácil mostrar que existe exatamente um tal produto interno. Mostraremos posteriormente que todo produto interno sôbre V é determinado por alguma base $\mathcal{B}$ da maneira acima.

(b) Consideremos novamente o Exemplo 5. Tomemos $V = W$, o espaço das funções contínuas sôbre o intervalo unitário. Seja T o operador linear "multiplicação por t", isto é, $(Tf)(t) = tf(t)$, $0 \leq t \leq 1$. É fácil ver que T é linear. T também é não-singular; de fato, suponhamos que $Tf = 0$. Então, $tf(t) = 0$ para $0 \leq t \leq 1$; logo, $f(t) = 0$ para $t > 0$. Como f é contínua temos também $f(0)$, ou seja, $f = 0$. Usando o produto interno do Exemplo 5, construamos um produto interno sôbre V por

$$\begin{aligned} p_T(f, g) &= \int_0^1 (Tf)(t)\overline{(Tg)(t)}dt \\ &= \int_0^1 f(t)\overline{g(t)}t^2dt. \end{aligned}$$

Voltemos agora a observações gerais sôbre produtos internos. Suponhamos que V seja um espaço vetorial complexo com produto interno. Então, para todos α, β em V

$$(\alpha, \beta) = Re(\alpha, \beta) + iIm(\alpha, \beta)$$

onde $Re(\alpha, \beta)$ e $Im(\alpha, \beta)$ são as partes real e imaginária do número complexo (α, β). Se z é um número complexo, então $Im(z) = Re(-iz)$. Decorre que

$$Im(\alpha, \beta) = Re[-i(\alpha, \beta)] = Re(\alpha, i\beta).$$

Assim, o produto interno é completamente determinado por sua "parte real" de acôrdo com

$$(8\text{-}2) \qquad (\alpha, \beta) = Re(\alpha, \beta) + iRe(\alpha, i\beta).$$

As vêzes é bastante útil saber que um produto interno sôbre um espaço vetorial real ou complexo é determinado por outra função, a chamada **forma quadrática** determinada pelo produto interno. Para defini-la, indiquemos primeiro a raiz quadrada positiva de (α, α)

por $||\alpha||$; $||\alpha||$ é denominada a **norma** de α em relação ao produto interno. Observando os produtos internos canônicos em R^1, C^1, R^2, e R^3, o leitor poderá se convencer de que é conveniente considerar a norma de α como o "comprimento" ou "magnitude" de α. *A* forma quadrática determinada pelo produto interno é a função que associa a cada vetor α o escalar $||\alpha||^2$. Decorre da definição que

$$||\alpha \pm \beta||^2 = ||\alpha||^2 \pm 2Re(\alpha, \beta) + ||\beta||^2.$$

Assim, no caso real,

$$(\alpha, \beta) = \frac{1}{4}||\alpha + \beta||^2 - \frac{1}{4}||\alpha - \beta||^2. \tag{8-3}$$

No caso complexo também precisamos usar (8-2) e obtemos uma expressão mais complicada:

$$(\alpha, \beta) = \frac{1}{4}||\alpha + \beta||^2 - \frac{1}{4}||\alpha - \beta||^2 + \frac{i}{4}||\alpha + i\beta||^2 - \frac{i}{4}||\alpha - i\beta||^2. \tag{8-4}$$

As equações (8-3) e (8-4) são denominadas as **identidades de polarização**. Notemos que (8-4) pode também ser escrita como segue:

$$(\alpha, \beta) = \frac{1}{4}\sum_{n=1}^{4} i^n||\alpha + i\beta||^2.$$

As observações acima valem para qualquer produto interno sôbre qualquer espaço vetorial real ou complexo, não importando sua dimensão. Voltamos agora ao caso em que V é de dimensão finita. Um produto interno sôbre um espaço de dimensão finita sempre pode ser descrito em têrmos de uma dada base por meio de uma matriz.

Suponhamos então que V seja de dimensão finita, que

$$\mathcal{B} = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$$

seja uma base ordenada de V e que nos seja dado um particular produto interno sôbre V; mostraremos que êle é completamente determinado pelos valores

$$G_{jk} = (\alpha_k, \alpha_j) \tag{8-5}$$

que assume sôbre pares de vetores em $\mathcal{B}$. Se

$$\alpha = \sum_k x_k\alpha_k \quad \text{e} \quad \beta = \sum_j y_j\alpha_j$$

então
$$\begin{aligned}(\alpha, \beta) &= (\sum_k x_k\alpha_k, \beta)\\ &= \sum_k x_k(\alpha_k, \beta)\\ &= \sum_k x_k \sum_j \bar{y}_j(\alpha_k, \alpha_j)\\ &= \sum_{j,k} \bar{y}_j G_{jk} x_k\\ &= Y^*GX\end{aligned}$$
onde X, Y são as matrizes das coordenadas de α, β em relação à base ordenada $\mathcal{B}$ e G é a matriz com elementos $G_{jk} = (\alpha_k, \alpha_j)$. Denominamos G **a matriz do produto interno em relação à base ordenada** $\mathcal{B}$. Decorre de (8-5) que $G = G^*$ ou, em outras palavras, que G é hermitiana; contudo, G não é uma matriz hermitiana típica. De fato, pois G deve satisfazer a condição adicional

(8-6) $$X^*GX > 0 \quad \text{se} \quad X \neq 0.$$

Em particular, G deve ser inversível. Caso contrário, existiria uma $X \neq 0$ tal que $GX = 0$. Quando escrita explìcitamente, (8-6) torna-se

(8-7) $$\sum_{j,k} \bar{x}_j G_{jk} x_k > 0, \quad X \neq 0.$$

Daqui vemos imediatamente que todo elemento diagonal de G deve ser positivo; no entanto, esta condição sôbre os elementos diagonais não é de forma alguma suficiente para assegurar a validez de (8-6).

Êste processo é reversível; isto é, se G é uma $n \times n$ matriz arbitrária sôbre F que satisfaz $G = G^*$ e (8-6), então G é a matriz, em relação à base ordenada $\mathcal{B}$, de algum produto interno sôbre V. Tal produto interno é aquêle definido por
$$(\alpha, \beta) = Y^*GX$$
onde X e Y são as matrizes das coordenadas de α e β em relação à base ordenada $\mathcal{B}$.

Assim, fixando-se uma base ordenada $\mathcal{B}$, pode-se obter uma descrição de todos os produtos internos possíveis sôbre o espaço V de dimensão finita. Por exemplo, podemos descrever todos os produtos internos sôbre F^n através de suas matrizes em relação à base ordenada canônica. Todo produto interno sôbre F^n é obtido tomando-se uma $n \times n$ matriz G sôbre F que satisfaça
$$G = G^*$$
$$\sum_{j,k} G_{jk}\bar{x}_j x_k > 0, \quad \text{se} \quad X \neq 0$$

e definindo então o produto interno de $\alpha = (x_1, \ldots, x_n)$ e $\beta = (y_1, \ldots, y_n)$ por

$$(\alpha, \beta) = \sum_{j,k} G_{jk}\bar{y}_j x_k = Y^*GX.$$

Para que uma tal descrição seja realmente útil, deve-se definir alguma maneira eficaz de se decidir quando é que uma dada matriz G satisfaz (8-7). Faremos isto posteriormente.

Exercícios

1. Seja V um espaço vetorial e $(\,,\,)$ um produto interno sôbre V.

(a) Mostrar que $(0, \beta) = 0$ para todo β em V.
(b) Mostrar que se $(\alpha, \beta) = 0$ para todo β em V, então $\alpha = 0$.

2. Seja V um espaço vetorial sôbre F. Mostrar que a soma de dois produtos internos sôbre V é um produto interno sôbre V. A diferença de dois produtos internos é um produto interno? Mostrar que um múltiplo positivo de um produto interno é um produto interno.

3. Descrever explìcitamente todos os produtos internos sôbre R_1 e sôbre C_1.

4. Verificar que o produto interno canônico sôbre F^n é um produto interno.

5. Seja $(\,,\,)$ o produto interno canônico sôbre R^2.

(a) Sejam $\alpha = (1, 2)$, $\beta = (-1, 1)$. Se γ é um vetor tal que $(\alpha, \gamma) = -1$ e $(\beta, \gamma) = 3$, determinar γ.

(b) Mostrar que para todo α em R^2 temos $\alpha = (\alpha, \epsilon_1)\epsilon_1 + (\alpha, \epsilon_2)\epsilon_2$.

6. Seja $(\,,\,)$ o produto interno canônico sôbre C^2 e seja T o operador linear $T(x_2, x_1) = (-x_2, x_1)$. Ora, T é "a rotação de 90°" e possui a propriedade de que $(\alpha, T\alpha) = 0$ para todo α em R^2. Determinar todos os produtos internos $[\,,\,]$ sôbre R^2 tais que $[\alpha, T\alpha] = 0$ para todo α.

7. Seja $(\,,\,)$ o produto interno canônico sôbre C^2. Demonstrar que não existe nenhum operador linear não-nulo sôbre C^2 tal que $(\alpha, T\alpha) = 0$ para todo α em C^2. Generalizar.

8. Seja V o espaço das 2×1 matrizes sôbre R e seja A uma 2×2 matriz com elementos reais. Para X, Y em V seja

$$f_A(X, Y) = Y^tAX.$$

Mostrar que f_A é um produto interno sôbre V se, e sòmente se, $A = A^t$, $A_{11} > 0$, $A_{22} > 0$ e $\det A > 0$.

9. Seja V um espaço vetorial real ou complexo com produto interno. Mostrar que a forma quadrática determinada pelo produto interno satisfaz a **regra do paralelogramo**

$$||\alpha + \beta||^2 + ||\alpha - \beta||^2 = 2||\alpha||^2 + 2||\beta||^2.$$

10. Seja $(\,,\,)$ o produto interno sôbre R^2 definido no Exemplo 2 e seja $\mathcal{B}$ a base ordenada canônica de R^2. Determinar a matriz dêste produto interno em relação a $\mathcal{B}$.

11. Mostrar que a fórmula

$$\left(\sum_j a_j x^j, \sum_k b_k x^k\right) = \sum_{j,k} \frac{a_j b_k}{j + k + 1}$$

define um produto interno sôbre o espaço $R[x]$ dos polinômios sôbre o corpo R. Seja W o subespaço dos polinômios de grau menor ou igual a n. Restringir o produto interno acima a W e determinar a matriz dêste produto interno sôbre W em relação à base ordenada $\{1, x, x^2, \ldots, x^n\}$. (*Sugestão*: Para mostrar que a fórmula define um produto interno, observar que

$$(f, g) = \int_0^1 f(t)g(t)dt$$

e trabalhar com a integral.)

12. Seja V um espaço vetorial de dimensão finita e seja $\mathfrak{B} = \{\alpha_1, \ldots, \alpha_n\}$ uma base de V. Seja $(\ ,\ )$ um produto interno sôbre V. Se $c_1, \ldots, c_n$ são n escalares arbitrários, mostrar que existe exatamente um vetor α em V tal que $(\alpha, \alpha_j) = c_j$, $j = 1, \ldots, n$.

13. Seja V um espaço vetorial complexo. Uma função J de V em V é denominada uma **conjugação** (também chamada função semilinear) se $J(\alpha + \beta) = J(\alpha) + J(\beta)$, $J(c\alpha) = \bar{c}J(\alpha)$ e $J(J(\alpha)) = \alpha$, para todos os escalares c e todos α, β em V. Se J é uma conjugação, mostrar que

(a) o conjunto W de todos α em V tais que $J\alpha = \alpha$ é um espaço vetorial sôbre R em relação às operações definidas em V;

(b) para cada α em V existe um único par de vetores β, γ em W tais que $\alpha = \beta + i\gamma$.

14. Seja V um espaço vetorial complexo e W um subconjunto de V com as seguintes propriedades:

(i) W é um espaço vetorial real em relação às operações definidas em V.
(ii) Para cada α em V existe um único par de vetores β, γ em W, tais que $\alpha = \beta + i\gamma$. Mostrar que a equação $J\alpha = \beta - i\gamma$ define uma conjugação sôbre V tal que $J\alpha = \alpha$ se, e sòmente se, α pertence a W e mostrar também que J é a única conjugação sôbre V com esta propriedade.

15. Determinar tôdas as conjugações sôbre C^1 e C^2.

16. Seja W um subespaço real de dimensão finita de um espaço vetorial complexo V. Mostrar que W satisfaz a condição (ii) do Exercício 14 se, e sòmente se, tôda base de W é também uma base de V.

17. Seja V um espaço vetorial complexo, J uma conjugação sôbre V, W o conjunto dos α em V tais que $J\alpha = \alpha$ e f um produto interno sôbre W. Mostrar que

(a) existe um único produto interno g sôbre V tal que $g(\alpha, \beta) = f(\alpha, \beta)$ para todos α, β em W,

(b) $g(J\alpha, J\beta) = g(\beta, \alpha)$ para todos α, β em V.

O que a parte (a) diz acêrca da relação entre os produtos internos canônicos sôbre R e C^1 ou sôbre R^n e C^n?

8.2 Espaços com Produto Interno

Agora que temos alguma idéia sôbre o que um produto inteno é, voltaremos nossa atenção para o que pode ser dito a respeito da combinação de um espaço vetorial e algum produto interno particular sôbre êle. Especìficamente, estabeleceremos as propriedades básicas dos conceitos de "comprimento" e "ortogonalidade" que são impostas ao espaço pelo produto interno.

Definição. *Um* **espaço com produto interno** *é um espaço vetorial real ou complexo, munido de um produto especificado sôbre aquêle espaço.*

Um espaço real com produto interno e de dimensão finita é freqüentemente denominado um **espaço euclidiano.** Um espaço complexo com produto interno é freqüentemente dito um **espaço unitário.**

Teorema 1. *Se* V *é um espaço com produto interno, então, para quaisquer vetores* α, β *em* V *e todo escalar* c

(a) $||c\alpha|| = |c|\ ||\alpha||$;
(b) $||\alpha|| > 0$ *para* $\alpha \neq 0$;
(c) $|(\alpha,\beta)| \leq ||\alpha||\ ||\beta||$;
(d) $||\alpha+\beta|| \leq ||\alpha|| + ||\beta||$.

Demonstração. As afirmações (a) e (b) decorrem quase imediatamente das diversas definições envolvidas. A desigualdade em (c) é evidentemente válida quando $\alpha = 0$. Se $\alpha \neq 0$, coloquemos

$$\gamma = \beta - \frac{(\beta,\alpha)}{||\alpha||^2}\alpha.$$

Então $(\gamma, \alpha) = 0$ e

$$\begin{aligned} 0 \leq ||\gamma||^2 &= \left(\beta - \frac{(\beta,\alpha)}{||\alpha||^2}\alpha, \beta - \frac{(\beta,\alpha)}{||\alpha||^2}\alpha\right) \\ &= (\beta, \beta) - \frac{\overline{(\beta,\alpha)}}{||\alpha||^2}(\beta, \alpha) \\ &= ||\beta||^2 - \frac{|(\alpha, \beta)|^2}{||\alpha||^2}. \end{aligned}$$

Logo $|(\alpha, \beta)|^2 \leq ||\alpha||^2||\beta||^2$. Usando agora (c) concluímos que

$$\begin{aligned} ||\alpha + \beta||^2 &= ||\alpha||^2 + (\alpha, \beta) + (\beta, \alpha) + ||\beta||^2 \\ &= ||\alpha||^2 + 2Re(\alpha, \beta) + ||\beta||^2 \\ &\leq ||\alpha||^2 + 2||\alpha||^2\ ||\beta|| + ||\beta||^2 \\ &= (||\alpha|| + ||\beta||)^2. \end{aligned}$$

Assim, $||\alpha + \beta|| \leq ||\alpha|| + ||\beta||$, o que demonstra (d).

Observação. *A desigualdade dada em* (c) *é denominada a desigualdade de* **Cauchy-Schwarz**. *Ela possui uma ampla gama de aplicações em matemática.*

Exemplo 7. Aplicando a desigualdade de Cauchy-Schwarz aos produtos internos dados nos Exemplos 1, 2, 3 e 5 obtemos o seguinte:

(a) $$|\Sigma\, x_k\bar{y}_k| \leq (\Sigma\, |x_k|^2)^{1/2}(\Sigma\, |y_k|^2)^{1/2}$$

(b) $$|x_1y_1 - x_2y_1 - x_1y_2 + 4x_2y_2| \leq ((x_1 - x_2)^2 + 3x_2^2)^{1/2}((y_1 - y_2)^2 + 3y_2^2)^{1/2}$$

(c) $$|\text{tr}\,(AB^*)| \leq |\text{tr}\,(AA^*)|^{1/2}|\text{tr}\,(BB^*)|^{1/2}$$

(d) $$\left|\int_0^1 f(x)g(x)dx\right| \leq \left(\int_0^1 |f(x)|^2dx\right)^{1/2}\left(\int_0^1 |g(x)|^2dx\right)^{1/2}.$$

Definição. *Sejam α e β vetores num espaço* V *com produto interno. Dizemos que α e β são* **ortogonais** *se $(\alpha, \beta) = 0$. Se* S *é um conjunto de vetores em* V, *dizemos que* S *é um* **conjunto ortogonal** *se dois quaisquer vetores distintos em* S *são ortogonais. Um* **conjunto ortonormal** *é um conjunto ortogonal* S, *com a propriedade adicional de que $||\alpha|| = 1$ para todo α em* S.

O vetor nulo é ortogonal a todo vetor em V e é o único vetor com esta propriedade. É conveniente pensar em um conjunto ortonormal como um conjunto de vetores mùtuamente perpendiculares, cada um tendo comprimento 1.

Exemplo 8. A base canônica de R^n ou de C^n é ortonormal em relação ao produto interno canônico sôbre R^n ou C^n.

Exemplo 9. Seja V o espaço das $n \times n$ matrizes complexas e E^{pq} a matriz cujo único elemento não-nulo é um 1 na linha p e coluna q. Então, o conjunto de tôdas estas matrizes E^{pq} é ortonormal em relação ao produto interno dado no Exemplo 3. De fato, pois

$$(E^{pq}, E^{rs}) = \text{tr}\,(E^{pq}E^{sr}) = \delta_{qs}\,\text{tr}\,(E^{pr}) = \delta_{qs}\,\delta_{pr}.$$

Exemplo 10. Se V é o espaço das funções contínuas definidas sôbre o intervalo $0 \leq x \leq 1$ e tomando valores complexos (ou valores reais), com o produto interno

$$(f, g) = \int_0^1 f(x)\overline{g(x)}\, dx$$

e se

$$f_n(x) = \sqrt{2}\cos 2\pi\, nx$$

$$g_n(x) = \sqrt{2}\,\text{sen}\, 2\,\pi\, nx$$

então $S = \{1, f_1, g_1, f_2, g_2, \ldots\}$ é um conjunto ortonormal. No caso complexo, se

$$h_n(x) = e^{2\pi inx}, \quad n = 0, \pm 1, \pm 2, \ldots$$

então $S = \{h_n\}$ é um conjunto ortonormal. Estamos supondo que o leitor tenha familiaridade com o cálculo das integrais acima. Deve-se notar que ambos os conjuntos há pouco exibidos são infinitos.

Teorema 2. *Um conjunto ortogonal de vetores não-nulos é linearmente independente.*

Demonstração. Seja S um conjunto ortogonal de vetores não-nulos num espaço com produto interno. Suponhamos que $\alpha_1, \alpha_2, \ldots, \alpha_m$ sejam vetores distintos em S e que

$$\beta = c_1\alpha_1 + c_2\alpha_2 + \ldots + c_m\alpha_m.$$

Então
$$\begin{aligned}(\beta, \alpha_k) &= \sum_j (c_j\alpha_j, \alpha_k)\\ &= \sum_j c_j(\alpha_j, \alpha_k)\\ &= c_k(\alpha_k, \alpha_k).\end{aligned}$$

Como $(\alpha_k, \alpha_k) \neq 0$, decorre que

$$c_k = \frac{(\beta, \alpha_k)}{||\alpha_k||^2}, \quad 1 \leq k \leq m.$$

Assim, quando $\beta = 0$, cada $c_k = 0$; logo S é um conjunto independente.

Corolário. *Se um vetor β é uma combinação linear de um conjunto ortogonal de vetores não-nulos $\alpha_1, \alpha_2, \ldots, \alpha_m$, então β é exatamente a combinação linear*

$$\beta = \sum_{k=1}^{m} \frac{(\beta, \alpha_k)}{||\alpha_k||^2}\alpha_k. \tag{8-8}$$

Êste corolário decorre da demonstração do teorema. Existe um outro corolário que, apesar de evidente, deve ser mencionado. Se $\{\alpha_1, \ldots, \alpha_m\}$ é um conjunto ortogonal de vetores não-nulos em um espaço V de dimensão finita com produto interno, então $m \leq \dim V$. Intuitivamente, isto diz que o número de dimensões mùtuamente ortogonais no espaço não pode exceder a dimensão do espaço, algèbricamente definida. O número máximo de direções ortogonais em V é o que provàvelmente se consideraria como sendo a dimensão geométrica de V, e acabamos de ver que esta não é maior que a dimensão algébrica. Evidentemente, veremos sem grande surprêsa que estas duas dimensões são idênticas.

Teorema 3. *Todo espaço de dimensão finita com produto interno possui uma base ortonormal.*

Demonstração. Seja V um espaço com produto interno e $\{\beta_1, \ldots, \beta_n\}$ uma base de V. A partir desta base, vamos obter uma base ortogonal, por meio de uma construção conhecida como o **processo de ortogonalização de Gram-Schmidt**. Em primeiro lugar, seja $\alpha_1 = \beta_1$. Então

$$\alpha_2 = \beta_2 - \frac{(\beta_2\ \alpha_1)}{||\alpha_1||^2}\,\alpha_1.$$

Como β_1 e β_2 são linearmente independentes, $\alpha_2 \neq 0$ e, por cálculos diretos, vemos que $(\alpha_2, \alpha_1) = 0$. Seja agora

$$\alpha_3 = \beta_3 - \frac{(\beta_3,\ \alpha_1)}{||\alpha_1||^2}\,\alpha_1 - \frac{(\beta_3,\ \alpha_2)}{||\alpha_2||^2}\,\alpha_2.$$

Então $\alpha_3 \neq 0$, pois caso contrário, β_3 seria uma combinação linear de β_1 e β_2; além disso, $(\alpha_3, \alpha_1) = (\alpha_3, \alpha_2) = 0$. Suponhamos agora que tenhamos construído vetores ortogonais não-nulos $\alpha_1, \alpha_2, \ldots, \alpha_k$ de maneira tal que α_j seja β_j menos uma certa combinação linear de $\beta_1, \beta_2, \ldots, \beta_{j-1}$ para $2 \leq j \leq k$. Seja

$$\alpha_{k+1} = \beta_{k+1} - \sum_{j=1}^{k} \frac{(\beta_{k+1}, \alpha_j)}{||\alpha_j||^2}\,\alpha_j. \tag{8-9}$$

Então

$$\begin{aligned}(\alpha_{k+1}, \alpha_i) &= (\beta_{k+1}, \alpha_i) - \sum_{j=1}^{k} \frac{(\beta_{k+1}, \alpha_j)}{||\alpha_j||^2}\,(\alpha_j, \alpha_i)\\ &= (\beta_{k+1}, \alpha_i) - (\beta_{k+1}, \alpha_i)\\ &= 0, \quad 1 \leq i \leq k.\end{aligned}$$

Assim, α_{k+1} é ortogonal a cada um dos vetores $\alpha_1, \alpha_2, \ldots, \alpha_k$. Suponhamos que $\alpha_{k+1} = 0$. Então β_{k+1} seria uma combinação de $\alpha_1, \alpha_2, \ldots, \alpha_k$ e, portanto, de $\beta_1, \beta_2, \ldots, \beta_k$. Assim, $\alpha_{k+1} \neq 0$ e, continuando desta forma, obteremos no final um conjunto ortogonal não-nulo $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ contendo n vetores distintos. Pelo Teorema 2, êste conjunto é independente, logo é uma base. Para obter uma base ortonormal basta substituir cada α_i por $\dfrac{\alpha_i}{||\alpha_i||}$.

Êste teorema nos mostra que o número máximo de vetores em um subconjunto ortonormal de V é dim V. Êle também nos mostra outro fato, que deve ser mencionado.

Suponhamos que V seja um espaço vetorial de dimensão finita sôbre o corpo dos números reais ou sôbre o corpo dos números complexos. Na última seção ressaltamos que todos os produtos internos possíveis sôbre V podem ser descritos por meio de certas matrizes hermitianas. Com o Teorema 3, pode-se fazer uma descrição diferente de todos os produtos internos sôbre V. Uma maneira de definir um produto interno sôbre V é esta: Tomar uma base *arbitrária* $\{\alpha_1, \ldots, \alpha_n\}$ de V e definir o produto interno de $\alpha = x_1\alpha_1 + x_n\alpha_n$ e $\beta = y_1\alpha_1 + \ldots + y_n\alpha_n$ por

$$(\alpha, \beta) = \sum_{j=1}^{n} x_j\bar{y}_i.$$

Em relação êsse produto interno, $\{\alpha_1, \ldots, \alpha_n\}$ é uma base ortonormal. O que o Teorema 3 nos diz é que todo produto interno sôbre V surge desta maneira, isto é, que êste é o único tipo de produto interno que existe sôbre V. De fato, pois se (,) é um produto interno, existe uma base $\{\alpha_1, \ldots, \alpha_n\}$ de V que é ortonormal em relação a êste produto interno. Como $(\alpha_j, \alpha_k) = \delta_{jk}$, é imediato que

$$(\sum_j x_j\alpha_j, \sum_k y_k\alpha_k) = \sum_{j=1}^{n} x_j\bar{y}_j.$$

Definição. *Seja* V *um espaço com produto interno e seja* S *um conjunto arbitrário de vetores em* V. *O* **suplementar ortogonal** *de* S *é o conjunto* $S^{\perp}$ *de todos os vetores em* V *que não são ortogonais a todo vetor em* S.

Se S é um subconjunto arbitrário de V, o suplementar ortogonal $S^{\perp}$ (S perpendicular) é um subespaço de V. De fato, pois $S^{\perp}$ contém o vetor nulo e é portanto não-vazio. Sejam α e β em $S^{\perp}$ e seja c um escalar arbitrário. Para todo vetor α em S temos

$$\begin{aligned}(c\alpha + \beta, \gamma) &= c(\alpha, \gamma) + (\beta, \gamma)\\ &= c0 + 0\\ &= 0\end{aligned}$$

o que mostra que $c\alpha + \beta$ está em $S^{\perp}$. O suplementar ortogonal de V é o subespaço nulo, e reciprocamente.

Teorema 4. *Se* W *é um subespaço de dimensão finita de um espaço* V *com produto interno, então*

$$V = W \oplus W^{\perp}.$$

Demonstração. Seja $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ uma base ortogonal qualquer de W; a existência de uma tal base é garantida pelo Teorema 3. Se β é um vetor qualquer em V,

$$\beta = \sum_{k=1}^{m} \frac{(\beta, \alpha_k)}{\|\alpha_k\|^2} \alpha_k + \gamma$$

onde

$$\gamma = \beta - \sum_{k=1}^{m} \frac{(\beta, \alpha_k)}{\|\alpha_k\|^2} \alpha_k.$$

Como $(\gamma, \alpha_j) = 0$ para todo α_j, decorre que γ é ortogonal a tôda combinação linear de $\alpha_1, \alpha_2, \ldots, \alpha_m$; logo, $V = W + W^{\perp}$. Se α pertence a W e também a $W^{\perp}$, então $(\alpha, \alpha) = 0$, o que mostra que $\alpha = 0$. Assim, V é a soma direta de W e $W^{\perp}$.

Corolário. *Seja $\{\alpha_1, \ldots, \alpha_m\}$ um conjunto ortogonal de vetores não-nulos em um espaço* V *com produto interno. Se β é um vetor arbitrário em* V, *então*

$$\sum_{k=1}^{m} \frac{|(\beta, \alpha_k)|^2}{\|\alpha_k\|^2} \leq \|\beta\|^2.$$

Além disso, a igualdade vale se, e sòmente se,

$$\beta = \sum_{k=1}^{m} \frac{(\beta, \alpha_k)}{\|\alpha_k\|^2} \alpha_k.$$

Demonstração. Seja

$$\delta = \sum_{k=1}^{m} \frac{(\beta, \alpha_k)}{\|\alpha_k\|^2} \alpha_k.$$

Então $\beta = \delta + \gamma$, onde $(\delta, \gamma) = 0$. Logo

$$\|\beta\|^2 = (\delta + \gamma, \delta + \gamma) = \|\delta\|^2 + \|\gamma\|^2.$$

Agora basta demonstrar que

$$\|\delta\|^2 = \sum_{k=1}^{m} \frac{|(\beta, \alpha_k)|^2}{\|\alpha_k\|^2}.$$

Isto é um cálculo imediato, usando o fato de que $(\alpha_j, \alpha_k) = 0$ para $j \neq k$.

A desigualdade no corolário cima é conhecida como a **desigualdade de Bessel**. No caso particular em que $\{\alpha_1, \ldots, \alpha_m\}$ é um conjunto ortonormal, a desigualdade de Bessel diz

$$\sum_{k=1}^{m} |(\beta, \alpha_k)|^2 \leq \|\beta\|^2.$$

O corolário também nos diz que, neste caso, β está no subespaço gerado por $\alpha_1, \ldots, \alpha_m$ se, e sòmente se,

$$\beta = \sum_{k=1}^{m} (\beta, \alpha_k)\alpha_k$$

ou seja, se, e sòmente se, a desigualdade de Bessel é, na verdade, uma igualdade. É claro que, no caso de V ser de dimensão finita e $\{\alpha_1, \ldots, \alpha_m\}$ ser uma base ortonormal de V, a fórmula acima vale para todo vetor β em V. Em outras palavras, se $\{\alpha_1, \ldots, \alpha_m\}$ é uma base *ortonormal* de V, a k-ésima coordenada de β em relação à base ordenada $\{\alpha_1, \ldots, \alpha_m\}$ é (β, α_k).

Exemplo 11. Aplicaremos o último corolário aos conjuntos ortonormais descritos no Exemplo 10. Concluímos que

(a) $$\sum_{k=-n}^{n} \left| \int_0^1 f(t)e^{-2\pi ikt}dt \right|^2 \leq \int_0^1 |f(t)|^2 dt$$

(b) $$\int_0^1 \left| \sum_{k=-n}^{n} c_k e^{2\pi ikt} \right|^2 dt = \sum_{k=-n}^{n} |c_k|^2$$

(c) $$\int_0^1 (\sqrt{2}\cos 2\pi t + \sqrt{2}\,\mathrm{sen}\, 4\pi t)^2 dt = 1 + 1 = 2.$$

Se W é um subespaço de dimensão finita de um espaço V com produto interno, já mostramos que $V = W \oplus W^\perp$. Decorre que existe uma única projeção E com imagem W e núcleo $W^\perp$. Denominamos esta projeção E a **projeção ortogonal** de V sôbre W. Lembramos ao leitor que E é o operador linear definido como segue: Se α está em V, escrevamos $\alpha = \beta + \gamma$, com β em W e γ em $W^\perp$. Então $E\alpha = \beta$. Se $\{\alpha_1, \ldots, \alpha_m\}$ é uma base ortogonal arbitrária de W, a demonstração do Teorema 4 nos mostra uma fórmula explícita para a projeção ortogonal E em têrmos dos vetores $\alpha_1, \ldots, \alpha_m$:

(8-10) $$E\beta = \sum_{k=1}^{m} \frac{(\beta, \alpha_k)}{||\alpha_k||^2}\alpha_k.$$

Exercícios

1. Consideremos R^4 com o produto interno canônico. Seja W o subespaço de R^4 formado pelos vetores que são ortogonais a $\alpha = (1, 0, -1, 1)$ e a $\beta = (2, 3, -1, 2)$. Determinar uma base de W.

2. Aplicar o processo de Gram-Schmidt aos vetores $\beta_1 = (1, 0, 1)$, $\beta_2 = (1, 0, -1)$, $\beta_3 = (0, 3, 4)$, para obter uma base ortonormal de R^3 com o produto interno canônico.

3. Consideremos C^3 com o produto interno canônico. Determinar uma base ortonormal do subespaço gerado por $\beta_1 = (1, 0, i)$ e $\beta_2 = (2, 1, 1 + i)$.

4. Seja V um espaço com produto interno. A **distância** entre dois vetores α e β em V é definida por

$$d(\alpha, \beta_2) = ||\alpha - \beta||.$$

Mostrar que

(a) $d(\alpha, \beta) \geq 0$;
(b) $d(\alpha, \beta) = 0$ se, e sòmente se, $\alpha = \beta$;
(c) $d(\alpha, \beta) = d(\beta, \alpha)$;
(d) $d(\alpha, \beta) \leq d(\alpha, \gamma) + d(\gamma, \beta)$.

5. Seja V um espaço com produto interno e sejam α, β vetores em V. Mostrar que $\alpha = \beta$ se, e sòmente se, $(\alpha, \gamma) = (\beta, \gamma)$ para todo γ em V.

6. Seja W o subespaço de R^2 gerado pelo vetor (3, 4). Usando o produto interno canônico, seja E a projeção ortogonal de R^2 sôbre W. Determinar:

(a) uma fórmula para $E(x_1, x_2)$;
(b) a matriz de E em relação à base ordenada canônica;
(c) $W^\perp$;
(d) uma base ortonormal em relação à qual E seja representada pela matriz

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

7. Seja V o espaço com produto interno que consiste de R^2 com o produto interno cuja forma quadrática é definida por

$$||(x_1, x_2)||^2 = (x_1 - x_2)^2 + 3x_2^2.$$

Seja E a projeção ortogonal de V sôbre o subespaço W gerado pelo vetor (3, 4). Responder agora às quatro questões do Exercício 6.

8. Determinar um produto interno sôbre R^2 tal que $(\epsilon_1, \epsilon_2) = 2$.

9. Seja V um espaço com produto interno, W o subespaço gerado por um vetor não-nulo α e E a projeção ortogonal de V sôbre W. Se β é um vetor em V, mostrar que

$$||\beta - E\beta|| \leq ||\beta - \gamma||$$

para todo γ em W. O que isto diz geomètricamente?

10. Seja V o subespaço de $R[x]$ formado pelos polinômios de grau no máximo 3. Equipemos V com o produto interno

$$(f, g) = \int_0^1 f(t)g(t)dt.$$

(a) Determinar o suplementar ortogonal do subespaço dos polinômios constantes.

(b) Aplicar o processo de Gram-Schmidt à base $\{1, x, x^2, x^3\}$.

11. Seja V o espaço vetorial das $n \times n$ matrizes sôbre C, com o produto interno $(A, B) = \operatorname{tr}(AB^*)$. Determinar o suplementar ortogonal do subespaço das matrizes diagonais.

12. Seja V um espaço de dimensão finita com produto interno e seja $\{\alpha_1, \ldots, \alpha_n\}$ uma base ortonormal de V. Mostrar que para quaisquer vetores α, β em V

$$(\alpha, \beta) = \sum_{k=1}^{n} (\alpha, \alpha_k)\overline{(\beta, \alpha_k)}.$$

13. Seja W um subespaço de dimensão finita de um espaço V com produto interno e seja E a projeção ortogonal de V sôbre W. Demonstrar que $(E\alpha, \beta) = (\alpha, E\beta)$ para todos α, β em V.

14. Seja S um subespaço de um espaço V com produto interno. Mostrar que $(S^{\perp})^{\perp}$ contém o subespaço gerado por S. Para V de dimensão finita, mostrar que $(S^{\perp})^{\perp}$ é o subespaço gerado por S.

15. Seja V um espaço de dimensão finita com produto interno e seja $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ uma base ortonormal de V. Seja T um operador linear sôbre V e A a matriz de T em relação à base ordenada $\mathcal{B}$. Demonstrar que

$$A_{ij} = (T\alpha_j, \alpha_i).$$

16. Suponhamos que $V = W_1 \oplus W_2$ e que f_1 e f_2 sejam produtos internos sôbre W_1 e W_2 respectivamente. Mostrar que existe um único produto interno f sôbre V tal que

(i) $W_2 = W_1{}^{\perp}$;

(ii) $f(\alpha, \beta) = f_k(\alpha, \beta)$, quando α, β estão em W_k, $k = 1, 2$.

17. Seja V um espaço com produto interno e W um subespaço de V de dimensão finita. Existem (em geral) muitas projeções que têm W por sua imagem. Uma destas, a projeção ortogonal sôbre W, tem a propriedade de que $||E\alpha|| \leq ||\alpha||$ para todo α em V. Demonstrar que se E é uma projeção com imagem W, tal que $||E\alpha|| \leq ||\alpha||$ para todo α em V, então E é a projeção ortogonal sôbre W.

18. Seja V o espaço real com produto interno que consiste do espaço das funções contínuas, definidas no intervalo $-1 \leq t \leq 1$, tomando valores reais, com o produto interno

$$(f, g) = \int_{-1}^{1} f(t)g(t)dt.$$

Seja W o subespaço das funções ímpares, isto é, funções que satisfazem $f(-t) = -f(t)$. Determinar o suplementar ortogonal de W.

8.3 Funcionais Lineares e Adjuntos

A primeira parte desta seção trata dos funcionais lineares sôbre um espaço com produto interno e de sua relação com o produto interno. O resultado fundamental é que todo funcional linear f sôbre um espaço de dimensão finita com produto interno é o "produto interno por um vetor fixo no espaço", isto é, que um tal f é da forma $f(\alpha) = (\alpha,\beta)$, para um certo β fixo em V. Usaremos êste resultado para demonstrar a existência do "adjunto" de um operador linear T sôbre V, sendo êste um operador linear T^* tal que $(T\alpha,\beta) = (\alpha,T^*\beta)$ para todos α e β em V. Através do uso de uma base ortonormal, esta operação de conjugação sôbre operadores lineares (passando de T a T^*) é identificada com a operação de se tomar a transposta conjugada de uma matriz. Vamos explorar superficialmente a analogia entre a operação de conjugação e a conjugação sôbre números complexos.

Seja V um espaço arbitrário com produto interno e seja β um certo vetor fixo em V. Definamos uma função f_β de V no corpo de escalares por

$$f_\beta(\alpha) = (\alpha,\beta).$$

Esta função f_β é um funcional linear sôbre V, pois, por sua própria definição, (α,β) é linear como uma função de α. Se V é de dimensão finita, todo funcional linear sôbre V provém desta maneira de algum β.

Teorema 5. *Seja* V *um espaço de dimensão finita com produto interno e* f *um funcional linear sôbre* V. *Então existe um único vetor* β *em* V *tal que* $f(\alpha) = (\alpha,\beta)$ *para todo* α *em* V.

Demonstração. Seja $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ um base ortonormal de V. Coloquemos

$$\beta = \sum_{j=1}^{n} \overline{f(\alpha_j)}\alpha_j \tag{8-11}$$

e seja f_β o funcional linear definido por

$$f_\beta(\alpha) = (\alpha,\beta).$$

Então

$$f_\beta(\alpha_k) = (\alpha_k, \sum_j \overline{f(\alpha_j)}\alpha_j) = f(\alpha_k).$$

Como isto é válido para todo α_k, decorre que $f = f_\beta$. Suponhamos agora que γ seja um vetor em V tal que $(\alpha,\beta) = (\alpha,\gamma)$ para todo α. Então $(\beta - \gamma,\beta - \gamma) = 0$ e $\beta = \gamma$. Assim, existe exatamente um vetor β que determina o funcional linear f da maneira afirmada.

A demonstração dêste teorema pode ser ligeiramente reformulada, em têrmos da representação de funcionais lineares em relação a uma base. Se tomarmos uma base ortonormal $\{\alpha_1, \ldots, \alpha_n\}$ de V, o produto interno de $\alpha = x_1\alpha_1 + \ldots + x_n\alpha_n$ e $\beta = y_1\alpha_1 + \ldots + y_n\alpha_n$ será

$$(\alpha,\beta) = x_1\bar{y}_1 + \ldots + x_n\bar{y}_n.$$

Se f é um funcional linear arbitrário sôbre V, então f é da forma

$$f(\alpha) = c_1x_1 + \ldots + c_nx_n$$

para certos escalares fixos $c_1, \ldots, c_n$ determinados pela base. É claro que $c_j = f(\alpha_j)$. Se desejamos encontrar um vetor β em V tal que $(\alpha,\beta) = f(\alpha)$ para todo α, então evidentemente as coordenadas y_j de β devem satisfazer $\bar{y}_j = c_j$, ou seja, $y_j = \overline{f(\alpha_j)}$. Conseqüentemente

$$\beta = \overline{f(\alpha_1)}\alpha_1 + \ldots + \overline{f(\alpha_n)}\alpha_n$$

é o vetor desejado.

Exemplo 12. Gostaríamos de dar um exemplo que mostre que o Teorema 5 não é válido sem a hipótese de V ser de dimensão finita. Seja V o espaço vetorial dos polinômios sôbre o corpo dos números complexos, com o produto interno

$$(f, g) = \int_0^1 {}_z f(t)\overline{g(t)}dt.$$

Êste produto interno pode também ser definido algèbricamente. Se $f = \Sigma\, a_k x^k$ e $g = \Sigma\, b_k x^k$, então

$$(f, g) = \sum_{j,k} \frac{1}{j + k + 1}\, a_j \overline{b_k}.$$

Seja z um número complexo fixo e seja L o funcional linear "valor que assume em z":

$$L(f) = f(z).$$

Existe um polinômio g tal que $(f, g) = L(f)$ para todo f? A resposta é negativa; de fato, suponhamos que se tenha

$$f(z) = \int_0^1 f(t)\overline{g(t)}dt$$

para todo f. Seja $h = x - z$, de modo que para todo f temos $(hf)(z) = 0$. Então

$$0 = \int_0^1 h(t)f(t)\overline{(gt)}dt$$

para todo f. Em particular, isto vale para $f = \bar{h}g$ de modo que

$$\int_0^1 |h(t)|^2|g(t)|^2dt = 0$$

e então $hg = 0$. Como $h \neq 0$, devemos ter que $g = 0$. Mas L não é o funcional nulo; logo, nenhum tal g existe.

Pode-se, num certo sentido, generalizar o exemplo, para o caso em que L é uma combinação linear de funcionais do tipo acima. Suponhamos que tomemos números complexos fixos $z_1, \ldots, z_n$ e escalares $c_1, \ldots, c_n$ e seja

$$L(f) = c_1 f(z_1) + \ldots + c_n f(z_n).$$

Então L é um funcional linear sôbre V, mas não existe nenhum g tal que $L(f) = (f, g)$, a menos que $c_1 = c_2 = \ldots = c_n = 0$. Bastar repetir o argumento acima com $h = (x - z_1) \ldots (x - z_n)$.

Voltamos agora ao conceito do adjunto de um operador linear. Demonstraremos primeiro o que segue.

Teorema 6. *Para qualquer operador linear* T *sôbre um espaço* V *de dimensão finita com produto interno, existe um único operador linear* T* *sôbre* V *tal que*

$$(8\text{-}12) \qquad (T\alpha, \beta) = (\alpha, T^*\beta)$$

para todos α,β *em* V.

Demonstração. Demonstraremos que existe um tal operador linear T^*. Seja β um vetor em V; definiremos $T^*\beta$. Ora, $f(\alpha) = (T\alpha,\beta)$ é um funcional linear sôbre V e o Teorema 5 nos diz que existe um vetor β' em V tal que $(T\alpha,\beta) = (\alpha,\beta')$ para todo α. Êste β' é determinado de modo único por β e a regra que associa β' a β será chamada de T^*:

$$\beta' = T^*\beta.$$

Temos (8-12), mas precisamos verificar que T^* é um operador linear. Sejam β, γ em V e seja c um escalar. Então, para qualquer α

$$\begin{aligned}
(\alpha, T^*(c\beta + \gamma)) &= (T\alpha, c\beta + \gamma) \\
&= (T\alpha, c\beta) + (T\alpha, \gamma) \\
&= \bar{c}(T\alpha, \beta) + (T\alpha, \gamma) \\
&= \bar{c}(\alpha, T^*\beta) + (\alpha, T^*\gamma) \\
&= (\alpha, cT^*\beta) + (\alpha, T^*\gamma) \\
&= (\alpha, cT^*\beta + T^*\gamma).
\end{aligned}$$

Assim, $T^*(c\beta + \gamma) = cT^*\beta + T^*\gamma$ e T^* é linear.

A unicidade de T^* é evidente. Para β arbitrário em V, o vetor $T^*\beta$ é determinado de modo único como sendo o vetor β' tal que $(T\alpha, \beta) = (\alpha, \beta')$ para todo α.

Teorema 7. *Seja* V *um espaço de dimensão finita com produto interno e seja* $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ *uma base ortonormal (ordenada) de* V. *Seja* T *um operador linear sôbre* V *e seja* A *a matriz de* T *em relação à base ordenada* $\mathcal{B}$. *Então* $A_{kj} = (T\alpha_j, \alpha_k)$.

Demonstração. Como $\mathcal{B}$ é uma base ortonormal, temos

$$\alpha = \sum_{k=1}^{n} (\alpha, \alpha_k)\alpha_k.$$

A matriz A é definida por

$$T\alpha_j = \sum_{k=1}^{n} A_{kj}\alpha_k$$

e como

$$T\alpha_j = \sum_{k=1}^{n} (T\alpha_j, \alpha_k)\alpha_k$$

temos $A_{kj} = (T\alpha_j, \alpha_k)$.

Corolário. *Seja* V *um espaço de dimensão finita com produto interno e seja* T *um operador linear sôbre* V. *Em relação a qualquer base ortonormal de* V, *a matriz de* T* *é a transposta conjugada da matriz de* T.

Demonstração. Seja $\mathfrak{B} = \{\alpha_1, \ldots, \alpha_n\}$ uma base ortonormal de V, seja $A = [T]_{\mathfrak{B}}$ e $B = [T^*]_{\mathfrak{B}}$. De acôrdo com o Teorema 7,

$$\begin{aligned} A_{kj} &= (T\alpha_j, \alpha_k) \\ B_{kj} &= (T^*\alpha_j, \alpha_k). \end{aligned}$$

Então, pela definição de T^*, temos

$$\begin{aligned} B_{kj} &= (T^*\alpha_j, \alpha_k) \\ &= \overline{(\alpha_k, T^*\alpha_j)} \\ &= \overline{(T\alpha_k, \alpha_j)} \\ &= \overline{A_{jk}}. \end{aligned}$$

Definição. *Seja* T *um operador linear sôbre um espaço* V *com produto interno. Dizemos que* T **possui um adjunto** *se existe um operador linear* T* *sôbre* V *tal que* $(T\alpha, \beta) = (\alpha, T^*\beta)$ *para todos* α *e* β *em* V.

A afirmação do Teorema 6 é que todo operador linear sôbre um espaço de dimensão finita com produto interno possui um adjunto. Isto não mais vale se V não é de dimensão finita. Em qualquer caso, existe no máximo um tal operador T^*; quando existe, denominamo-lo o **adjunto** de T.

Dois comentários devem ser feitos acêrca do caso de dimensão finita. (1) O adjunto de T depende não só de T, mas também do produto interno. (2) Para um base ordenada arbitrária $\mathfrak{B}$, a relação $[T]_{\mathfrak{B}}$ e $[T^*]_{\mathfrak{B}}$ é mais complicada que a apresentada no corolário acima.

Exemplo 13. Seja V o espaço das $n \times 1$ matrizes complexas, com o produto interno $(X, Y) = Y^*X$. Se A é uma $n \times n$ matriz com elementos complexos, o adjunto do operador linear $X \to AX$ é o operador $X \to A^*X$. De fato, pois

$$(AX, Y) = Y^*AX = (A^*Y)^*X = (X, A^*Y).$$

O leitor deverá se convencer de que isto é na realidade uma reformulação do último corolário.

Exemplo 14. Êste é semelhante ao Exemplo 13. Seja V o espaço das $n \times n$ matrizes complexas, com produto interno $(A, B) = \operatorname{tr}(B^*A)$. Seja M uma $n \times n$ matriz fixa sôbre C. O adjunto da multiplicação à esquerda por M é a multiplicação à esquerda por M^*. Evidentemente, "multiplicação à esquerda por M" é o operador linear L_M definido por $L_M(A) = MA$.

$$\begin{aligned}(L_M(A), B) &= \operatorname{tr}(B^*(MA))\\ &= \operatorname{tr}(MAB^*)\\ &= \operatorname{tr}(AB^*M)\\ &= \operatorname{tr}(A(M^*B)^*)\\ &= (A, L_{M^*}(B)).\end{aligned}$$

Assim, $(L_M)^* = L_{M^*}$. No cálculo acima, usamos duas vêzes a propriedade característica da função traço: $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Exemplo 15. Seja V o espaço dos polinômios sôbre o corpo dos números complexos, com o produto interno

$$(f, g) = \int_0^1 f(t)\overline{g(t)}dt.$$

Se f é um polinômio, $f = \Sigma\, a_k x^k$, seja $\bar{f} = \Sigma\, \bar{a}_k x^k$. Isto é, $\bar{f}$ é o polinômio cuja função polinomial associada é a complexa conjugada da de f:

$$\bar{f}(t) = \overline{f(t)}, \quad t \text{ real}$$

Consideremos o operador "multiplicação por f", isto é, o operador linear m_f definido por $m_f(g) = fg$. Então êste operador possui um adjunto, a saber, a multiplicação por $\bar{f}$. De fato, pois

$$\begin{aligned}(M_f(g), h) &= (fg, h)\\ &= \int_0^1 f(t)g(t)\overline{h(t)}dt\\ &= \int_0^1 g(t)[\overline{\bar{f}(t)h(t)}]dt\\ &= (g, \bar{f}h)\\ &= (g, M_{\bar{f}}(h))\end{aligned}$$

e portanto $(M_{\bar{f}})^* = M_f$.

Exemplo 16. No Exemplo 15, vimos que alguns operadores lineares sôbre um espaço de dimensão finita com produto interno possuem um adjunto. Como comentamos anteriormente, outros não o têm. Seja V o espaço com produto interno do Exemplo 15 e seja D o operador derivação sôbre $C[x]$. A integração por partes mostra que

$$(Df, g) = f(1)g(1) - f(0)g(0) - (f, Dg).$$

Fixemos g e perguntemos quando é que existe um polinômio D^*g tal que $(Df, g) = (f, D^*g)$ para todo f. Se um tal D^*g existe, temos

$$(f, D^*g) = f(1)g(1) - f(0)g(0) - (f, Dg)$$

ou seja, $$(f, D^*g + Dg) = f(1)g(1) - f(0)g(0).$$

Com g fixo, $L(f) = f(1)g(1) - f(0)g(0)$ é um funcional linear do tipo considerado no Exemplo 12 e não pode ser da forma $L(f) = (f, h)$ a menos que $L = 0$. Se D^*g existe, então com $h = D^*g + Dg$ temos de fato $L(f) = (f, h)$, e então $g(0) = g(1) = 0$. A existência de um polinômio adequado D^*g implica $g(0) = g(1) = 0$. Reciprocamente, se $g(0) = g(1) = 0$, o polinômio $D^*g = - Dg$ satisfaz $(Df, g) = (f, D^*g)$ para todo f. Tomando um g qualquer para o qual $g(0) \neq 0$ ou $g(1) \neq 0$, não podemos definir D^*g de modo conveniente, portanto, concluímos que D não possui adjunto.

Esperamos que êstes exemplos aumentem a compreensão do leitor quanto ao adjunto de um operador linear. Vemos que a operação de conjugação, que faz passar de T a T^*, se comporta um pouco como a conjugação sôbre números complexos. O teorema seguinte fortalece esta analogia.

Teorema 8. *Seja* V *um espaço de dimensão finita com produto interno. Se* T *e* U *são operadores lineares sôbre* V *e* c *é um escalar*

(a) $(T + U)^* = T^* + U^*$;

(b) $(cT)^* = \bar{c}T^*$;

(c) $(TU)^* = U^*T^*$;

(d) $(T^*)^* = T$.

Demonstração. Para demonstrar (a)

$$\begin{aligned}((T + U)\alpha, \beta) &= (T\alpha + U\alpha, \beta)\\ &= (T\alpha, \beta) + (U\alpha, \beta)\\ &= (\alpha, T^*\beta) + (\alpha, U^*\beta)\\ &= (\alpha, T^*\beta + U^*\beta)\\ &= (\alpha, (T^* + U^*)\beta).\end{aligned}$$

Pela unicidade do adjunto temos $(T + U)^* = T^* + U^*$. Deixamos a demonstração de (b) a cargo do leitor. Obtemos (c) e (d) a partir de

$$(TU\alpha, \beta) = (U\alpha, T^*\beta) = (\alpha, U^*T^*\beta)$$
$$(T^*\alpha, \beta) = \overline{(\beta, T^*\alpha)} = \overline{(T\beta, \alpha)} = (\alpha, T\beta).$$

O Teorema 8 é freqüentemente formulado como segue: a aplicação $T \to T^*$ é um anti-isomorfismo linear-conjugado de período 2. A analogia com a conjugação complexa que mencionamos acima é, evidentemente, baseada na observação de que a conjugação complexa tem as propriedades $\overline{(z_1 + z_2)} = \bar{z}_1 + \bar{z}_2$, $\overline{(z_1 z_2)} = \bar{z}_1\bar{z}_2, \bar{\bar{z}} = z$. Deve-se ter o cuidado de observar a inversão da ordem num produto, imposta pela operação de conjugação: $(UT)^* = T^*U^*$. Mencionaremos extensões da analogia à medida que prossigamos nosso estudo de operadores lineares sôbre um espaço com produto interno. Podemos mencionar alguma coisa nesse sentido agora. Um número complexo z é real se, e sòmente se, $z = \bar{z}$. É de esperar que os operadores lineares T tais que $T = T^*$ se comportem, de certa maneira, como os números reais. É isto o que realmente ocorre. Por exemplo, se T é um operador linear sôbre um espaço *complexo* de dimensão finita com produto interno, então

$$T = U_1 + iU_2 \tag{8-13}$$

onde $U_1 = U^*_1$ e $U_2 = U^*_2$. Assim, de certa forma, T possui uma "parte real" e uma "parte imaginária". Os operadores U_1 e U_2 que satisfazem $U_1 = U^*_2$, $U_2 = U^*_1$ e (8-13) são únicos e são dados por

$$U_1 = \frac{1}{2}(T + T^*)$$

$$U_2 = \frac{1}{2i}(T - T^*).$$

Um operador linear T tal que $T = T^*$ é dito **auto-adjunto** ou **(hermitiano)**. Se $\mathcal{B}$ é uma base ortonormal de V, então

$$[T^*]_{\mathcal{B}} = [T]^*_{\mathcal{B}}$$

e então T é auto-adjunto se, e sòmente se, sua matriz em relação a tôda base ortonormal é uma matriz auto-adjunta. Operadores auto-adjuntos são importantes, não só porque nos fornecem uma espécie de partes real e imaginária de um operador linear arbitrário, mas também pelas seguintes razões: (1) Operadores auto-adjuntos possuem muitas propriedades especiais. Por exemplo, para um operador dêste tipo, existe uma base ortonormal formada por vetores característicos. (2) Muitos operadores que surgem na prática são auto-adjuntos. Consideraremos posteriormente as propriedades especiais dos operadores auto-adjuntos.

Exercícios

1. Seja V o espaço C^2, com o produto interno canônico. Seja T o operador definido por $T\epsilon_1 = (1, -2)$, $T\epsilon_2 = (i, -1)$. Se $\alpha = (x^1, x_2)$, determinar $T^*\alpha$.

2. Seja T o operador linear sôbre C^2 definido por $T\epsilon_1 = (1 + i, 2)$, $T\epsilon_2 = (i, i)$. Usando o produto interno canônico, determinar a matriz de T^* em relação à base ordenada canônica. T comuta com T^*?

3. Suponhamos que V seja C^2 com o produto interno canônico. Seja T o operador linear sôbre V cuja matriz em relação à base ordenada canônica é definida por

$$A_{jk} = i^{j+k}, \quad (i^2 = -1).$$

Determinar uma base do núcleo de T^*.

4. Seja V um espaço de dimensão finita com produto interno e T um operador linear sôbre V. Se T é inversível, mostrar que T^* é inversível e $(T^*)^{-1} = (T^{-1})^*$.

5. Seja V um espaço com produto interno e β, γ vetores fixos em V. Mostrar que $T\alpha = (\alpha, \beta)\gamma$ define um operador linear sôbre V. Mostrar que T possui um adjunto e descrever T^* explìcitamente.

Suponhamos agora que V seja C^n com o produto interno canônico, $\beta = (y_1, \ldots, y_n)$ e $\gamma = (x_1, \ldots, x_n)$. Qual é o elemento j, k da matriz de T em relação à base ordenada canônica? Qual é o pôsto desta matriz?

6. Mostrar que o produto de dois operadores auto-adjuntos é auto-adjunto se, e sòmente se, os dois operadores comutam.

7. Seja V o espaço vetorial dos polinômios sôbre R de grau menor ou igual a 3, com o produto interno

$$(f, g) = \int_0^1 f(t)g(t)dt.$$

Se t é um número real, determinar o polinômio g_t em V tal que $(f, g_t) = (ft)$ para todo f em V.

8. Seja V o espaço com produto interno do Exercício 7 e seja D operador derivação sôbre V. Determinar D^*.

9. Seja V um espaço de dimensão finita com produto interno e T um operador linear sôbre V. Mostrar que a imagem de T^* é o suplementar ortogonal do núcleo de T.

10. Seja V o espaço das $n \times n$ matrizes sôbre o corpo dos números complexos, com o produto interno $(A, B) = \operatorname{tr}(AB^*)$. Seja P uma matriz inversível fixa em V e seja T_P o operador linear sôbre V definido por $T_P(A) = P^{-1}AP$. Determinar o adjunto de T_P.

11. Seja V um espaço de dimensão finita com produto interno e seja E um operador linear idempotente sôbre V, isto é, $E^2 = E$. Demonstrar que E é auto-adjunto se, e sòmente se, $EE^* — E^*E$.

12. Seja V um espaço *complexo* de dimensão finita com produto interno e seja T um operador linear sôbre V. Demonstrar que T é auto-adjunto se, e sòmente se, $(T\alpha, \alpha)$ é real para todo α em V.

8.4 Operadores Positivos

Seja V um espaço com produto interno e seja T um operador linear sôbre V. Seja p a função definida sôbre os pares ordenados de vetores α, β em V por

$$p(\alpha, \beta) = (T\alpha, \beta).$$

Estamos interessados em obter condições necessárias e suficientes que T precisa satisfazer para p ser um produto interno. Certamente $p(\alpha, \beta)$ é linear como uma função de α, portanto o que precisamos fazer é traduzir as propriedades

$$\begin{aligned} &p(\alpha, \beta) = \overline{p(\beta, \alpha)} \\ &p(\alpha, \alpha) > 0, \quad \text{para } \alpha \neq 0 \end{aligned}$$

em afirmações sôbre T. Ora, $p(\alpha, \beta) = (T\alpha, \beta)$ e $\overline{p(\beta, \alpha)} = \overline{(T\beta, \alpha)} = (\alpha, T\beta)$. Como $p(\alpha, \alpha) = (T\alpha, \alpha)$, vemos que a função p é um produto interno se, e sòmente se, o operador linear T satisfaz

$$\begin{aligned} &(T\alpha, \beta) = (\alpha, T\beta) \\ &(T\alpha, \alpha) > 0 \quad \text{se } \alpha \neq 0. \end{aligned} \tag{8-14}$$

A primeira condição diz que T é auto-adjunto. Um operador linear T que satisfaz as condições (8-14) é dito **positivo**. O que acabamos de observar pode ser reformulado como segue: O operador linear T é positivo, se, e sòmente se, a função p definida por $p(\alpha, \beta) = (T\alpha, \beta)$ é um produto interno.

Desejamos agora mostrar que, no caso de dimensão finita, todo produto interno sôbre V é do tipo há pouco descrito.

Teorema 9. *Seja* V *um espaço de dimensão finita com produto interno* (,). *Se* p *é um produto interno arbitrário sôbre* V, *existe um único operador linear positivo* T *sôbre* V *tal que* $p(\alpha, \beta) = (T\alpha, \beta)$ *para quaisquer* α, β *em* V.

Demonstração. Fixemos um vetor β em V. Então $p(\alpha, \beta)$ é uma função linear de α. Pelo Teorema 5 existe um único vetor β' em V tal que $p(\alpha, \beta) = (\alpha, \beta')$ para todo α. Definamos uma função T de V em V por $T\beta = B'$. Em vista da maneira como T é definida, temos $p(\alpha, \beta) = \overline{(\alpha, T\beta)}$ para todos α, β em V. Temos também $p(\alpha, \beta) = (\alpha, T\beta) = \overline{p(\beta, \alpha)} = \overline{(\beta, T\alpha)} = (T\alpha, \beta)$ para quaisquer α, β em V. Para mostrar que T é linear, observemos que

$$\begin{aligned}(T(c\,\alpha + \beta), \gamma) &= p(c\,\alpha + \beta, \gamma)\\ &= cp(\alpha, \gamma) + p(\beta, \gamma)\\ &= c(T\alpha, \gamma) + (T\beta, \gamma)\\ &= (cT\alpha + T\beta, \gamma)\end{aligned}$$

para todos α, β, γ em V e todos escalares c. Logo, $T(c\alpha + \beta) = cT\alpha + T\beta$. Então, demonstraremos a existência de um operador linear T tal que $p(\alpha, \beta) = (T\alpha, \beta)$. É claro que T é positivo, pois p é um produto interno. Precisamos mostrar que T é único. Suponhamos que $p(\alpha, \beta) = (U\alpha, \beta)$. Então

$$(T\alpha, \beta) = (U\alpha, \beta), \text{ ou } (T\alpha - U\alpha, \beta) = 0$$

para todos α e β. Para α fixo, o vetor $T\alpha - U\alpha$ é ortogonal a todo o espaço V e é portanto o vetor nulo; logo $T\alpha = U\alpha$ para todo α.

Sabemos agora que todos os produtos internos sôbre um espaço de V de dimensão finita com produto interno podem ser descritos em têrmos dos operadores lineares positivos sôbre V. Passamos então a estudar os operadores lineares positivos de modo a tornar mais significativa nossa descrição de produtos internos. Alguns dêstes resultados deverão tornar claras as razões para o nome "operador positivo".

Teorema 10. *Seja* V *um espaço de dimensão finita com produto interno e* T *um operador linear sôbre* V. *Então* T *é positivo se, e sòmente se, existe um operador linear inversível* U *sôbre* V *tal que* T = U*U.

Demonstração. Suponhamos que $T = U^*U$, onde U é operador linear inversível sôbre V. Então $T^* = (U^*U)^* = U^*U = T$, de modo que T é auto-adjunto. Além disso, $(T\alpha,\alpha) = (U^*U\alpha, \alpha) = (U\alpha, U\alpha) \geq 0$. Ora, U é inversível; portanto se $\alpha \neq 0$ temos $U\alpha \neq 0$ e $(T\alpha, \alpha) > 0$. Então T é positivo.

Suponhamos agora que T seja positivo. Então $p(\alpha, \beta) = (T\alpha, \beta)$ é um produto interno sôbre V. Seja $\{\alpha_1, \ldots, \alpha_n\}$ uma base de V que seja ortonormal em relação ao produto interno $(\ ,\ )$ e seja $\{\beta_1, \ldots, \beta_n\}$ uma base ortonormal em relação a p. Então

$$p(\beta_j, \beta_k) = \delta_{jk} = (\alpha_j, \alpha_k).$$

Seja agora U o único operador linear sôbre V tal que $U\beta_j = \alpha_j$,

$j = 1, \ldots, n$. Por levar uma base em outra base U é inversível. Temos

$$p(\beta_j, \beta_k) = (U\beta_j, U\beta_k) = (\alpha_j, \alpha_k).$$

Sejam $\alpha = \Sigma\, x_j\beta_j$ e $\beta = \Sigma\, y_j\beta_j$ vetores arbitrários em V. Então

$$\begin{aligned} p(\alpha, \beta) &= p(\sum_j x_j\beta_j, \sum_k y_k\beta_k) \\ &= \sum_j \sum_k x_j\bar{y}_k p(\beta_j, \beta_k) \\ &= \sum_j \sum_k x_j\bar{y}_k(U\beta_j, U\beta_k) \\ &= (\sum_j x_jU\beta_j, \sum_k y_kU\beta_k) \\ &= (U\alpha, U\beta). \end{aligned}$$

Pela definição de p temos $((T\alpha, \beta) = (U\alpha, U\beta) = (U^*U\alpha, \beta)$ para todos α, β em V. Assim, $T = U^*U$.

Se considerarmos a operação de conjugação como sendo análoga à conjugação sôbre números complexos, veremos que "operador positivo" é mais ou menos análogo a "número positivo". De fato, um número complexo z é positivo se, e sòmente se, é da forma $z = \bar{w}w$ para algum número complexo não-nulo w.

É interessante o fato de que num espaço *complexo* com produto interno, a condição $T = T^*$ pode ser retirada da definição de um operador linear positivo. Em outras palavras, se $(T\alpha, \alpha) > 0$ para $\alpha \neq 0$, então T é necessàriamente auto-adjunto. Isto decorre do lema seguinte:

Lema. *Seja* V *um espaço complexo com produto interno e* T *um operador linear sôbre* V. *Se* $(T\alpha, \alpha)$ *é real para todo* α *em* V, *então* T *é auto-adjunto.*

Demonstração. Sejam α e β vetores em V. Precisamos mostrar que $(T\alpha, \beta) = (\alpha, T\beta)$. Ora,

$$(T(\alpha + \beta), \alpha + \beta) = (T\alpha, \alpha) + (T\alpha, \beta) + (T\beta, \alpha) + (T\beta, \beta).$$

Como $(T(\alpha + \beta), \alpha + \beta)$, $(T\alpha, \alpha)$ e $(T\beta, \beta)$ são reais, o número $(T\alpha, \beta) + (T\beta, \alpha)$ é real. Usando o mesmo argumento para $\alpha + i\beta$ em vez de $\alpha + \beta$, temos,

$$(T(\alpha + i\beta), \alpha + i\beta) = (T\alpha, \alpha) - i(T\alpha, \beta) + i(T\beta, \alpha) + (T\beta, \beta)$$

e então concluímos que $-i(T\alpha, \beta) + i(T\beta, \alpha)$ é real. Tendo concluído que dois números são reais, igualamo-los aos seus complexos conjugados e obtemos

$$(T\alpha, \beta) + (T\beta, \alpha) = (\beta, T\alpha) + (\alpha, T\beta)$$
$$-i(T\alpha, \beta) + i(T\beta, \alpha) = i(\beta, T\alpha) - i(\alpha, T\beta).$$

Multiplicando a segunda equação por i e somando o resultado à primeira equação, obtemos

$$2(T\alpha, \beta) = 2(\alpha, T\beta).$$

Evidentemente, êste lema é falso para um espaço real com produto interno, onde $(T\alpha, \alpha)$ é real para qualquer T. Se T é auto-adjunto, então $(T\alpha, \alpha)$ é real porque $(T\alpha, \alpha) = (\alpha, T\alpha) = \overline{(T\alpha, \alpha)}$. Assim, num espaço complexo com produto interno, os operadores auto-adjuntos são caracterizados pelo fato de que $(T\alpha, \alpha)$ é real para todo α e os operadores positivos pelo fato de que $(T\alpha, \alpha)$ é positivo para todo α não-nulo.

A fim de obter informações mais detalhadas a respeito de operadores lineares positivos, observemos a matriz de um tal operador T em relação a uma base ortonormal (ordenada) $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$. Se A é a matriz de T em relação à base ordenada $\mathcal{B}$, a condição $T = T^*$ diz simplesmente que $A = A^*$, ou seja, que $A_{jk} = \overline{A_{jk}}$. Lembramos ao leitor que isto usa, de modo fundamental, o fato de $\mathcal{B}$ ser ortonormal e o conseqüente fato de que $A_{jk} = (T\alpha_k, \alpha_j)$. A condição $(T\alpha, \alpha) > 0$ é fàcilmente interpretada em têrmos de A. Se $\alpha = x_1\alpha_1 + \ldots + x_n\alpha_n$

$$\begin{aligned}(T\alpha, \alpha) &= \Sigma_j(x_jT\alpha_j, \sum_k x_k\alpha_k)\\ &= \sum_j \sum_k x_j\bar{x}_k (T\alpha_j, \alpha_k)\\ &= \sum_j \sum_k A_{kj}x_j\bar{x}_k.\end{aligned}$$

Portanto, vemos que T é positivo se, e sòmente se, sua matriz em relação à base ortonormal $\mathcal{B}$ satisfaz

$$\text{(8-15)}\qquad \begin{aligned}&A = A^*\\ &\sum_j \sum_k A_{kj}x_j\bar{x}_k > 0, \quad \text{se} \quad (x_1, \ldots, x_n) \neq 0.\end{aligned}$$

Definição. *Seja* A *uma* n $\times$ n *matriz com elementos complexos. Dizemos que* A *é* **positiva** *se vale o seguinte: Sempre que* $x_1, \ldots, x_n$ *são números complexos, não todos nulos, então*

$$\sum_j \sum_k A_{kj}x_j\bar{x}_k > 0.$$

Teorema 11. *Seja* V *um espaço de dimensão finita com produto interno e* $\mathcal{B}$ *uma base ortonormal ordenada de* V. *Se* T *é um operador linear sôbre* V, *então* T *é positivo se, e sòmente se, a matriz de* T *em relação a esta base ordenada é positiva.*

Demonstração. Na verdade definimos uma matriz positiva de maneira tal que êste teorema valesse. No entanto, existem alguns pontos que decididamente requerem comentários.

Primeiro, suponhamos que V seja um espaço complexo com produto interno. Se A é a matriz de T em relação a $\mathcal{B}$, a afirmação de que A é positiva diz simplesmente que $(T\alpha, \alpha) > 0$ para todo α em V. Como uma conseqüêndia do último lema, observamos que esta condição vale se, e sòmente se, T é positivo. Êste é o argumento completo para o caso complexo. Notemos que uma conseqüência disto é que uma matriz positiva é automàticamente auto-adjunta.

Suponhamos agora que V seja um espaço real com produto interno e seja $A = [T]_{\mathcal{B}}$. Precisamos mostrar que se A é positiva, então T é positivo e que se T é positivo, então A é positiva. Em cada uma destas demonstrações surge um ponto sutil, que não ocorrera no caso complexo. Suponhamos que A seja positiva. Como estamos em um espaço real com produto interno, A é uma matriz com elementos reais; contudo, a afirmação de que A é positiva significa simplesmente que

$$\sum_j \sum_k A_{kj} x_j \bar{x}_k > 0 \tag{8-16}$$

para quaisquer números *complexos* $x_1, \ldots, x_n$ não todos nulos. Como isto é válido em particular quando $x_1, \ldots, x_n$ são reais, o argumento que conduziu a (8-15) mostra que $(T\alpha, \alpha) > 0$ para qualquer vetor não-nulo α em V. Mas num espaço real com produto interno, isto não implica necessàriamente que T seja positivo. Precisamos mostrar também que T é auto-adjunto. É aqui que usamos a hipótese de que (8-16) vale para todos os complexos $x_1, \ldots, x_n$. Esta hipótese mais forte implica que A é auto-adjunta e, portanto, que T é auto-adjunto.

Novamente para o caso real, suponhamos que T seja positivo. Então nosso argumento que levou a (8-15) mostra que A é auto-adjunta e que (8-16) vale para números *reais* arbitrários $x_1, \ldots, x_n$, não todos nulos. Mas precisamos mostrar que (8-16) também vale para $x_1, \ldots, x_n$ complexos. Pelo Teorema 10, existe um operador linear inversível U sôbre V, tal que $T = U^*U$. Se P é a matriz de U em relação à base $\mathcal{B}$, temos

$$A = P^*P = P^tP.$$

Se X é uma $n \times 1$ matriz complexa, então

$$\sum_j \sum_k A_{kj}x_j\bar{x}_k = X^*AX = X^*P^*PX = (PX)^*(PX).$$

Seja $Y = PX$. Se $X \neq 0$, então como P é inversível, $Y \neq 0$ e então $Y^*Y > 0$. Logo, A é positiva.

Corolário. *Se* A *é uma* n $\times$ n *matriz com elementos complexos, então* A *é positiva se, e sòmente se, existe uma* n $\times$ n *matriz inversível* P *tal que* A = P*P. *Se* A *é positiva e possui elementos reais então* A = P^tP *onde* P *é uma* n $\times$ n *matriz inversível com elementos reais.*

O restante desta seção será dedicado à obtenção de um teste para a positividade de uma dada matriz. Pràticamente por definição, a $n \times n$ matriz A é positiva se, e sòmente se,

$$(X, Y) = Y^*AX$$

define um produto interno sôbre o espaço das $n \times 1$ matrizes sôbre C, isto é,

$$(\sum_j x_j\epsilon_j, \sum_k y_k\epsilon_k) = \Sigma_j \Sigma_k A_{jk}x_k\bar{y}_j \tag{8-17}$$

define um produto interno sôbre C^n. O teste para a positividade que vamos obter é baseado em duas observações.

(i) Se A é uma matriz positiva, então det $A > 0$.

(ii) Se A é uma matriz positiva e $1 \leq k \leq n$, a matriz

$$A^{(k)} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ A_{21} & A_{22} & \cdots & A_{2k} \\ \vdots & \vdots & & \vdots \\ A_{k1} & A_{k2} & \cdots & A_{kk} \end{bmatrix} \tag{8-18}$$

é uma $k \times k$ matriz positiva.

A afirmação (i) é demonstrada como segue: Se A é positiva então $A = P^*P$ para alguma $n \times n$ matriz inversível P. Assim,

$$\det A = \det(P^*P) = \det(P^*)\det(P) = \overline{(\det P)}\det P > 0.$$

A afirmação (ii) decorre desta observação: Suponhamos que V seja um espaço vetorial de dimensão finita e que (,) seja um produto interno sôbre V. Se $\{\alpha_1, \ldots, \alpha_n\}$ é uma base *arbitrária* de V, a matriz do produto interno em relação a esta base é uma matriz positiva. Observamos êste fato na primeira seção dêste capítulo. Lembramos ao leitor que a matriz G do produto interno $(\, , \,)$ em relação à base $\{\alpha_1, \ldots, \alpha_n\}$ é definida por

$$G_{jk} = (\alpha_k, \alpha_j).$$

Esta matriz é positiva simplesmente porque

$$(\sum_j x_j\alpha_j, \sum_k y_k\alpha_k) = \sum_j \sum_k G_{kj}x_j\bar{y}_k$$

e então $[X, Y] = Y^*GX$ é um produto interno sôbre o espaço das $n \times 1$ matrizes. Suponhamos agora que nos seja dada uma $n \times n$ matriz positiva A. Se $1 \leq k \leq n$, seja W_k o subespaço de C^n gerado por $\epsilon_1, \epsilon_2, \ldots, \epsilon_k$. Ora, (8-17) define um produto interno $(\, , \,)$ sôbre C^n. Se restringirmos êste produto interno ao subespaço W_k, obteremos um produto interno sôbre W_k, cuja matriz em relação à base $\{\epsilon_1, \ldots, \epsilon_k\}$ será $A^{(k)}$

$$A_{jk} = (\epsilon_k, \epsilon_j).$$

O que as observações (i) e (ii) nos dizem é que, se A é uma matriz positiva, então $\det A^{(k)} > 0$ para $k = 1, \ldots, n$:

$$A_{11} > 0, \; A_{11}A_{22} - A_{12}A_{21} > 0, \quad \ldots, \det A > 0.$$

Para matrizes auto-adjuntas, vale a recíproca.

Definição. *Seja* A *uma* n $\times$ n *matriz. Os* **menores principais** *de* A *são os* n *escalares definidos por*

$$\det A^{(k)} = \det \begin{bmatrix} A_{11} & \ldots & A_{1k} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ A_{k1} & \ldots & A_{kk} \end{bmatrix}, \quad k = 1, \ldots, n.$$

Teorema 12. *Seja* A *uma* n $\times$ n *matriz sôbre o corpo dos números complexos e suponhamos que* A *seja auto-adjunta (hermitiana). Então* A *é positiva se, e sòmente se, os menores principais de* A *são todos positivos.*

Demonstração. Acabamos de observar que se A é uma matriz positiva, os menores principais de A são números positivos.

Suponhamos agora que A seja auto-adjunta e que $\det A^{(k)} > 0$ para $k = 1, \ldots, n$. Demonstraremos que A é positiva. Seja $(\, , \,)$ a função definida sôbre os pares ordenados de vetores em C^n por (8-17).

Certamente $(c\alpha + \beta, \gamma) = c(\alpha\, \gamma) + (\beta, \gamma)$; e, como estamos supondo que A é auto-adjunta, temos $(\beta, \alpha) = \overline{(\alpha, \beta)}$.

Demonstraremos que A é positiva, por indução sôbre n. Notemos primeiro que uma 1×1 matriz auto-adjunta com menores principais positivos é positiva. Suponhamos que o resultado seja verdadeiro para $(n-1) \times (n-1)$ matrizes e seja A uma $n \times n$ matriz auto-adjunta com menores principais positivos. Seja

$$\epsilon_j' = \epsilon_j - \frac{A_{1j}}{A_{11}}\epsilon_1, \quad 2 \leq j \leq n.$$

A divisão por A_{11} não causa complicações porque $A_{11} > 0$. Então $\mathcal{B}' = \{\epsilon_1, \epsilon_2, \ldots, \epsilon_n\}$ é uma base ordenada de C^n. Usando o fato de que $A_{jk} = (\epsilon_k, \epsilon_j)$ e que $(\beta, \alpha) = \overline{(\alpha, \beta)}$, obtém-se $(\epsilon_j', \epsilon_1) = (\epsilon_1, \epsilon_j') = 0$ para $j \geq 2$. Se indicarmos por A' a matriz de (,) em relação à base $\mathcal{B}'$, veremos então que

$$A' = \begin{bmatrix} A_{11} & 0 & 0 & \ldots & 0 \\ 0 & A_{22} & A_{23}' & \ldots & A_{2n}' \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 0 & A_{n2}' & A_{n3} & \ldots & A_{nn}' \end{bmatrix}.$$

De fato, se $j \geq 2$ e $k \geq 2$, temos por definição

$$\begin{aligned} A_{jk}' &= (\epsilon_k, \epsilon_j) \\ &= \left(\epsilon_k - \frac{A_{1k}}{A_{11}}\epsilon_1, \epsilon_j - \frac{A_{1j}}{A_{11}}\epsilon_1\right) \\ &= (\epsilon_k, \epsilon_j) - \frac{A_{1k}}{A_{11}}(\epsilon_1, \epsilon_j) \\ &= A_{jk} - \frac{A_{1k}}{A_{11}}A_{j1}. \end{aligned}$$

Esta relação também é válida para $j = 1$ e $k \geq 2$. Assim, temos

$$A_k' = A_k - \frac{A_{1k}}{A_{11}}A_1, \quad 2 \leq k \leq n$$

onde A_k' é a k-ésima coluna de A', A_k é a k-ésima coluna de A e A_1 é a primeira coluna de A. Em outras palavras, subtraindo das outras colunas múltiplos convenientes da primeira coluna de A, obtemos a matriz

$$\begin{bmatrix} A_{11} & 0 & \dots & 0 \\ A_{21} & A'_{22} & \dots & A'_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ A_{n1} & A'_{n2} & \dots & A'_{nn} \end{bmatrix}.$$

Podemos obter A' a partir desta matriz subtraindo das outras linhas múltiplos convenientes da primeira linha. O fato crucial agora é observar que os menores principais não são alterados por êstas operações elementares sôbre colunas e linhas. Assim, A e A' possuem os mesmos menores principais. Como êsses menores são positivos

$$A_{11} \det \begin{bmatrix} A'_{22} & \dots & A'_{2k} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ A'_{k2} & \dots & A'_{kk} \end{bmatrix} > 0, \quad k = 2, \dots, n.$$

Dividindo por A_{11}, concluímos que os menores principais da $(n-1) \times (n-1)$ matriz

$$B = \begin{bmatrix} A'_{22} & \dots & A'_{2n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ A'_{n2} & \dots & A'_{nn} \end{bmatrix}$$

são todos positivos. Ora, B também é auto-adjunta, porque para $j \geq 2$ e $k \geq 2$

$$A'_{kj} = (\epsilon'_j, \epsilon'_k) = \overline{(\epsilon'_k, \epsilon'_j)} = \overline{A'_{jk}}.$$

Pela hipótese de indução, B é uma matriz positiva. Como

$$\sum_j \sum_k A'_{kj} x_j \bar{x}_k = A_{11}|x_1|^2 + \sum_{k=2}^{n} \sum_{j=2}^{n} A'_{kj} x_j \bar{x}_k$$

então é evidente que A é uma matriz positiva. Mas A' é a matriz de $(\ ,\ )$ em relação à base ordenada $\{\epsilon_1, \epsilon_2, \dots, \epsilon'_n\}$ e então $(\ ,\ )$ deve ser um produto interno. Assim, A é positiva.

Resumindo, se A é uma $n \times n$ matriz sôbre o corpo dos números complexos, as seguintes afirmações são equivalentes:

(i) A é positiva, isto é, $\sum_j \sum_k A_{kj} x_j \bar{x}_k > 0$ sempre que $x_1, \ldots, x_n$ são números complexos, não todos nulos.

(ii) $(X, Y) = Y^*AX$ é um produto interno sôbre o espaço das $n \times 1$ matrizes complexas.

(iii) Em relação ao produto interno canônico $(X, Y) = Y^*X$ sôbre $n \times 1$ matrizes, o operador linear $X \rightarrow AX$ é positivo.

(iv) $A = P^*P$ para alguma $n \times n$ matriz inversível P sôbre C.

(v) $A = A^*$, e os menores principais de A são positivos.

Se todo elemento de A é real, estas são equivalentes a:

(vi) $A = A^t$ e $\sum_j \sum_k A_{kj} x_j x_k > 0$ sempre que $x_1, \ldots, x_n$ são números reais não todos nulos.

(vii) $(X, Y) = Y^tAX$ é um produto interno sôbre o espaço das $n \times 1$ matrizes reais.

(viii) Em relação ao produto interno canônico $(X, Y) = Y^tX$ sôbre $n \times 1$ matrizes reais, o operador linear $X \rightarrow AX$ é positivo.

(ix) Existe uma $n \times n$ matriz inversível P, com elementos reais, tal que $A = P^tP$.

Exercícios

1. Seja V igual a C^2, com o produto interno canônico. Para que vetores α em V existe um operador linear positivo T tal que $\alpha = T\epsilon_1$?

2. Suponhamos que V seja R^2, com o produto interno canônico. Se θ é um número real, seja T_θ o operador linear "rotação de ângulo θ"

$$T_\theta(x_1, x_2) = (x_1 \cos \theta - x_2 \operatorname{sen} \theta, x_1 \operatorname{sen} \theta + x_2 \cos \theta).$$

Para que valores de θ se tem T_θ um operador positivo?

3. Seja V o espaço das $n \times 1$ matrizes sôbre C, com o produto interno $(X, Y) = Y^*GX$ (onde G é uma $n \times n$ matriz tal que isto seja um produto interno). Seja A uma $n \times n$ matriz e T o operador linear $T(X) = AX$. Determinar T^*. Se Y é um elemento fixo de V, encontrar o elemento Z de V que determina o funcional linear $X \rightarrow Y^*X$. Em outras palavras, encontrar Z tal que $Y^*Z = (X, Z)$ para tôda X em V.

4. Seja V um espaço de dimensão finita com produto interno. Se T e V são operadores lineares positivos sôbre V, demonstrar que $(T + U)$ é positivo. Dar um exemplo que mostre que TU não é necessàriamente positivo.

5. Seja

$$A = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{bmatrix}.$$

(a) Mostrar que A é positiva.

(b) Seja V o espaço das 2×1 matrizes reais, com o produto interno $(X, Y) = Y^t AX$. Determinar uma base ortonormal de V, aplicando o processo de Gram-Schmidt à base $\{X_1, X_2\}$ definida por

$$X_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad X_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

(c) Determinar uma 2×2 matriz real P tal que $A = P^t P$.

6. Quais das matrizes são positivas?

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \begin{bmatrix} 1 & 1+i \\ 1-i & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & -1 & 1 \\ 2 & -1 & 1 \\ 3 & -1 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}$$

7. Dar um exemplo de uma $n \times n$ matrizes cujos menores principais sejam todos positivos, mas que não seja matriz positiva.

8. Verificar se $((x_1, x_2), (y_1, y_2)) = x_1\bar{y}_1 + 2x_2\bar{y}_1 + 2x_1\bar{y}_2 + x_2\bar{y}_2$ define um produto interno sôbre C^2.

9. Demonstrar que todo elemento da diagonal principal de uma matriz positiva é positivo.

10. Seja V um espaço de dimensão finita com produto interno. Se T e U são operadores lineares sôbre V, coloquemos $T < U$ se $U - T$ é um operador positivo. Demonstrar o seguinte:

(a) É impossível que $T < U$ e $U < T$.
(b) Se $T < U$ e $U < S$, então $T < S$.
(c) Se $T < U$ e $0 < S$, não é necessário que $ST < SU$.

11. Seja V um espaço de dimensão finita com produto interno e E a projeção ortogonal de V sôbre algum subespaço.

(a) Demonstrar que, para todo número positivo arbitrário c, o operador $cI + E$ é positivo.

(b) Exprimir em têrmos de E um operador linear auto-adjunto T tal que $T^2 = I + E$.

12. Seja n um inteiro positivo e A a $n \times n$ matriz

$$A = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \frac{1}{2n-1} \end{bmatrix}$$

Demonstrar que A é positiva.

13. Seja A uma $n \times n$ matriz auto-adjunta. Demonstrar que existe um número real c tal que a matriz $cI + A$ seja positiva.

14. Demonstrar que o produto de dois operadores lineares é positivo se, e sòmente se, êles comutam.

8.5 Operadores Unitários

Nesta seção, vamos considerar o conceito de um isomorfismo entre dois espaços com produto interno. Se V e W são espaços vetoriais, um isomorfismo de V em W é uma transformação linear bijetora de V em W, isto é, uma correspondência bijetora entre os elementos de V e os de W, a qual "conserva" as operações de espaço vetorial. Ora, um espaço com produto interno consiste de um espaço vetorial e um produto interno especificado sôbre aquêle espaço. Assim, quando V e W são espaço com produto interno, exigiremos que um isomorfismo de V em W não só conserve as operações lineares, mas também conserve produtos internos. Um isomorfismo de um espaço com produto interno em si mesmo é denominado um "operador unitário" sôbre aquêle espaço. Consideraremos vários exemplos de operadores unitários e estabeleceremos suas propriedades fundamentais.

Definição. *Sejam* V *e* W *espaços com produto interno sôbre o mesmo corpo e seja* T *uma transformação linear de* V *em* W. *Dizemos que* T **conserva produtos internos** *se* $(T\alpha, T\beta) = (\alpha, \beta)$ *para todos* α, β *em* V. *Um isomorfismo de* V *em* W *é um* **isomorfismo** T *de espaço vetorial de* V *em* W *que também conserva produtos internos.*

Se T conserva produtos internos, então $||T\alpha|| = ||\alpha||$, portanto T é, necessàriamente, não-singular. Assim, um isomorfismo de V em W pode também ser definido como uma transformação linear de V em W que conserva produtos internos. Se T é um isomorfismo de V em W, então TP^1 é um isomorfismo de W em V; logo, quando um tal T existir, diremos simplesmente que V e W são **isomorfos**. É claro que o isomorfismo de espaços com produto interno é uma relação de equivalência.

Teorema 13. *Sejam* V *e* W *espaços de dimensão finita com pro duto interno sobre o mesmo corpo, que tenham a mesma dimensão Se* T *é uma transformação linear de* V *em* W, *as seguintes afirmações são equivalentes:*

(i) T *conserva produtos internos.*

(ii) T *é um isomorfismo (de espaço com produto interno).*

(iii) T *leva tôda base ortonormal de* V *em uma base ortonormal* de W.

(iv) T *leva alguma base ortonormal de* V *em alguma base ortonormal de* W.

Demonstração. (i) $\rightarrow$ (ii) Se T conserva produtos internos, então $\|T\alpha\| = \|\alpha\|$ para todo α em V. Assim, T é não-singular e como dim V = dim W, sabemos que T é um isomorfismo de espaço vetorial.

(ii) $\rightarrow$ (iii) Suponhamos que T seja um isomorfismo. Seja $\{\alpha_1, \ldots, \ldots, \alpha_n\}$ uma base ortonormal de V. Como T é um isomorfismo de espaço vetorial e dim W = dim V, decorre que $\{T\alpha_1, \ldots, T\alpha_n\}$ é uma base de W. Como T conserva também produtos internos, $(T\alpha_j, T\alpha_k) = (\alpha_j, \alpha_k) = \delta_{jk}$.

(iii) $\rightarrow$ (iv) Não requer comentários.

(iv) $\rightarrow$ (i) Seja $\{\alpha_1, \ldots, \alpha_n\}$ uma base ortonormal de V tal que $\{\alpha_1, \ldots, T\alpha_n\}$ seja uma base ortonormal de W. Então

$$(T\alpha_j, T\alpha_k) = (\alpha_j, \alpha_k) = \delta_{jk}.$$

Para todos $\alpha = x_1\alpha_1 + \ldots + x_n\alpha_n$ e $\beta = y_1\alpha_1 + \ldots + y_n\alpha_n$ em V, temos

$$\begin{aligned}(\alpha, \beta) &= \sum_{j=1}^{n} x_j\bar{y}_j \\ (T\alpha, T\beta) &= (\sum_j x_jT\alpha_j, \sum_k y_kT\alpha_k) \\ &= \sum_j \sum_k x_{jk}\bar{y}(T\alpha_j, T\alpha_k) \\ &= \sum_{j=1}^{n} x_j\bar{y}_j\end{aligned}$$

logo T conserva produtos internos.

Corolário. *Sejam* V *e* W *espaços de dimensão finita como produto interno sôbre o mesmo corpo. Então* V *e* W *são isomorfos se, e sòmente se, tem a mesma dimensão.*

Demonstração. Se $\{\alpha_1, \ldots, \alpha_n\}$ é uma base ortonormal de V e $\{\beta_1, \ldots, \beta_n\}$ é uma base ortonormal de W, seja T a transformação linear de V em W definida por $T\alpha_j = \beta_j$. Então T é um isomorfismo de V em W.

Exemplo 17. Se V é um espaço n-dimensional com produto interno, então tôda base ortonormal ordenada $\mathcal{B} = \{\alpha, \ldots, \alpha_n\}$ determina um isomorfismo de V em F^n com o produto interno canônico. O isomorfismo é simplesmente

$$T(x_1\alpha_1 + \ldots + x_n\alpha_n) = (x_1, \ldots, x_n).$$

Existe o isomorfismo, superficialmente diferente, determinado por $\mathcal{B}$, de V no espaço das $n \times 1$ matrizes com $(X, Y) = Y^*X$ como produto interno. O isomorfismo é

$$\alpha \rightarrow [\alpha]_{\mathcal{B}}$$

isto é, a transformação que leva α na matriz de suas coordenadas em relação à base ordenada $\mathcal{B}$. Para qualquer base ordenada $\mathcal{B}$, êste é um isomorfismo de espaço vetorial; no entanto, êle é um isomorfismo dos dois espaços com produto interno se, e sòmente se, $\mathcal{B}$ é ortonormal.

Exemplo 18. Eis um isomorfismo um pouco menos superficial. Seja W o espaço das 3×3 matrizes A sôbre R que sejam anti-simétricas, isto é, $A^t = -A$. Vamos equipar W com o produto interno $(A, B) = \frac{1}{2} \operatorname{tr}(AB^t)$, sendo o $\frac{1}{2}$ colocado por conveniência. Seja V o espaço R^3 com o produto interno canônico. Seja T a transformação linear de V em W definida por

$$T(x_1, x_2, x_3) = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}.$$

Então T leva V sôbre W, e colocando

$$A = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & -y_3 & y_2 \\ y_3 & 0 & -y_1 \\ -y_2 & y_1 & 0 \end{bmatrix}$$

temos

$$\begin{aligned} \operatorname{tr}(AB^t) &= x_3y_3 + x_2y_2 + x_3y_3 + x_2y_2 + x_1y_1 \\ &= 2(x_1y_1 + x_2y_2 + x_3y_3). \end{aligned}$$

Assim, $(\alpha, \beta) = (T\alpha, T\beta)$ e T é um isoformismo de espaço vetorial. Notemos que T leva a base canônica $\{\epsilon_1, \epsilon_2, \epsilon_3\}$ na base ortonormal formada pelas três matrizes

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Exemplo 19. Nem sempre é particularmente conveniente descrever um isomorfismo em têrmos de bases ortonormais. Por exemplo, seja G uma $n \times n$ matriz positiva, isto é, uma matriz auto-adjunta (hermitiana) com os menores principais positivos. Seja V o espaço das $n \times 1$ matrizes complexas, com o produto interno $[X, Y] =$

$= Y^*GX$. Seja W o mesmo espaço vetorial, com o produto interno canônico $(X, Y) = Y^*X$. Sabemos que V e W são espaços com produto interno que são isomorfos. Deveria parecer que a maneira mais conveniente de descrever um isomorfismo entre V e W seja a seguinte: Como G é positiva, existe uma $n \times n$ matriz inversível P tal que $G = P^*P$. Seja T a transformação linear de V em W definida por $T(X) = PX$. Então

$$\begin{aligned}(TX, TY) &= (PX, PY)\\ &= (PY)^*(PX)\\ &\ Y^*P^*PX\\ &= Y^*GX\\ &= [X, Y].\end{aligned}$$

Logo, T é um isomorfismo.

Exemplo 20. Seja V o espaço das funções contínuas, definidas sôbre o intervalo unitário, $0 \leq t \leq 1$ e tomando valores reais, com o produto interno

$$(f, g) = \int_0^1 f(t)g(t)t^2dt.$$

Seja W o mesmo espaço vetorial com o produto interno

$$(f, g) = \int_0^1 f(t)g(t)dt.$$

Seja T a transformação linear de V em W dada por

$$(Tf)(t) = tf(t).$$

Então $(Tf, Tg) = [f, g]$, portanto T conserva produtos internos; contudo, T *não* é um isomorfismo de V em W, porque a imagem de T não é todo o espaço W. Evidentemente, isto ocorre porque o espaço vetorial subjacente não é de dimensão finita.

Teorema 14. *Sejam* V *e* W *espaços com produto interno sôbre o mesmo corpo e seja* T *uma transformação linear de* V *em* W. *Então,* T *conserva produtos internos se, e sòmente se,* $||T\alpha|| = ||\alpha||$ *para todo* α *em* V.

Demonstração. Se T conserva produtos internos, T "conserva normas". Suponhamos que $||T\alpha|| = ||\alpha||$ para todo α em V. Então $||T\alpha||^2 = ||\alpha||^2$. Usando agora a identidade de polarização conveniente, (8-3) ou (8-4) e o fato de que T é linear, obtém-se fàcilmente $(\alpha, \beta) = (T\alpha, T\beta)$ para quaisquer α, β em V.

Definição. *Um* **operador unitário** *sôbre um espaço com produto interno é um isomorfismo do espaço em si mesmo.*

O produto de dois operadores unitários é unitário. De fato, se U_1 e U_2 são unitários, então U_2U_1 é inversível e $\|U_2U_1\alpha\| = \|U_1\alpha\| = \|\alpha\|$ para todo α. Além disso, o inverso de um operador unitário é unitário, pois $\|U\alpha\| = \|\alpha\|$ diz que $\|U^{-1}\beta\| = \|\beta\|$, onde $\beta = U\alpha$. Como o operador idêntico é òbviamente unitário, vemos que o conjunto dos operadores unitários sôbre um espaço com produto interno é um grupo, com a operação de composição.

Se V é um espaço de dimensão finita com produto interno e U é um operador linear sôbre V, o Teorema 13 nos diz que U é unitário se, e sòmente se, $(U\alpha, U\beta) = (\alpha, \beta)$ para todos α, β em V; ou seja, se, e sòmente se, para alguma (tôda) base ortonormal $\{\alpha_1, \ldots, \alpha_n\}$ é verdade que $\{U\alpha_1, \ldots, U\alpha_n\}$ é uma base ortonormal.

Teorema 15. *Seja* U *um operador linear sôbre um espaço* V *com produto interno. Então* U *é unitário se, e sòmente se, o adjunto* U* *de* U *existe e* UU* = U*U = I.

Demonstração. Suponhamos que U seja unitário. Então U é inversível e

$$(U\alpha, \beta) = (U\alpha, UU^{-1}\beta) = (\alpha, U^{-1}\beta)$$

para todos α, β. Logo, U^{-1} é o adjunto de U.

Recìprocamente, suponhamos que U^* exista e $UU^* = U^*U = I$. Então U é inversível, com $U^{-1} = U^*$. Portanto, basta mostrar que U conserva produtos internos. Temos

$$\begin{aligned}(U\alpha, U\beta) &= (\alpha, U^*U\beta)\\ &= (\alpha, I\beta)\\ &= (\alpha, \beta)\end{aligned}$$

para todos α, β.

Exemplo 21. Seja V o espaço das $n \times 1$ matrizes sôbre C, com o produto interno $(X, Y) = Y^*X$. Seja A uma $n \times n$ matriz sôbre C e seja U o operador linear definido por $U(X) = AX$. Então

$$(UX, UY) = (AX, AY) = Y^*A^*AX$$

para tôdas X, Y. Logo, U é unitário se, e sòmente se, $A^*A = I$.

Definição. *Uma* $n \times n$ *matriz complexa* A *é dita* **unitária** *se* $A^*A = I$.

Teorema 16. *Seja* V *um espaço de dimensão finita com produto interno e seja* U *um operador linear sôbre* V. *Então* U *é unitário se, e sòmente se, a matriz de* U *em relação a alguma (tôda) base ortonormal ordenada é uma matriz unitária.*

Demonstração. A esta altura, isto não é bem um teorema e só o enunciamos por questão de ênfase. Se $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ é uma base ortonormal ordenada de V e A é a matriz de U em relação a $\mathcal{B}$, então $A^*A = I$ se, e sòmente se, $U^*U = I$. O resultado decorre do Teorema 15.

Seja A uma $n \times n$ matriz. A afirmação de que A é unitária diz simplesmente

$$(A^*A)_{jk} = \delta_{jk}$$

ou

$$\sum_{r=1}^{n} A_{rj}A_{rk} = \delta_{jk}.$$

Em outras palavras, diz que as colunas de A formam um conjunto ortonormal de matrizes-colunas, usando o produto interno canônico $(X, Y) = Y^*X$. Como $A^*A = I$ se, e sòmente se, $AA^* = I$ vemos que A é unitária exatamente quando as linhas de A constituem um conjunto ortonormal de n-uplas em C^n (com o produto interno canônico). Portanto, usando produtos internos canônicos, A é unitária se, e sòmente se, as linhas (colunas) de A são ortonormais. Vê-se aqui um exemplo da fôrça do teorema que afirma que uma inversa unilateral de uma matriz é, na verdade, uma inversa bilateral. Aplicando êste teorema como fizemos acima, digamos, a matrizes reais, temos o seguinte: Suponhamos que exista uma tabela quadrada de números reais tal que a soma dos quadrados dos elementos de cada linha seja 1 e tal que linhas distintas sejam ortogonais. Então a soma dos quadrados dos elementos de cada coluna é 1 e colunas distintas são ortogonais. Basta escrever a demonstração dêste fato para uma tabela 3×3, sem usar nenhum conhecimento sôbre matrizes, para se ficar razoàvelmente impressionado.

Definição. *Uma* n $\times$ n *matriz* A, *real ou complexa, é dita* **ortogonal** *se* $A^tA = I$.

Uma matriz ortogonal real é unitária; uma matriz unitária é ortogonal se, e sòmente se, cada um dos seus elementos é real.

Exemplo 22. Vejamos alguns exemplos de matrizes unitárias e ortogonais.

(a) Uma 1×1 matriz $[c]$ é ortogonal se, e sòmente se, $c = \pm 1$ e é unitária se, e sòmente se, $\bar{c}c = 1$. A última condição significa (é claro) que $|c| = 1$, ou seja, $c = e^{i\theta}$, com θ real.

(b) Seja

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Então A é ortogonal se, e sòmente se,

$$A^t = A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

O determinante de qualquer matriz ortogonal, como se pode ver fàcilmente, é ± 1. Assim A é ortogonal se, e sòmente se,

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}.$$

ou

$$A = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$$

onde $a^2 + b^2 = 1$. Os dois casos são distinguidos pelo valor de det A.

(c) As bem conhecidas relações entre as funções trigonométricas mostram que a matriz

$$\mathrm{A}_\theta = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

é ortogonal. Se θ é um número real, então A_θ é a matriz, em relação à base ordenada canônica de R^2, do operador linear U_θ, que é a rotação do ângulo θ. A afirmação de que A_θ é uma matriz ortogonal real (logo unitária) significa simplesmente que U_θ é um operador unitário, isto é, conserva produtos escalares.

(d) Seja

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Então A é unitária se, e sòmente se,

$$\begin{bmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{bmatrix} = \frac{1}{ad - bc}\begin{bmatrix} d & b \\ -c & a \end{bmatrix}.$$

O determinante de uma matriz unitária tem valor absoluto 1 e é portanto um número complexo da forma $e^{i\theta}$, θ real. Assim, A é unitária se, e sòmente se,

$$A = \begin{bmatrix} a & b \\ -e^{i\theta}\bar{b} & e^{i\theta}\bar{a} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{bmatrix} \begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}$$

onde θ é um número real e a, b são números complexos tais que $|a|^2 + |b|^2 = 1$.

Consideremos agora abreviadamente uma mudança de coordenadas em um espaço com produto interno. Suponhamos que V seja um espaço de dimensão finita com produto interno e que $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ e $\mathcal{B}' = \{\alpha'_1, \ldots, \alpha'_n\}$ sejam duas bases *ortonormais* ordenadas de V. Existe uma única $n \times n$ matriz P (necessàriamente inversível) tal que

$$[\alpha]_{\mathcal{B}'} = P^{-1}[\alpha]_{\mathcal{B}}$$

para todo α em V. Se U é o bem determinado operador linear sôbre V definido por $U\alpha_j = \alpha'_j$, então P é a matriz de U em relação à base ordenada Σ:

$$\alpha'_k = \sum_{j=1}^{n} P_{jk}\alpha_j.$$

como $\mathcal{B}$ e $\mathcal{B}'$ são bases ortonormais, U é um operador unitário e P é uma matriz unitária. Se T é um operador linear arbitrário sôbre V, então

$$[T]_{\mathcal{B}'} = P^{-1}[T]_{\mathcal{B}}P = P^*[T]_{\mathcal{B}}P.$$

Definição. *Sejam* A *e* B n $\times$ n *matrizes complexas. Dizemos que* B *é* **unitàriamente equivalente** *a* A *se existe uma* n $\times$ n *matriz unitária* P *tal que* B = P*AP. *Dizemos que* B *é* **ortogonalmente equivalente** *a* A *se existe uma* n $\times$ n *matriz ortogonal* P *tal que* B = P^tAP.

Com essa definição, o que observamos acima pode ser enunciado como segue: Se $\mathcal{B}$ e $\mathcal{B}'$ são duas bases ortonormais ordenadas de V, para todo operador linear T sôbre V, a matriz $[T]_{\mathcal{B}'}$ é unitàriamente equivalente à matriz $[T]_{\mathcal{B}}$. No caso de V ser um espaço real com produto interno, estas matrizes são ortogonalmente equivalentes, através de uma matriz ortogonal real.

Exercícios

1. Determinar uma matriz unitária que não seja ortogonal e determinar uma matriz ortogonal que não seja unitária.

2. Seja V o espaço das $n \times n$ matrizes complexas com o produto interno $(A, B) = \operatorname{tr}(AB^*)$. Para cada M em V, seja T_M o operador linear definido por $T_M(A) = MA$. Mostrar que T_M é unitário se, e sòmente se, M é uma matriz unitária.

3. Seja V o conjunto dos números complexos, considerado como um espaço vetorial *real*.

(a) Mostrar que $(\alpha, \beta) = Re\,(\alpha\bar{\beta})$ define um produto interno sôbre V.

(b) Exibir um isomorfismo (de espaço com produto interno) de V em R^2 com o produto interno canônico.

(c) Para cada γ em V, seja M_γ o operador linear sôbre V definido por $M_\gamma(\alpha) = \gamma\alpha$. Mostrar que $(M_\gamma)^* = M_{\bar{\gamma}}$.

(d) Para quais números complexos γ se tem M_γ auto-adjunto?

(e) Para quais γ, M_γ é unitário?

(f) Para quais γ, M_γ é positivo?

(g) Qual é o $\det(M_\gamma)$?

(h) Determinar a matriz de M_γ em relação à base $\{1, i.\}$

(i) Se T é um operador linear sôbre V, encontrar condições necessárias e suficientes para que T seja um M_γ.

(j) Encontrar um operador unitário sôbre V que não seja um M_γ.

4. Seja V o espaço R^2, com o produto interno canônico. Se U é um operador unitário sôbre V, mostrar que a matriz de U em relação à base ordenada canônica é

$$\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \quad \text{ou} \quad \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{bmatrix}$$

para algum θ, $0 \leq \theta < 2$. Seja U_θ o operador linear correspondente à primeira matriz, isto é, U_θ é uma rotação de um ângulo θ. Agora é possível convencer-se de que todo operador unitário sôbre V é uma rotação ou uma reflexão em relação ao eixo β_1 seguida de uma rotação.

(a) O que é $U_\theta U_\phi$?

(b) Mostrar que $U^*_\theta = U_{-\theta}$.

(c) Seja ϕ um número real fixo e $\mathcal{B} = \{\alpha_1, \alpha_2\}$ a base ortonormal obtida girando $\{\epsilon_1, \epsilon_2\}$ de um ângulo ϕ, isto é, $\alpha_j = U_\phi \epsilon_j$. Se θ é um outro número real, qual é a matriz de U_θ em relação à base ordenada $\mathcal{B}$?

5. Seja V o espaço R^3, com o produto interno canônico. Seja W o plano gerado por $\alpha = (1, 1, 1)$ e $\beta = (1, 1, -2)$. Seja U o operador linear definido geomètricamente como segue: U é uma rotação de um ângulo θ, em tôrno de uma reta que passa pela origem e é ortogonal a W. Existem na verdade duas tais rotações: tome-se uma delas. Determinar a matriz de U em relação à base ordenada canônica. (Eis um modo possível de se proceder: Determinar α_1 e α_2 que formem uma base ortonormal de W. Seja α_3 um vetor de norma 1, ortogonal a W. Determinar a matriz de U em relação à base $\{\alpha_1, \alpha_2, \alpha_3\}$. Efetuar uma mudança de base.)

6. Seja V um espaço de dimensão finita com produto interno e seja W um subespaço de V. Então $V = W \oplus W^\perp$, isto é, cada α em V pode ser expresso de um único modo sob a forma $\alpha = \beta + \gamma$, com β em W e γ em $W^\perp$. Definamos um operador linear U por $U\alpha = \beta - \gamma$.

(a) Demonstrar que U é auto-adjunto e unitário.

(b) Se V é R^3 com o produto interno canônico e W é o subespaço gerado por $(1, 0, 1)$, encontrar a matriz de U em relação à base ordenada canônica.

7. Seja V um espaço *complexo* com produto interno e T um operador linear *auto-adjunto* sôbre V. Mostrar que

(a) $||\alpha + iT\alpha|| = ||\alpha - iT\alpha||$ para todo α em V.
(b) $\alpha + iT\alpha = \beta + iT\beta$ se, e sòmente se, $\alpha = \beta$.
(c) $I + iT$ é não-singular
(d) $I - iT$ é não-singular.
(e) Suponhamos que V seja de dimensão finita; demonstrar que $U = (I - iT)(I + iT)^{-1}$ é um operador unitário; U é denominado a **transformada de Cayley** de T. Num certo sentido, $U = f(T)$, onde $f(x) = (\sigma - ix)/(1 + ix)$.

8. Se θ é um número real, demonstrar que as matrizes seguintes são unitàriamente equivalentes.

$$\begin{bmatrix} \cos\theta & -\text{sen}\,\theta \\ \text{sen}\,\theta & \cos\theta \end{bmatrix} \quad \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix}.$$

9. Seja V um espaço de dimensão finita com produto interno e T um operador linear positivo sôbre V. Seja p_T o produto interno sôbre V definido por $p_T(\alpha, \beta) = (T\alpha, \beta)$. Seja U um operador linear sôbre V e U^* o seu adjunto em relação a $(\ ,\ )$. Demonstrar que U é unitário em relação ao produto interno p_T se, e sòmente se, $T = U^*TU$.

10. Seja V um espaço de dimensão finita com produto interno. Para cada par α, β em V, seja $T_{\alpha,\beta}$ o operador linear sôbre V definido por $T_{\alpha,\beta}(\gamma) = (\gamma, \beta)\alpha$. Mostrar que

(a) $T^*_{\alpha,\beta} = T_{\beta,\alpha}$;
(b) traço $(T_{\alpha,\beta}) = (\alpha, \beta)$;
(c) $T_{\alpha,\beta}\, T_{\gamma,\delta} = T_{\alpha,(\beta,\gamma)\delta}$.
(d) Em que condições $T_{\alpha,\beta}$ é auto-adjunto?

11. Seja V um espaço n-dimensional com produto interno sôbre o corpo F e seja $L(V, V)$ o espaço dos operadores lineares sôbre V. Mostrar que existe um único produto interno sôbre $L(V, V)$ com a propriedade que $||T_{\alpha,\beta}||^2 = ||\alpha||^2||\beta||^2$ para todos α,β em V. ($T_{\alpha,\beta}$ é o operador definido no Exercício 10.) Encontrar um isomorfismo entre $L(V, V)$ com êste produto interno e o espaço das $n \times n$ matrizes sôbre F, com o produto interno $(A, B) = \text{tr}\,(AB^*)$.

12. Seja V um espaço de dimensão finita com produto interno. No Exercício 6, mostramos como construir certos operadores lineares sôbre V que são auto-adjuntos e unitários. Demonstrar então que não existem outros, isto é, que todo operador auto-adjunto unitário provém de algum subespaço W como descrevemos no Exercício 6.

13. Sejam V e W espaços de mesma dimensão finita com produto interno. Seja U um isomorfismo de V em W. Mostrar que

(a) a aplicação $T \to UTU^{-1}$ é um isomorfismo do espaço vetorial $L(V, V)$ no espaço vetorial $L(W, W)$;
(b) traço $(UTU^{-1}) =$ traço(T) para todo T em $L(V, V)$;
(c) $UT_{\alpha,\beta}U^{-1} = T_{U\alpha,\, U\beta}$ ($T_{\alpha,\beta}$ definido no Exercício 10);
(d) $(UTU^{-1})^* = UT^*U^{-1}$;

(e) se equiparmos $L(V, V)$ com o produto interno $(T_1, T_2) =$ traço $(T_1 T_2^*)$. e anàlogamente para $L(W, W)$, então $T \rightarrow UTU^{-1}$ é um isomorfismo de espaço com produto interno.

14. Se V é um espaço com produto interno, um **movimento rígido** é uma função qualquer T de V em V (não necessàriamente linear) tal que $||T\alpha - T\beta|| = \alpha - \beta||$ para todos α, β em V. Um exemplo de um movimento rígido é um operador linear unitário. Outro exemplo é uma translação por um vetor fixo γ.

$$T_\gamma(\alpha) = \alpha + \gamma$$

(a) Seja V o espaço R^2 com o produto interno canônico. Suponhamos que T seja um movimento rígido de V e que $T(0) = 0$. Demonstrar que T é linear e é um operador unitário.

(b) Usar o resultado da parte (a) para demonstrar que todo movimento rígido de R^2 é composto de uma translação, seguida de um operador unitário.

(c) Mostrar agora que um movimento rígido de R^2 é uma translação seguida de uma rotação ou então uma translação seguida de uma reflexão seguida de uma rotação.

15. Um operador unitário sôbre R^4 (com o produto interno canônico) é simplesmente um operador linear que conserva a forma quadrática

$$||(x, y, z, t)||^2 = x^2 + y^2 + z^2 + t^2$$

isto é, um operador linear U tal que $||U\alpha||^2 = ||\alpha||^2$ para todo α em R^4. Numa certa parte da teoria da relatividade, é de interêsse determinar os operadores lineares T que conservam a forma

$$||(x, y, z, t)||_L^2 = t^2 - x^2 - y^2 - z^2.$$

Contudo, $||\ ||_L^2$ não vem de um produto interno, mas de algo chamado "métrica de Lorentz" (a qual não estudaremos). Por esta razão, um operador linear T sôbre R^4 tal que $||T\alpha||_L^2 = ||\alpha||_L^2$, para todo α em R^4, é denominado **uma transformação de Lorentz.**

(a) Mostrar que a função U definida por

$$U(x, y, z, t) = \begin{bmatrix} t + z & y + iz \\ y - iz & t - x \end{bmatrix}$$

é um isomorfismo de R^4 no espaço vetorial H das 2×2 matrizes complexas auto-adjuntas.

(b) Mostrar que $||\alpha||_L^2 = \det\,(U\alpha)$.

(c) Suponhamos que T seja um operador linear (real) sôbre o espaço H das 2×2 matrizes auto-adjuntas. Mostrar que $L = U^{-1}TU$ é um operador linear sôbre R^4.

(d) Seja M uma 2×2 matriz complexa arbitrária. Mostrar que $T_M(A) = M^*AM$ define um operador linear T_M sôbre H. (É necessário verificar que T_M leva H em H.)

(e) Se M é uma 2×2 matriz tal que $|\det M| = 1$, mostrar que $L_M = U^{-1}T_M U$ é uma transformaçãp de Lorentz sôbre R^4.

(f) Encontrar uma transformação de Lorentz que não seja uma L_M.

8.6 Operadores Normais

O objetivo principal desta seção é a resolução do problema seguinte: Se T é um operador linear sôbre um espaço V de dimensão finita com produto interno, sob que condições V possui uma base ortonormal formada por vetores característicos de T? Em outras palavras, quando é que existe uma base *ortonormal* $\mathcal{B}$ de V, tal que a matriz de T em relação à base $\mathcal{B}$ seja diagonal?

Vamos iniciar deduzindo algumas condições necessárias sôbre T, que mostraremos subseqüentemente serem suficientes. Suponhamos que $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ seja uma base ortonormal de V com a propriedade

(8-19) $$T\alpha_j = c_j\alpha_j, \quad j = 1, \ldots, n.$$

Isto diz simplesmente que a matriz de T em relação à base ordenada $\mathcal{B}$ é a matriz diagonal com elementos diagonais $c_1, \ldots, c_n$. O operador adjunto T^* é representado em relação a esta mesma base ordenada pela matriz transposta conjugada, isto é, a matriz diagonal com elementos diagonais $\bar{c}_1, \ldots, \bar{c}_n$. Se V é um espaço real com produto interno, os escalares $c_1, \ldots, c_n$ são (evidentemente) reais e então temos $T = T^*$. Em outras palavras, se V é um espaço *real* de dimensão finita com produto interno e T é um operador linear para o qual existe uma base ortonormal de vetores característicos, então T deve ser auto-adjunto. Se V é um espaço complexo com produto interno, os escalares $c_1, \ldots, c_n$ não são necessàriamente reais, isto é, T não é necessàriamente auto-adjunto. Mas notemos que T deve satisfazer

(8-20) $$TT^* = T^*T$$

De fato, duas matrizes diagonais quaisquer comutam e como T e T^* são ambas representadas por matrizes diagonais em relação à base ordenada $\mathcal{B}$, temos (8-20). É um fato bastante notável que esta condição também seja suficiente para implicar a existência de uma base ortonormal formada por vetores característicos.

Definição. *Seja* V *um espaço de dimensão finita com produto interno e* T *um operador linear sôbre* V. *Dizemos que* T *é* **normal** *se comuta com seu adjunto, isto é,* TT* = T*T.

Todo operador auto-adjunto é normal, como também o é todo operador unitário. Todo múltiplo escalar de um operador normal é normal; contudo, somas e produtos de operadores normais não são em geral normais. Embora isso não seja de forma alguma necessário, iniciaremos nosso estudo de operadores normais considerando operadores auto-adjuntos.

Teorema 17. *Seja* V *um espaço com produto interno e* T *um operador linear auto-adjunto sôbre* V. *Todo valor característico de* T *é real. Vetores característicos de* T *associados a valores característicos distintos são ortogonais.*

Demonstração. Suponhamos que c seja um valor característico de T, isto é, que $T\alpha = c\alpha$ para algum vetor não-nulo α. Então

$$\begin{aligned} c(\alpha, \alpha) &= (c\alpha, \alpha) \\ &= (T\alpha, \alpha) \\ &= (\alpha, T\alpha) \\ &= (\alpha, c\alpha) \\ &= \bar{c}(\alpha, \alpha). \end{aligned}$$

Como $(\alpha, \alpha) \neq 0$, devemos ter $c = \bar{c}$. Suponhamos também que $T\beta = d\beta$ com $\beta \neq 0$. Então

$$\begin{aligned} c(\alpha, \beta) &= (T\alpha, \beta) \\ &= (\alpha, T\beta) \\ &= (\beta, d\beta) \\ &= \bar{d}(\alpha, \beta) \\ &= d(\alpha, \beta). \end{aligned}$$

Se $c \neq d$, então $(\alpha, \beta) = 0$.

Deve-se salientar que o Teorema 17 nada diz a respeito da existência de valores característicos ou de vetores característicos.

Teorema 18. *Em um espaço de dimensão finita positiva com produto interno, todo operador auto-adjunto possui um vetor característico (não-nulo).*

Demonstração. Seja V um espaço de dimensão n com produto interno, sendo $n > 0$ e seja T um operador auto-adjunto sôbre V. Tomemos uma base ortonormal $\mathcal{B}$ de V e seja $A = [T]_{\mathcal{B}}$. Como $T = T^*$ temos $A = A^*$. Seja agora W o espaço das $n \times 1$ matrizes sôbre C, com produto interno $(X, Y) = Y^*X$. Então $U(X) = AX$ define um operador linear auto-adjunto U sôbre W. O polinômio característico, det $(xI - A)$, é um polinômio de grau n sôbre o corpo dos números complexos; todo polinômio sôbre C de grau positivo possui uma raiz. Assim, existe um número complexo c tal que det $(cI - A) = 0$. Isto significa que $A - cI$ é singular, ou que existe uma X não-nula tal que $AX = cX$. Como o operador U (multiplicação por A) é auto-adjunto, decorre do Teorema 17 que c é real. Se V é um espaço vetorial real, podemos tomar X com elementos reais. De fato, nesse caso A e $A - cI$ têm elementos reais e como $A - cI$ é singular, o sistema $(A - cI)X = 0$ possui uma solução real não-nula X. Decorre que existe um vetor não-nulo α em V tal que $T\alpha = c\alpha$.

Diversos comentários devem ser feitos a respeito da demonstração. (1) A demonstração da existência de um X não-nulo tal que $AX = cX$ nada teve que ver com o fato de A ser hermitiana (auto-adjunta). Ela mostra que todo operador linear sôbre um espaço vetorial complexo de dimensão finita possui um vetor característico. No caso de um espaço real com produto interno, a auto-adjunção de A é usada de modo fundamental para nos dizer que cada valor característico de A é real e que portanto podemos encontrar um X conveniente com valores reais. (2) O argumento mostra que o polinômio característico de uma matriz auto-adjunta tem coeficientes reais, a despeito do fato de que a matriz possa não ter elementos reais. (3) A hipótese de V ser de dimensão finita é necessária para o teorema; um operador auto-adjunto sôbre um espaço de dimensão infinita com produto interno pode não ter nenhum valor característico.

Exemplo 23. Seja V o espaço vetorial das funções complexas (ou reais) contínuas, definidas sôbre o intervalo unitário $0 \leq t \leq 1$, com o produto interno

$$(f, g) = \int_0^1 f(t)\overline{g(t)}dt.$$

O operador "multiplicação por t", $(Tf)(t) = tf(t)$, é auto-adjunto. Suponhamos que $Tf = cf$. Então

$$(t - c)f(t) = 0, \quad 0 \leq t \leq 1$$

e então $f(t) = 0$ para $t \neq c$. Como f é contínua, $f = 0$. Logo T não possui valores (vetores) característicos.

Teorema 19. *Seja* V *um espaço de dimensão finita com produto interno e seja* T *um operador linear arbitrário sôbre* V. *Suponhamos que* W *seja um subespaço de* V *que seja invariante sob* T. *Então o suplementar ortogonal de* W *é invariante sob* T.

Demonstração. Recordamos que o fato de W ser invariante sob T não quer dizer que cada vetor em W permaneça fixo por meio de T; significa que se α está em W então $T\alpha$ está em W. Seja β em $W^{\perp}$. Precisamos mostrar que $T^*\beta$ está em $W^{\perp}$, isto é, que $(\alpha, T^*\beta) = 0$ para todo α em W. Se α está em W, então $T\alpha$ está em W, portanto $(T\alpha, \beta) = 0$. Mas $(T\alpha, \beta) = (\alpha, T^*\beta)^*$.

Teorema 20. *Seja* V *um espaço de dimensão finita com produto interno e seja* T *um operador linear auto-adjunto sôbre* V. *Então existe uma base ortonormal de* V, *cujos vetores são vetores característicos de* T.

Demonstração. Estamos supondo dim $V > 0$. Pelo Teorema 18, T possui um vetor característico α. Seja $\alpha_1 = \alpha/||\alpha||$ de modo que α_1 também é um vetor característico de T e $||\alpha_1|| = 1$. Se dim $V = 1$, já terminamos. Vamos agora proceder por indução sôbre a dimensão de V. Suponhamos que o teorema seja válido para espaços com produto interno de dimensão menor que dim V. Seja W o subespaço unidimensional gerado pelo vetor α_1. A afirmação de que α_1 é um vetor característico de T significa simplesmente que W é invariante sob T. Pelo Teorema 19, o suplementar ortogonal $W^{\perp}$ é invariante sob $T^* = T$. Ora $W^{\perp}$, com o produto interno de V, é um espaço com produto interno de dimensão um a menos que a dimensão de V. Seja U o operador linear induzido sôbre $W^{\perp}$ por T, isto é, a restrição de T a $W^{\perp}$. Então U é auto-adjunto e, pela hipótese de indução, $W^{\perp}$ possui uma base ortonormal $\{\alpha_2, \ldots, \alpha_n\}$ formada por vetores característicos de U. Ora, cada um dêsses vetores também é um vetor característico de T e como $V = W \oplus W^{\perp}$, concluímos que $\{\alpha_1, \ldots, \alpha_n\}$ é a desejada base de V.

Corolário. *Seja* A *uma* n $\times$ n *matriz hermitiana (auto-adjunta). Então existe uma matriz unitária* P *tal que* P*AP *seja diagonal* (A *é unitàriamente equivalente a uma matriz diagonal). Se* A *é uma matriz simétrica real, existe uma matriz ortogonal real* P *tal que* P^tAP *seja diagonal.*

Demonstração. Seja V o espaço C^n, com o produto interno canônico e seja T o operador linear sôbre V que é representado por A em relação à base ordenada canônica. Como $A = A^*$, temos $T = T^*$. Seja $\mathfrak{B} = \{\alpha_1, \ldots, \alpha_n\}$ uma base ortonormal ordenada de V, tal que $T\alpha_j = c_j\alpha_j$, $j = 1, \ldots, n$. Se $D = [T]_{\mathfrak{B}}$, então D é a matriz diagonal com elementos diagonais $c_1, \ldots, c_n$. Seja P a matriz com vetores-colunas $\alpha_1, \ldots, \alpha_n$. Então $D = P^*AP$.

Caso todo elemento de A seja real, podemos tomar V como sendo R^n, com o produto interno canônico e repetir o argumento. Neste caso, P será uma matriz unitária com elementos reais, ou seja, uma matriz ortogonal real.

Caso todo elemento de A seja real, podemos tomar V como sendo R^n, com o produto interno canônico e repetir o argumento.

Neste caso, P será uma matriz unitária com elementos reais, ou seja, uma matriz ortogonal real.

Combinando o Teorema 20 com nossos comentários no início desta seção, temos o seguinte: Se V é um espaço *real* de dimensão finita com produto interno e T é um operador linear sôbre V, então V possui uma base ortonormal formada por vetores característicos de T se, e sòmente se, T é auto-adjunto. Equivalentemente, se A é uma $n \times n$ matriz com elementos *reais*, existe uma matriz ortogonal real P tal que P^tAP seja diagonal se, e sòmente se, $A = A^t$. Não existe nenhum resultado semelhante para matrizes simétricas complexas. Em outras palavras, para matrizes complexas, uma diferença significativa entre as condições $A = A^t$ e $A = A^*$.

Tendo resolvido o caso de operadores auto-adjuntos, voltamos ao estudo dos operadores normais em geral. Vamos demonstrar o análogo do Teorema 20 para operadores normais, no caso *complexo*.

Teorema 21. *Seja* V *um espaço de dimensão finita com produto interno e* T *um operador normal sôbre* V. *Então todo vetor característico de* T *também é um vetor característico de* T*.

Demonstração. Suponhamos que U seja um operador normal arbitrário sôbre V. Então $||U\alpha|| = ||U^*\alpha||$ para todo α em V. De fato, usando $UU^* = U^*U$ vemos que

$$\begin{aligned} ||U\alpha||^2 &= (U\alpha, U\alpha) = (\alpha, U^*U\alpha) \\ &= (\alpha, UU^*\alpha) = (U^*\alpha, U^*\alpha) = ||U^*\alpha||^2. \end{aligned}$$

Se T é normal e c é um escalar, o operador $U = T - cI$ é normal, pois $(T - cI)^* = T^* - \bar{c}I$. Então, para cada α em V

$$||(T - cI)\alpha|| = ||(T^* - \bar{c}I)||.$$

Em particular, $(T - cI)\alpha = 0$ se, e sòmente se, $(T^* - \bar{c}I)\alpha = 0$. Assim, α é um vetor característico de T com valor característico c se, e sòmente se, α é um vetor característico de T^* com valor característico $\bar{c}$.

Teorema 22. *Seja* V *um espaço complexo de dimensão finita com produto interno e seja* T *um operador normal sôbre* V. *Então* V *possui uma base ortonormal, onde cada vetor é um vetor característico de* T.

Demonstração. Como estamos trabalhando sôbre um espaço vetorial complexo, T possui um vetor característico α_1, que, podemos supor, satisfaz $||\alpha_1|| = 1$. Seja W o subespaço gerado por α_1, de modo que W é invariante sob T. Pelo Teorema 21, W também é invariante sob T^*. Decorre do Teorema 19 que $W^\perp$ é invariante sob T^{**}. Mas $T^{**} = T$ e então $W^\perp$ é invariante sob T. A restrição de T a $W^\perp$ é um operador normal, sendo o adjunto desta restrição a restrição de T^* a $W^\perp$. Agora basta repetir o argumento usado na demonstração do Teorema 20.

Novamente existe a interpretação matricial. Talvez esta definição proceda.

Definição. *Uma* $n \times n$ *matriz complexa* A *é dita* **normal** *se* $AA^* = A^*A$.

Corolário. *Seja* A *uma* $n \times n$ *matriz com elementos complexos. Existe uma matriz unitária* P *tal que* P^*AP *seja diagonal se, e sòmente se,* $AA^* = A^*A$. *Em outras palavras,* A *é unitàriamente equivalente a uma matriz diagonal se, e sòmente se,* A *é normal.*

Exercícios

1. Para cada uma das seguintes matrizes simétricas reais A, encontrar uma matriz ortogonal real P tal que P^tAP seja diagonal.

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \quad \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{bmatrix}$$

2. Uma matriz simétrica complexa é auto-adjunta? É normal?

3. Para

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}$$

existe uma matriz ortogonal real P tal que $P^tAP = D$ seja diagonal. Determinar esta matriz diagonal D.

4. Seja V o espaço C^2, com o produto interno canônico. Seja T o operador linear sôbre V que é representado em relação à base ordenada canônica pela matriz

$$A = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}.$$

Mostrar que T é normal e determinar uma base ortonormal de V, constituída de vetores característicos de T.

5. Dar um exemplo de uma 2×2 matriz A tal que A^2 seja normal, mas A não seja normal.

6. Seja T um operador normal sôbre um espaço complexo de dimensão finita com produto interno. Demonstrar que T é auto-adjunto, positivo ou unitário conforme todo valor característico de T seja real, positivo ou de valor absoluto 1. (Usar o Teorema 22 para reduzir o problema a um semelhante relativo a matrizes diagonais.)

7. Seja T um operador linear sôbre o espaço V de dimensão finita com produto interno e suponhamos que T seja positivo e unitário. Demonstrar que $T = I$.

8. Demonstrar que T é normal se, e sòmente se, $T = T_1 + iT_2$, onde T_1 e T_2 são operadores auto-adjuntos que comutam.

9. Demonstrar que uma matriz simétrica real possui uma raiz cúbica simétrica real; isto, é se A é simétrica real, existe uma B simétrica real tal que $B^3 = A$.

10. Demonstrar que tôda matriz positiva é o quadrado de uma raiz positiva.

11. Demonstrar que se um operador é normal e nilpotente então êle é o operador nulo.

12. Se T é um operador normal, demonstrar que vetores característicos de T associados a valores característicos distintos são ortogonais.

13. Seja T um operador normal sôbre um espaço complexo de dimensão finita com produto interno. Demonstrar que existe um polinômio f, com coeficientes complexos, tal que $T^* = f(t)$. (Representar T por meio de uma matriz diagonal e verificar o que f deve ser.)

14. Se dois operadores normais comutam, demonstrar que o seu produto é normal.

8.7 Teorema Espectral

Em tôda esta seção, V será um espaço complexo de dimensão finita com produto interno. Vamos demonstrar o teorema espectral que afirma que todo operador normal T pode ser expresso sob a forma

$$T = c_1E_1 + c_2E_2 + \ldots + c_kE_k$$

onde $E_1, \ldots, E_k$ são projeções ortogonais duas a duas ortogonais entre si, $E_i E_j = 0$ para $i \neq j$. Abreviadamente (e um tanto despreocupadamente) todo operador normal é uma combinação linear de projeções ortogonais. Ora, êste teorema pode ser deduzido de modo bastante fácil a partir do teorema de diagonalização para operadores normais (Teorema 22); contudo, gostaríamos de fazer uma demonstração independente. A demonstração que fizemos do Teorema 22 foi uma demonstração "especial", isto é, ela lidou diretamente com a geometria do espaço com produto interno e com o modo como o operador normal se comporta em relação a esta geometria. Faremos uma demonstração do teorema espectral que é quase puramente algébrica. Nenhuma demonstração dêste teorema pode ser (ou deve ser) completamente divorciada da geometria do espaço subjacente V; entretanto, sentimos que uma "demonstração algébrica" merece ser feita, por duas razões. (1) Essa demonstração vai relacionar o tratamento dos operadores normais com o tratamento, no Capítulo 6, dos operadores diagonalizáveis sôbre um espaço vetorial sem nenhum produto interno particular. (2) Em anos recentes, métodos algébricos colaboraram para uma demonstração elegante do teorema espectral para operadores normais em certos espaços particulares de dimensão infinita com produto interno, conhecidos como "espaços de Hilbert".

Seja T um operador normal sôbre V. Desejamos escrever T como uma combinação linear de projeções ortogonais. Antes de chegarmos ao âmago da demonstração de que isto é possível, parece aconselhável que nos familiarizemos um pouco mais com projeções ortogonais. Recordemos que se W é um subespaço de V, então $V = W \oplus W^{\perp}$, ou seja, cada vetor α em V pode ser expresso de um único modo sob a forma $\alpha = \beta + \gamma$, com β em W e γ no suplementar ortogonal $W^{\perp}$. A projeção ortogonal de V sôbre W é simplesmente o operador linear E definido por $E\alpha = \beta$. Em geral, existem muitas maneiras de "projetar" V sôbre W. Todo operador linear E que tem W por sua imagem e satisfaz $E^2 = E$ é uma tal projeção. A projeção ortogonal é distinguida pelo fato de que seu núcleo é $W^{\perp}$, o suplementar ortogonal de sua imagem. É útil ter um critério para distinguir as projeções ortogonais e com tal fim demonstraremos o que segue.

Teorema 23. *Seja* E *uma projeção,* $E^2 = E$. *As seguintes afirmações são equivalentes.*

(i) E *é normal,* $EE^* = E^*E$;

(ii) E *é auto-adjunto,* $E = E^*$;

(iii) E *é a projeção ortogonal sôbre sua imagem.*

Demonstração. (i) $\rightarrow$ (ii). De $EE^* = E^*E$, temos $||E\beta|| = ||E^*\beta||$ para todo β em V. De fato

$$(E\beta, E\beta) = (\beta, E^*E\beta) = (\beta, EE^*\beta) = (E^*\beta, E^*\beta).$$

Em particular, $E\beta = 0$ se, e sòmente se, $E^*\beta = 0$. Seja α um vetor arbitrário em V e seja $\beta = \alpha - E\alpha$. Então

$$E\beta = E\alpha - E^2\alpha = E\alpha - E\alpha = 0$$

logo

$$0 = E^*\beta = E^*\alpha - E^*E\alpha.$$

Portanto $E^* = E^*E$ e

$$E = E^{**} = (E^*E)^* = E^*E = E^*.$$

(ii) $\rightarrow$ (iii). Suponhamos que $E = E^*$. Então, o núcleo de E é o suplementar ortogonal de sua imagem. De fato, $(E\alpha, \beta) = (\alpha, E\beta)$, portanto, β é ortogonal a todo $E\alpha$ se, e sòmente se, $E\beta$ é ortogonal a V, isto é, $E\beta = 0$. Logo, E é a projeção ortogonal sôbre sua imagem.

(iii) $\rightarrow$ (i). Suponhamos que E seja uma projeção ortogonal. Sejam α, β vetores arbitrários em V. Então $\alpha - E\alpha$ está no núcleo de E e $E\beta$ está na imagem de E, logo

$$\begin{aligned} 0 &= (\alpha - E\alpha, E\beta) \\ &= (\alpha, E\beta) - (E\alpha, E\beta) \\ &= (\alpha, E\beta) - (\alpha, E^*E\beta). \end{aligned}$$

Como isto vale para todos α, β, concluímos que $E = E^*E$. Assim, E é auto-adjunta, logo normal.

Sejam $W_1, \ldots, W_k$ subespaços de V. Lebremos (Seção 6.1) que V é a soma direta de $W_1, \ldots, W_k$, isto é, $V = W_1 \oplus \ldots \oplus W_k$, se todo vetor α em V pode ser expresso *de um único modo* sob a forma

$$\alpha = \alpha_1 + \ldots + \alpha_k, \quad \alpha_j \text{ em } W_j.$$

Em um espaço com produto interno, interessam-nos de maneira especial as decomposições em soma direta que são **ortogonais**, isto é, nas quais todo vetor em W_i é ortogonal a todo vetor em W_i para i $\neq j$.

Teorema 24. *Sejam* $W_1, \ldots, W_k$ *subespaços de* V *e seja* E_j *a projeção sôbre* W_j, $j = 1, \ldots, k$. *As seguintes condições são equivalentes.*

(i) $V = W_1 \oplus \ldots \oplus W_k$, *e esta é uma soma direta ortogonal.*

(ii) $I = E_1 + \ldots + E_k$, *e* $E_iE_j = 0$ *para* $i \neq j$.

(iii) *Se* $\mathcal{B}_j$ *é uma base ortonormal de* W_j, $j = 1, \ldots k$, *então a reunião das* $\mathcal{B}_j$ *é uma base ortònormal de* V.

Demonstração. Vamos esboçar a demonstração. Se o leitor encontrar dificuldades em alguns detalhes, deverá ver as demonstrações dos Teoremas 1, 2 e 3 do Capítulo 6, que contém todo o argumento, exceto pelas questões sôbre ortogonalidade.

Suponhamos que (ii) seja válida. Se α está em V, então, de $I = E_1 + \ldots + E_k$ temos

$$\alpha = E_1\alpha + \ldots + E_k\alpha.$$

Assim, α é uma soma de vetores, um de cada subespaço W_j. Se $\alpha = \alpha_1 + \ldots + \alpha_k$ com α_j em W_j, então

$$E_j\alpha_j = \alpha_j \quad \text{e} \quad \alpha = E_1\alpha_1 + \ldots + E_k\alpha_k.$$

Portanto

$$E_j\alpha = E_jE_1\alpha + \ldots + E_jE_k\alpha = E_j^2\alpha_j = E_j\alpha_j = \alpha_j$$

Conseqüentemente, a única expressão de α com uma soma de vetores dos subespaços W_j é $\alpha = E_1\alpha + \ldots + E_k\alpha$; logo

$$V = W_1 \oplus \ldots \oplus W_k.$$

Para ver que esta é uma soma direta ortogonal, observemos que

$$(E_i\alpha, E_j\beta) = (\alpha, E_i^* E_j\beta) = (\alpha, E_iE_j\beta) = (\alpha, 0) = 0.$$

Usamos aqui o fato de E_i ser auto-adjunto, por ser uma projeção ortogonal. Então (ii) implica (i). Não é difícil inverter a argumentação, para mostrar que (i) implica (ii).

Vejamos por que é que (iii) é equivalente a (i). Se (i) vale, cada α em V é da forma $\alpha = \alpha_1 + \ldots + \alpha_k$ com α_j em W_j e $(\alpha_i, \alpha_j) = 0$ para $i \neq j$. Se $\mathcal{B}_j$ é uma base ortonormal de W_j, $j = 1, \ldots, k$, seja $\mathcal{B}$ a reunião das $\mathcal{B}_j$. Então $\mathcal{B}$ gera V, pois se α está em V, podemos escrever α_j como uma combinação linear dos vetores em $\mathcal{B}_j$ e a expressão $\alpha = \alpha_1 + \ldots + \alpha_k$ fornece α como uma combinação linear dos vetores em $\mathcal{B}$. Além disso, $\mathcal{B}$ é um conjunto ortonormal, pois $\mathcal{B}_j$ é ortonormal e todo vetor em $\mathcal{B}_i$ é ortogonal a todo vetor em $\mathcal{B}_j$ se $i \neq j$. Logo, $\mathcal{B}$ é uma base ortonormal. Novamente, deixamos a cargo do leitor a inversão desta argumentação.

A grosso modo, os Teoremas 23 e 24 são a parte geométrica de nossa demonstração do teorema espectral. O efeito dêsses teoremas é traduzir afirmações sôbre a ortogonalidade de subespaços em afirmações algébricas sôbre projeções e suas adjuntas.

Voltemos agora ao nosso operador normal T. Vamos estudar T considerando a álgebra dos polinômios em T, isto é, a coleção de operadores lineares da forma $f(T)$, onde f é um polinômio com coeficientes complexos. Necessitaremos de três fatos básicos, que enunciaremos sob a forma de lemas.

Lema. *Seja* T *um operador normal e* α *um vetor tal que* $T^2\alpha = 0$. *Então* $T\alpha = 0$. *Em outras palavras, a imagem e o núcleo de um operador normal são disjuntos.*

Demonstração. Suponhamos que $T^2\alpha = 0$. Seja $\beta = T\alpha$, de modo que $T\beta = 0$. Usando $TT^* = T^*T$ temos $||T\beta|| = ||T^*\beta||$ e assim $T^*\beta = 0$. Conseqüentemente,

$$\begin{aligned} 0 &= (T^*\beta, \alpha) \\ &= (\beta, T\alpha) \\ &= (\beta, \beta) \end{aligned}$$

e $\beta = 0$.

A afirmação de que $T^2\alpha = 0$ implica $T\alpha = 0$ diz simplesmente que a imagem e o núcleo de T são disjuntos. De fato, ela diz que se β está na imagem de T, $\beta = T\alpha$, e $T\beta = 0$, então $\beta = 0$.

Lema. *Seja* T *um operador normal e* f *um polinômio com coeficientes complexos. Então o operador* f(T) *é normal.*

Demonstração. Se $f = a_0 + a_1x + \ldots + a_kx^k$, então

$$f(T) = a_0I + a_1T + \ldots + a_kT^k.$$

Assim,

$$f(T)^* = \bar{a}_0I + \bar{a}_1T^* + \ldots + \bar{a}_k(T^*)^k.$$

Como $TT^* = T^*T$ é fácil ver que $f(T)$ comuta com $f(T)^*$.

Lema. *O polinômio minimal de um operador normal possui raízes distintas.*

Demonstração. Lembramos (Seção 6.2) que o polinômio minimal de T é o polinômio unitário de menor grau entre os polinômios f tais que $f(T) = 0$. Seja p êste polinômio minimal. Vamos mostrar que $p = (x - c_1) \ldots (x - c_k)$, onde $c_1, \ldots, c_k$ são números complexos distintos. Suponhamos que isto não ocorra, isto é, que alguma raiz c de p seja múltipla. Isto significa que $p = (x - c)^2g$ para algum polinômio g. Como $p(T) = 0$, temos $(T - cI)^2g(T) = 0$, ou

$$(T - cI)^2g(T)\alpha = 0$$

para todo α em V. Pelo último lema, o operador $U = T - cI$ é normal. Seja α em V e seja $\beta = g(T)\alpha$. Então

$$U^2\beta = (T - cI)^2g(T)\alpha = 0.$$

Então, pelo primeiro lema acima, temos $U\beta = 0$. Assim,

$$(T - cI)g(T)\alpha = 0$$

para todo α, ou seja $(T - cI)g(T) = 0$. Mas isto contradiz a hipótese de que p tem o menor grau dentre todos f tais que $f(T) = 0$.

Vamos agora escrever T como uma combinação linear de projeções ortogonais. O método que vamos usar será exatamente o mesmo que usamos no estudo de operadores diagonalizáveis no Capítulo 6. Se alguns detalhes não estiverem claros, insistimos em que o leitor consulte as demonstrações dos Teoremas 8, 9 e 10 do Capítulo 6.

Como T é normal, o polinômio minimal p de T é da forma

$$p = (x - c_1) \ldots (x - c_k)$$

onde $c_1, \ldots, c_k$ são números complexos *distintos*. Ora, $c_1, \ldots, c_k$ são, na verdade. os valores característicos de T, mas não precisamos saber isto agora, uma vez que logo se tornará evidente. Vamos exprimir T sob a forma $T = c_1E_1 + \ldots + c_kE_k$, onde $E_1, \ldots, E_k$ são projeções ortogonais que satisfazem $I = E_1 + \ldots + E_k$ e $E_iE_j = 0$ para $i \neq j$. Para compreender o método que usamos, suponhamos que tivéssemos as projeções $E_1, \ldots, E_k$. Então teríamos

$$\begin{aligned} T^2 &= \left(\sum_{i=1}^{k} c_iE_i\right)\left(\sum_{j=1}^{k} c_jE_j\right) \\ &= \sum_i \sum_j c_ic_jE_iE_j \\ &= \sum_j c_j^2E_j^2 \\ &= \sum_j c_j^2E_j. \end{aligned}$$

Da forma mais geral, teríamos

$$T^r = \sum_{j=1}^{k} c_j^rE_j$$

e, tomando combinações lineares,

$$g(T) = \sum_{j=1}^{k} g(c_j)E_j$$

para todo polinômio g. Isto nos diz como conseguir as projeções $E_1, \ldots, E_k$. De fato, suponhamos que p_j seja um polinômio tal que $p_j(c_i) = \delta_{ij}$. Então

$$\begin{aligned} p_j(T) &= \sum_{i=1}^{n} p_j(c_i)E_i \\ &= \sum_{i=1}^{n} \delta_{ij}E_i \\ &= E_j. \end{aligned}$$

Ora, começamos com T e seu polinômio minimal

$$p = (x - c_1) \ldots (x - c_k),$$

seja $\quad p_j = \dfrac{(x - c_1) \ldots (x - c_{j-1})(x - c_{j+1}) \ldots (x - c_k)}{(c_j - c_1) \ldots (c_j - c_{j-1})(c_j - c_{j+1}) \ldots (c_j - c_k)}.$

Então $p_j(c_i) = \delta_{ij}$. Os polinômios $p_1, \ldots, p_k$ são os polinômios de

Lagrange determinados por $c_1, \ldots, c_k$. Êstes polinômios têm a propriedade de que se f é um polinômio qualquer de *grau no máximo* $k - 1$, então

$$f = f(c_1)p_1 + \ldots + f(c_k)p_k.$$

Isto decorre do fato de que

$$f - f(c_1)p_1 - \ldots - f(c_k)p_k$$

é um polinômio de grau máximo $k - 1$ que, em vista de $p_j(c_i) = \delta_{ij}$, tem as k raízes distantes $c_1, \ldots, c_k$. Em particular, temos

$$1 = p_1 + \ldots + p_k$$
$$x = c_1p_1 + \ldots c_kp_k.$$

(Estamos supondo $k \leqslant 2$; em caso contrário, T é um múltiplo escalar do operador idêntico e então tudo está certo.) Definamos agora $E_j = p_j(T)$. A equação acima nos fornece então

$$I = E_1 + \ldots + E_k$$
$$T = c_1E_1 + \ldots + c_kE_k.$$

Façamos algumas observações.

(i) $E_j \neq 0$, para cada j. De fato, $E_j = p_j(T)$ e p_j é um polinômio de grau $k - 1$. Como p é o polinômio minimal de T, não podemos ter $p_j(T) = 0$.

(ii) $E_iE_j = 0$ se $i \neq j$. De fato, $E_iE_j = p_i(T)p_j(T)$. Quando $i \neq j$, o polinômio $h = p_ip_j$ é divisível por p. Assim, $h(T) = p_i(T)$ $p_j(T) = 0$.

(iii) Cada E_j é uma projeção. Como

$$I = E_1 + \ldots + E_k \text{ e } E_iE_j = 0$$

para $i \neq j$, temos

$$\begin{aligned} E_j &= E_jE_1 + \ldots + E_jE_k \\ &= E_j^2. \end{aligned}$$

(iv) $c_1, \ldots, c_k$ são exatamente os valores característicos distintos de T. Primeiro, c_j é um valor característico pela seguinte razão: Seja α um vetor não-nulo da imagem de E_j. Então $\alpha = E_j\alpha$, portanto

$$\begin{aligned} T\alpha &= TE_j\alpha \\ &= E_jT\alpha \\ &= E_j(c_1E_1\alpha + \ldots + c_kE_k\alpha) \\ &= c_1E_jE_1\alpha + \ldots + c_kE_jE_k\alpha \\ &= c_jE_j\alpha \\ &= c_j\alpha. \end{aligned}$$

Recìprocamente, suponhamos que c seja um valor característico de T. Seja α um vetor tal que $(T - cI)\alpha = 0$. Isto diz

$$\sum_{j=1}^{k} (c_j -)E_j\alpha = 0.$$

Aplicando E_i a esta expressão obtemos

$$(c_i - c)E_i\alpha = 0$$

para $i = 1, \ldots, k$. Se $c_i \neq c$ para cada i, então temos $E_i\alpha = 0$ para todo α, logo

$$\alpha = E_1\alpha + \ldots + E_k\alpha = 0.$$

Assim, c sendo um valor característico, deve ser igual a algum c_i.

(v) A imagem de E_j é o espaço dos vetores característicos de T que são associados ao valor característico c_j. Em outras palavras, α está na imagem de E_j se, e sòmente se, $T\alpha = c_j\alpha$. Primeiro, suponhamos que $T\alpha = c_j\alpha$. Então

$$\sum_{i=1}^{k} c_iE_i\alpha = c_j \sum_{i=1}^{k} E_i\alpha$$

ou

$$\sum_{i=1}^{k} (c_i - c_j)E_i\alpha = 0.$$

A partir disto concluímos que $E_i\alpha = 0$ para $i \neq j$ e portanto que $\alpha = E_j\alpha$. Suponhamos agora que α esteja na imagem de E_j, isto é, que $\alpha = E_j\alpha$. Na parte (iv) mostraremos que $T\alpha = c_j\alpha$.

(vi) E_j é a projeção ortogonal sôbre sua imagem. De fato, E_j é um polinômio no operador normal T. Assim, E_j é uma projeção normal e, pelo Teorema 23, é uma projeção ortogonal.

Teorema 25 (Teorema Espectral). *Seja* T *um operador normal sôbre o espaço complexo* V *de dimensão finita com produto interno. Sejam* $c_1, \ldots, c_k$ *os valores característicos distintos de* T *e seja* E_j *a* *projeção ortogonal sôbre o espaço dos vetores característicos associados ao valor característico* c_j. *Então*

(a) $T = c_1E_1 + \ldots + c_kE_k$;

(b) $I = E_1 + \ldots + E_k$;

(c) $E_iE_j = 0$, *se* $i \neq j$.

Além disso, a decomposição na parte (a) é única, no seguinte sentido. Suponhamos que $c_1, \ldots, c_k$ *sejam números complexos distintos e* $E_1, \ldots, E_k$, *operadores lineares não-nulos sôbre* V *tais que as condições (a), (b) e (c) estejam satisfeitas. Então* $c_1, \ldots, c_k$ *são exatamente os valores característicos distintos de* T *e, para cada* i, E_i *é a projeção ortogonal de* V *sôbre o espaço dos vetores característicos de* T *associado ao valor característico* c_j.

Demonstração. Já fizemos acima tôda a demonstração. Talvez sejam oportunos alguns comentários sôbre a unicidade. Salientamos que se c_j e E_j satisfazem (a), (b) e (c) e se $p_j(c_i) = \delta_{ij}$, então $E_j = p_j(T)$. Isto determina E_j de um único modo como um polinômio bem determinado em T, uma vez que já mostramos que $c_1, \ldots, c_k$ são os valores característicos. Isto, e mais o restante da demonstração da unidade, estão contidos nas observações (iii)-(iv) acima. Basta notar que (iii)-(vi) foram demonstradas usando apenas (a), (b), (c), o fato dos c_i serem distintos e mais o fato de que nenhum dos E_j é o operador nulo.

Denominaremos a decomposição $T = c_1E_1 + \ldots + c_kE_k$ a **resolução espectral** de T. Deveríamos ressaltar que se T é um operador linear arbitrário que tem uma resolução

$$T = c_1E_1 + \ldots + c_kE_k$$

tal que $E_j = E_j^*$ e $E_iE_j = 0$ para $i \neq j$, então T é normal. De fato, $E_j^* = E_j$ implica

$$T^* = \bar{c}_1E_1 + \ldots + \bar{c}_kE_k$$

e como $E_1, \ldots, E_k$ comutam porque $E_iE_j = E_jE_i = 0$ para $i \neq j$, é claro que T e T^* comutam. Pode-se descrever isto dizendo que operadores normais sôbre V são caracterizados como sendo os operadores que podem ser expressos como uma combinação linear de projeções ortogonais que comutam.

Corolário. *Se* T *é um operador normal, então* V *possui uma base ortonormal formada de vetores característicos de* T.

Demonstração. Seja W_j o espaço dos α tais que $T\alpha = c_j\alpha$, e seja $\mathcal{B}_j$ uma base ortonormal de W_j. Então, a reunião das $\mathcal{B}_j$ é uma base ortonormal de V. Isto decorre das partes (b) e (c) do teorema espectral e do Teorema 24. Como todo vetor em W_j é um vetor característico de T, a demonstração está completa.

Façamos alguns comentários. Na demonstração do corolário, que é exatamente o Teorema 22, o fato de os diversos espaços W_j serem ortogonais entre si é exatamente a afirmação de que vetores característicos de um operador normal, que pertençam a valores característicos distintos, são ortogonais. Notemos também que o teorema espectral se esquece de mencionar um dos fatos importantes que demonstramos, a saber, que as projeções $E_1, \ldots, E_k$ são polinômios em T. Evidentemente, isto decorre das condições (a), (b) e (c) do teorema. De maneira mais geral, temos

$$g(T) = g(c_1)E_1 + \ldots + g(c_k)E_k$$

para todo polinômio g. A partir disto temos o seguinte:

Corolário. *Seja* T *um operador linear sôbre* V. *Então* T *é normal se, e sòmente se, o adjunto* T* *é um polinômio em* T.

Demonstração. Se existe um polinômio f tal que $T^* = f(T)$, então, òbviamente, T e T^* comutam. Recìprocamente, suponhamos que T seja normal. Seja

$$T = c_1E_1 + \ldots + c_kE_k$$

a resolução espectral de T. Como $E_j^* = E_j$, temos

$$T^* = \bar{c}_1E_1 + \ldots + \bar{c}_kE_k.$$

Ora, $c_1, \ldots, c_k$ são números complexos distintos, portanto existe um polinômio f tal que $f(c_j) = \bar{c}_j$, $j = 1, \ldots, k$. Podemos tomar

$$f = \bar{c}_1p_1 + \ldots + \bar{c}_kp_k,$$

onde $p_1, \ldots, p_k$ são os polinômios utilizados na demonstração do Teorema 25. De $f(c_j) = \bar{c}_j$, é imediato que $T^* = f(T)$.

Talvez devamos observar agora os tipos especiais de operadores normais com os quais trabalhamos neste capítulo: auto-adjuntos, positivos e unitários. Se T é um operador normal com resolução espectral $T = c_1E_1 + \ldots + c_kE_k$, então $T^* = \bar{c}_1E_1 + \ldots + \bar{c}_kE_k$. Dizer que T é auto-adjunto é dizer que $T = T^*$, ou seja,

$$(c_1 - \bar{c}_1)E_1 + \ldots + (c_k - \bar{c}_k)E_k = 0.$$

Usando o fato de que $E_iE_j = 0$ para $i \neq j$ e o fato de que nenhum E_j é o operador nulo, vemos que T é auto-adjunto se, e sòmente se,

$c_j = \bar{c}_j$, $j = 1, \ldots, k$. Para distinguir os operadores normais que são positivos, observemos

$$\begin{aligned}(T\alpha, \alpha) &= \left(\sum_{j=1}^{k} c_j E_j\alpha, \sum_{i=1}^{k} E_i\alpha\right)\\ &= \sum_i \sum_j c_j(E_j\alpha, E_i\alpha)\\ &= \sum_j c_j\|E_j\alpha\|^2.\end{aligned}$$

Usamos o fato de que $(E_j\alpha, E_i\alpha) = 0$ para $i \neq j$. A partir disto é evidente que a condição $(T\alpha, \alpha) > 0$ para $\alpha > 0$ é satisfeita se, e sòmente se, $c_j > 0$ para todo j. Para distinguir os operadores unitários, observamos que

$$\begin{aligned}TT^* &= c_1\bar{c}_1E_1 + \ldots + c_k\bar{c}_kE_k\\ &= |c_1|^2E_1 + \ldots + |c_k|^2E_k.\end{aligned}$$

Se $TT^* = I$, então $I = |c_1|^2E_1 + \ldots + |c_k|^2E_k$ e, aplicando E_j,

$$E_j = |c_j|^2E_j.$$

Como $E_j \neq 0$, temos $|c_j|^2 = 1$ ou $|c_j| = 1$. Recìprocamente, se $|c_j|^2 = 1$ para cada j, é claro que $TT^* = I$.

Teorema 26. *Seja* T *um operador normal sôbre o espaço complexo* V *de dimensão finita com produto interno. Então,* T *é auto-adjunto, positivo ou unitário conforme tôdo valor característico de* T *seja real, positivo ou de valor absoluto* 1.

É importante notar que êste é um teorema sôbre operadores normais. Se T é um operador linear genérico sôbre V que possui valores característicos reais, não decorre que T é auto-adjunto. O teorema afirma que se T possui valores característicos reais e *se* T é normal, então T é auto-adjunto. Um teorema dêste tipo serve para fortalecer a analogia entre a operação de conjugação (operação de se tomar adjuntos) e o processo de se formar o conjugado de um número complexo. Um número complexo z é real, positivo ou de valor absoluto 1 conforme $z = \bar{z}$, $z = \bar{w}w$ para algum $w \neq 0$, ou $\bar{z}z = 1$, respectivamente. Um operador T é auto-adjunto, positivo ou unitário se, respectivamente, $T = T^*$, $T = U^*U$ para algum U inversível, ou $T^*T = I$.

Vamos agora demonstrar dois teoremas, que são os análogos destas duas afirmações. (1) Todo número não-negativo possui uma única raiz quadrada não-negativa. (2) Todo número complexo pode ser expresso sob a forma ru onde r é não-negativo e $|u| = 1$. Esta é a decomposição polar $z = re^{i\theta}$ para números complexos. Para êstes dois resultados vamos precisar da seguinte:

Definição. *Seja* T *um operador linear sôbre o espaço complexo* V *com produto interno. Dizemos que* T *é* **não-negativo** *se* $(T\alpha, \alpha) \geq 0$ *para todo* α *em* V.

Se $(T\alpha, \alpha) \geq 0$ para todo α, então, em particular, $(T\alpha, \alpha)$ é real para todo α. Assim T é auto-adjunto. Em um espaço real com produto interno, deveríamos acrescentar a condição $T = T^*$ à definição de um operador não-negativo. Se U é um operador linear arbitrário sôbre V, então, $T = U^*U$ é, como se vê fàcilmente, não-negativo:

$$(U^*U\alpha, \alpha) = (U\alpha, U\alpha) \geq 0.$$

Se U é inversível, então (òbviamente) T é positivo.

Se T é um operador normal sôbre V, então T é não-negativo se, e sòmente se, todo valor característico de T é não-negativo. Isto se demonstra da mesma maneira que o resultado análogo para operadores positivos.

Teorema 27. *Seja* V *um espaço complexo de dimensão finita com produto interno e seja* T *um operador não-negativo sôbre* V. *Então* T *possui uma única raiz quadrada não-negativa, isto é, existe um, e sòmente um, operador não-negativo* N *sôbre* V *tal que* $N^2 = T$.

Demonstração. Seja $T = c_1E_1 + \ldots + c_kE_k$ a resolução espectral de T. Existe essa resolução, pois T é não-negativo, portanto normal. O fato de T ser não-negativo nos diz que $c_j \geq 0$ para cada j. Indiquemos por $\sqrt{c_j}$ a única raiz quadrada não-negativa de c_j e seja

$$N = \sqrt{c_j}\, E_1 + \ldots + \sqrt{c_k}\, E_k.$$

Certamente N é normal e afirmamos que esta expressão é a resolução espectral de N. De fato, $I = E_1 + \ldots + E_k$, $E_iE_j = 0$ para $i \neq j$, nenhum E_j é 0 e $\sqrt{c_1}, \ldots, \sqrt{c_k}$ são distintos. Como $\sqrt{c_j} \geq 0$ para cada j, N é um operador não-negativo. É óbvio que $N^2 = T$.

Seja agora P um operador não-negativo sôbre V tal que $p^2 = T$. Vamos demonstrar que $P = N$. Seja

$$P = d_1F_1 + \ldots + d_rF_r$$

a resolução espectral de P. Então $d_j \geq 0$ para cada j, pois P é não-negativo. De $P^2 = T$ resulta

$$T = d_1^2F_1 + \ldots + d_r^2F_r.$$

Ora, $F_1, \dots, F_r$ satisfazem as condições $I = F_1 + \dots + F_r$, $F_iF_j = 0$ para $i \neq j$ e nenhum F_j é 0. Os números $d_1^2, \dots, d_r^2$ são distintos, porque números não-negativos distintos possuem quadrados distintos. Pela unicidade da resolução espectral de T, devemos ter $r = k$ e (talvez reordenando) $F_j = E_j$, $d_j^2 = c_j$. Assim $P = N$.

Teorema 28. *Seja* V *um espaço complexo de dimensão finita com produto interno e seja* T *um operador linear arbitrário sôbre* V. *Então, existe um operador unitário* U *sôbre* V *e um operador não-negativo* N *sôbre* V *tal que* T = UN. *O operador não-negativo* N *é único. Se* T *é inversível, o operador* U *tambem é único.*

Demonstração. Suponhamos que $T = UN$, onde U é unitário e N não-negativo. Então $T^* = (UN)^* = N^*U^* = NU^*$. Assim, $T^*T = NU^*UN = N^2$. Isto mostra que N é determinado de um único modo como a raiz quadrada não-negativa do operador não-negativo T^*T.

Então, para iniciarmos a demonstração da existência de U e N, usaremos o Teorema 27 para definirmos N como a única raiz quadrada não-negativa de T^*T. Se T é inversível, então N também o é porque

$$(N\alpha, N\alpha) = (N^2\alpha, \alpha) = (T^*T\alpha, \alpha) = (T\alpha, T\alpha).$$

Neste caso, definamos $U = TX^{-1}$ e demonstraremos que U é unitário. Ora, $U^* = (TN^{-1})^* = (N^{-1})^*T^* = (N^*)^{-1}T^* = N^{-1}T^*$. Assim,

$$\begin{aligned} UU^* &= TN^{-1}N^{-1}T^* \\ &= T(N^{-1})^2T^* \\ &= T(N^2)^{-1}T^* \\ &= T(T^*T)^{-1}T^* \\ &= TT^{-1}(T^*)^{-1}T^* \\ &= I \end{aligned}$$

e U é unitário.

Se T é não inversível, teremos de realizar um pouco mais de trabalho para definir U. Definamos primeiro U sôbre a imagem de N. Seja α um vetor na imagem de N, digamos, $\alpha = N\beta$. Definamos $U\alpha = T\beta$, motivados pelo fato de que queremos $UN\beta = T\beta$. Precisamos verificar que U está bem definida sôbre a imagem de N; em outras palavras, se $N\beta' = N\beta$, então $T\beta' = T\beta$. Verificamos acima que $||N\gamma||^2 = ||T\gamma||^2$ para todo γ em V. Assim, com $\alpha = \beta - \beta'$, vemos que $N(\beta - \beta') = 0$ se, e sòmente se, $T(\beta - \beta') = 0$. Portanto, U

está bem definida sôbre a imagem de N e é evidentemente linear onde definida. Se W é a imagem de N, vamos agora definir U sôbre $W^\perp$. Para fazer isto precisamos da seguinte observação: Como T e N possuem o mesmo núcleo, suas imagens têm a mesma dimensão. Assim, $W^\perp$ possui mesma dimensão que o suplementar ortogonal da imagem de T. Portanto, existe um isomorfismo (de espaço com produto interno) U_0 de $W^\perp$ em $T(V)^\perp$. Agora já definimos U sôbre W e vamos definir U sôbre $W^\perp$ como sendo V_0.

Repitamos a definição de U. Como $V = W \oplus W^\perp$, cada α em V pode ser expresso de um único modo sob a forma $\alpha = N\beta + \gamma$, onde $N\beta$ está na imagem W de N e γ está em $W^\perp$. Definamos

$$U\alpha = T\beta + U_0\gamma.$$

Êste U é evidentemente linear e, como verificamos acima, está bem definido. Além disso

$$\begin{aligned}(U\alpha, U\alpha) &= (T\beta + U_0\gamma, T\beta + U_0\gamma) \\ &= (T\beta, T\beta) + (U_0\gamma, U_0\gamma) \\ &= (N\beta, N\beta) + (\gamma, \gamma) \\ &= (\alpha, \alpha)\end{aligned}$$

portanto U é unitário. Também $UN\beta = T\beta$ para cada β.

Denominamos $T = UN$ a **decomposição polar** de T. Certamente não dizemos que é *a* decomposição polar, pois U não é único. Mesmo quando T é inversível, de modo que U é único, temos a dificuldade de que U e N podem não comutar. Na verdade, êles comutam se, e sòmente se, T é normal. Por exemplo, se $T = UN = NU$, com N não-negativo e U unitário.

$$TT^* = (NU)(NU)^* = NUU^*N = N^2 = T^*T.$$

O operador arbitrário T também possui uma decomposição $T = N_1U_1$, com N_1 não-negativo e U_1 unitário. Neste caso, N_1 será a raiz quadrada não-negativa de TT^*. Êste resultado pode ser obtido aplicando-se o teorema há pouco demonstrado ao operador T^* e depois tomando adjuntos.

Exercícios

1. Dar uma definição razoável de uma $n \times n$ matriz não-negativa e depois demonstrar que uma tal matriz possui uma única raiz quadrada não-negativa.

2. Seja T um operador normal e seja U um operador arbitrário que comute com T. Demonstrar que U cómuta com T^*.

3. Se U e T são operadores normais que comutam, demonstrar que $U + T$ e UT são normais.

4. Seja T um operador linear sôbre o espaço complexo V de dimensão finita com produto interno. Demonstrar que as dez afirmações seguintes sôbre T são equivalentes:

(i) T é normal.
(ii) $||T\alpha|| = ||T^*\alpha||$ para todos α em V.
(iii) $T = T_1 + iT_2$, onde T_1 e T_2 são auto-adjuntos e $T_1T_2 = T_2T_1$.
(iv) Se α é um vetor e c um escalar tal que $T\alpha = c\alpha$, então $T^*\alpha = \bar{c}\alpha$.
(v) Existe uma base ortonormal de V formada por vetores característicos de T.
(vi) Existe uma base ortonormal $\mathcal{B}$ tal que $[T]_{\mathcal{B}}$ seja diagonal.
(vii) Existe um polinômio g com coeficientes complexos tal que $T^* = g(T)$.
(viii) Todo subespaço que é invariante sob T também é invariante sob T^*.
(ix) $T = NU$, onde N é não-negativo, U é unitário e N comuta com U.
(x) $T = c_1E_1 + \ldots + c_kE_k$, onde $I = E_1 + \ldots + E_k$, $E_iE_j = 0$ para $i \neq j$, e $E_j^2 = E_j = E_j^*$.

8.8 Diagonalização Simultânea de Operadores Normais

Novamente, seja V um espaço complexo de dimensão finita com produto interno. O que vamos demonstrar é que se tivermos uma família arbitrária de operadores normais sôbre V, todos os quais comutam, existe um base ortonormal de V que diagonaliza todos êsses operadores simultâneamente.

Teorema 29. *Sejam* T *e* U *operadores normais sôbre* V *que comutam. Então existe um operador normal* S *sôbre* V *com estas propriedades*:

(i) *Tanto* T *como* U *são polinômios em* S.

(ii) S *comuta com todo operador linear sôbre* V *que comute com* T *e* U.

Demonstração. Sejam $T = c_1E_1 + \ldots + c_kE_k$ e $U = d_1F_1 + \ldots + d_rF_r$ as resoluções especiais de T e U, respectivamente. Cada E_i é um polinômio em T; logo, U comuta com cada E_i. Como cada F_j é um polinômio em U vemos que E_i e F_j comutam. Seja agora

$$G_{ij} = E_iF_j, \quad 1 \leq i \leq k, \quad 1 \leq j \leq r.$$

Então, G_{ij} é uma projeção ortogonal. De fato, como $E_iF_j = F_jE_i$, temos

$$G_{ij}^2 = E_iF_jE_iF_j = E_iE_iF_jF_j = E_iF_j = G_{ij}$$
$$G_{ij}^* = (E_iF_j)^* = F^*_jE^*_i = F_jE_i = E_iF_j = G_{ij}.$$

Além disso
$$\sum_{i=1}^{k} \sum_{j=1}^{r} G_{ij} = \sum_i \sum_j E_i F_j$$
$$= \left(\sum_{i=1}^{k} E_i \right) \left(\sum_{j=1}^{r} F_j \right)$$
$$= I \; . \; I$$
$$= I.$$

Se $i \neq p$ ou $j \neq q$, então
$$\begin{aligned} G_{ij}G_{pq} &= E_iE_jE_pE_q \\ &= E_iE_pF_jF_q \\ &= 0. \end{aligned}$$

Portanto, temos aqui kr projeções ortogonais G_{ij} tais que $G_{ij}G_{pq} = 0$ a menos que $i = p$ e $j = q$ (e talvez mesmo nesse caso). Ora, T e U são combinações lineares das G_{ij}:
$$T = TI = (\sum_i c_i E_i)(\sum_j F_j) = \sum_i \sum_j c_i G_{ij}$$
$$U = UI = (\sum_j d_j F_j)(\sum_i E_i) = \sum_i \sum_j d_j G_{ij}.$$

Sejam agora e_{ij}, $1 \leq i \leq k$, $1 \leq j \leq r$, kr números complexos arbitrários *distintos* e seja
$$S = \sum_i \sum_j e_{ij} G_{ij}.$$

As condições que estabelecemos sôbre as G_{ij} nos mostra que S é normal e que
$$f(S) = \sum_i \sum_j f(e_{ij}) G_{ij}$$
para todo polinômio f. Tomemos agora um polinômio f tal que $f(e_{ij}) = c_i$ e um polinômio g tal que $g(e_{ij}) = d_j$ e temos $T = f(S)$ e $U = g(S)$.

Todo operador linear que comuta com T e com U comuta com E_i e com F_j, logo comuta com cada G_{ij}, portanto comuta com S. Poderíamos destacar que S é um polinômio em T e U conjuntamente, isto é, S tem a forma
$$S = \sum_{m,n=0}^{N} a_{mn} T^m U^n.$$

Isto é evidente, pois cada G_{ij} é o produto de um polinômio em T por um polinômio em U.

Corolário. *Sejam* $T_1, \ldots, T_n$ *um número finito de operadores normais sôbre* V *tais que* T_i *comuta com* T_j *para todos* i, j. *Então existe um operador normal* T *sôbre* V *tal que cada* T_j *seja um polinômio em* T.

Demonstração. A demonstração será por indução sôbre n, usando o Teorema 29. Se $n = 1$, basta tomar $T = T_1$. Suponhamos que $n > 1$ e que o teorema seja válido para quaisquer $n - 1$ operadores normais que comutam. Pelo Teorema 29, existe um operador normal S sôbre V tal que T_1 e T_2 sejam polinômios em S e S comuta com todo operador que comute com T_1 e com T_2. Ora, $\{S, T_3, \ldots, T_n\}$ é uma coleção de $n - 1$ operadores normais sôbre V que comutam. Pela hipótese de indução, existe um operador normal T tal que $S, T_3, \ldots, T_n$ sejam polinômios em T. Como T_1 e T_2 são polinômios em S, vemos que cada T_j é um polinômio em T.

Teorema 30. *Seja* $\mathfrak{F}$ *uma família arbitrária de operadores normais sôbre* V, *todos os quais comutam. Então existe um operador normal* T *sôbre* V *tal que todo operador na família* $\mathfrak{F}$ *seja um polinômio em* T.

Demonstração. Seja W o subespaço de $L(V, V)$, gerado por $\mathfrak{F}$, isto é, o conjunto das combinações lineares (finitas) de operadores em $\mathfrak{F}$. Então, todo operador em W é normal. De fato, sejam $U_1, \ldots, U_k$ operadores em $\mathfrak{F}$ e $c_1, \ldots, c_k$ números complexos. Pelo último corolário, existe um operador normal S tal que cada U_j seja um polinômio em S. Portanto, a combinação linear $c_1U_1 + \ldots + c_kU_k$ é um polinômio no operador normal S e é também normal. Além disso, é evidente que todos os operadores em W comutam.

Seja agora $\{T_1, \ldots, T_n\}$ uma base de W. Pelo último corolário, existe um operador normal T tal que cada T_j seja um polinômio em T. Como todo operador em W é uma combinação linear de $T_1, \ldots, T_n$, todo operador em W (em particular, em $\mathfrak{F}$) é um polinômio em T.

Corolário. *Se* $\mathfrak{F}$ *é uma família arbitrária de operadores normais sôbre* V *que comutam, existe um base ortonormal ordenada* $\mathcal{B}$ *de* V *tal que todo operador em* $\mathfrak{F}$ *seja representado por uma matriz diagonal em relação à base ordenada* $\mathcal{B}$.

Demonstração. Tomemos T como no Teorema 30 e uma base ortonormal $\mathcal{B}$ tal que $[T]_{\mathcal{B}}$ seja diagonal. Se D é uma matriz diagonal e f é um polinômio, então $f(D)$ é diagonal.

CAPÍTULO 9

FORMAS BILINEARES

9.1 Formas Bilineares

Neste capítulo vamos tratar das formas bilineares sôbre espaços vetoriais de dimensão finita. O leitor provàvelmente observará uma semelhança entre uma parte da matéria e a discussão dos determinantes no Capítulo 5 e dos produtos internos no Capítulo 8. A relação entre formas bilineares e produtos internos é particularmente forte; no entanto, êste capítulo não pressupõe nada da matéria do Capítulo 8. O leitor que não tiver familiaridade com produtos internos provàvelmente lucraria lendo a primeira parte do Capítulo 8 à medida que lesse a discussão de formas bilineares.

A primeira seção trata do espaço das formas bilineares sôbre um espaço vetorial de dimensão n. A matriz de uma forma bilinear em relação a uma base ordenada é introduzida e é estabelecido o isomorfismo entre o espaço das formas e o espaço das $n \times n$ matrizes. Define-se o pôsto de uma forma bilinear e são introduzidas as formas bilineares degeneradas. A segunda seção discute as formas bilineares simétricas e sua diagonalização. A terceira seção estuda as formas bilineares anti-simétricas. A quarta seção discute o grupo que conserva uma forma bilinear não-degenerada, com atenção especial prestada aos grupos ortogonais, os grupos pseudo-ortogonais e um grupo pseudo-ortogonal particular — o grupo de Lorentz.

Definição. *Seja* V *um espaço vetorial sôbre o corpo* F. *Uma* **forma bilinear** *sôbre* V *é uma função* f, *que associa a cada par ordenado de vetores* α, β *em* V *um escalar* $f(\alpha, \beta)$ *em* F, *e que satisfaz*

$$(9\text{-}1)\qquad \begin{aligned} f(c\alpha_1 + \alpha_2, \beta) &= cf(\alpha_1, \beta) + f(\alpha_2, \beta) \\ f(\alpha, c\beta_1 + \beta_2) &= cf(\alpha, \beta_1) + f(\alpha, \beta_2). \end{aligned}$$

Indicando por $V \times V$ o conjunto de todos os pares ordenados de vetores em V, esta definição pode ser reformulada como segue: Uma forma bilinear sôbre V é uma função f de $V \times V$ em F que é linear como uma função de qualquer dos seus argumentos quando o outro é deixado fixo. A função nula de $V \times V$ em F é evidentemente uma forma bilinear. É também verdade que tôda combinação linear de formas bilineares sôbre V é ainda uma forma bilinear. Para demonstrar êste fato, basta considerar combinações lineares do tipo $cf + g$, sendo f e g formas bilineares sôbre V. A demonstração de que $cf + g$ satisfaz (9-1) é semelhante a muitas outras que fizemos e vamos omiti-la. Tudo isto pode ser resumido dizendo-se que o conjunto das formas bilineares sôbre V é um subespaço do espaço das funções de $V \times V$ em F (Exemplo 3, capítulo 2). Indicaremos o espaço das formas bilineares sôbre V por $L(V, V, F)$.

Exemplo 1. Seja V um espaço vetorial sôbre o corpo F e sejam L_1 e L_2 funcionais lineares sôbre V. Definamos f por

$$f(\alpha, \beta) = L_1(\alpha)L_2(\beta).$$

Fixando β e considerando f como uma função de α, então temos simplesmente um múltiplo escalar do funcional linear L_1. Com α fixo, f é um múltiplo escalar de L_2. Assim, é evidente que f é uma forma bilinear sôbre V.

Exemplo 2. Sejam m e n inteiros positivos e F um corpo. Seja V o espaço vetorial das $m \times n$ matrizes sôbre F. Seja A uma $m \times m$ matriz fixa sôbre F. Definamos

$$f_A(X, Y) = \text{traço } (X^tAY).$$

Então f_A é uma forma bilinear sôbre V. De fato, se X, Y e Z são $m \times n$ matrizes sôbre F.

$$\begin{aligned} f_A(cX + Z, Y) &= \text{traço } [(cX + Z)^tAY] \\ &= \text{traço } (cX^tAY) + \text{traço } (Z^tAY) \\ &= cf_A(X, Y) + f_A(Z, Y). \end{aligned}$$

Evidentemente, utilizamos o fato de que a operação transposta e a função traço são lineares. É ainda mais fácil mostrar que f_A é linear como uma função do seu segundo argumento. No caso particular $n = 1$, a matriz X^tAY é 1×1, isto é, um escalar, e a forma bilinear é simplesmente

$$\begin{aligned} f_A(X, Y) &= X^tAY \\ &= \sum_i \sum_j A_{ij}x_iy_j. \end{aligned}$$

Mostraremos em breve que tôda forma bilinear sôbre o espaço das $m \times 1$ matrizes é dêsse tipo, isto é, é f_A para alguma $m \times m$ matriz A.

Exemplo 3. Seja F um corpo. Vamos determinar tôdas as formas bilineares sôbre o espaço F^2. Suponhamos que f seja uma tal forma bilinear. Se $\alpha = (x_1, x_2)$ e $\beta = (y_1, \alpha_2)$ são vetores em F^2, então

$$\begin{aligned} f(\alpha, \beta) &= f(x_1\epsilon_1 + x_2\epsilon_2, \beta) \\ &= x_1 f(\epsilon_1, \beta) + x_2 f(\epsilon_2, \beta) \\ &= x_1 f(\epsilon_1, y_1\epsilon_1 + y_2\epsilon_2) + x_2 f(\epsilon_2, y_1\epsilon_1 + y_2\epsilon_2) \\ &= x_1y_1 f(\epsilon, \epsilon_1) + x_1y_2 f(\epsilon_1, \epsilon_2) + x_2y_1 f(\epsilon_2, \epsilon_1) + x_2y_2 f(\epsilon_2, \epsilon_2). \end{aligned}$$

Assim, f é completamente determinada pelos quatro escalares $A_{ij} = f(\epsilon_i, \epsilon_j)$ da seguinte maneira:

$$\begin{aligned} f(\alpha, \beta) &= A_{11}x_1y_1 + A_{12}x_1y_2 + A_{21}x_2y_1 + A_{22}x_2y_2 \\ &= \sum_{i,j} A_{ij}x_iy_j. \end{aligned}$$

Se X e Y são as matrizes das coordenadas de α e β e se A é a 2×2 matriz com elementos $A(i,j) = A_{ij} = f(\epsilon_i, \epsilon_j)$, então

$$f(\alpha, \beta) = X^tAY. \tag{9-2}$$

Observemos no Exemplo 2 que se A é uma 2×2 matriz arbitrária sôbre F, então (9-2) define uma forma bilinear sôbre F^2. Vemos que as formas bilineares sôbre F^2 são exatamente as obtidas por meio de uma 2×2 matriz como em (9-2).

A discussão no Exemplo 3 pode ser generalizada de modo a descrever tôdas as formas bilineares sôbre um espaço vetorial de dimensão finita. Seja V um espaço vetorial de dimensão finita sôbre o corpo F e seja $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ uma base ordenada de V. Suponhamos que f seja uma forma bilinear sôbre V. Se

$$\alpha = x_1\alpha_1 + \ldots + x_n\alpha_n \quad \text{e} \quad \beta = y_1\alpha_1 + \ldots + y_n\alpha_n$$

são vetores em V, então

$$\begin{aligned} f(\alpha, \beta) &= f\left(\sum_i x_i\alpha_i, \beta\right) \\ &= \sum_i x_i f(\alpha_i, \beta) \\ &= \sum_i x_i f\left(\alpha_i, \sum_j y_j\alpha_j\right) \\ &= \sum_i \sum_j x_iy_j f(\alpha_i, \alpha_j). \end{aligned}$$

Se fizermos $A_{ij} = f(\alpha_i, \alpha_j)$, então

$$f(\alpha, \beta) = \sum_i \sum_j A_{ij} x_i y_j$$
$$= X^t A Y$$

onde X e Y são as matrizes das coordenadas de α e β em relação à base ordenada $\mathcal{B}$. Assim, tôda forma bilinear sôbre V é do tipo

$$f(\alpha, \beta) = [\alpha]_{\mathcal{B}}^t A[\beta]_{\mathcal{B}} \tag{9-3}$$

para alguma $n \times n$ matriz A sôbre F. Recìprocamente, se temos uma $n \times n$ matriz arbitrária A, é fácil ver que (9-3) define uma forma bilinear sôbre V, tal que $A_{ij} = f(\alpha_i, \alpha_j)$.

Definição. *Seja* V *um espaço vetorial de dimensão finita e seja* $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ *uma base ordenada de* V. *Se* f *é uma forma bilinear sôbre* V, *a* **matriz de** f **em relação à base ordenada** $\mathcal{B}$ *e a* $n \times n$ *matriz* A *com elementos* $A_{ij} = f(\alpha_i, \alpha_j)$. *Às vêzes, indicaremos esta matriz por* $[f]_{\mathcal{B}}$.

Teorema 1. *Seja* V *um espaço vetorial de dimensão finita sôbre o corpo* F. *Para cada base ordenada* $\mathcal{B}$ *de* V, *a função que associa a cada forma bilinear sôbre* V *sua matriz em relação à base ordenada* $\mathcal{B}$ *é um isomorfismo do espaço* L(V, V, F) *no espaço das* $n \times n$ *matrizes sôbre o corpo* F.

Demonstração. Observamos acima que $f \to [f]_{\mathcal{B}}$ é uma correspondência bijetora entre o conjunto das formas bilineares sôbre V e o conjunto de tôdas as $n \times n$ matrizes sôbre F. Que isso é uma transformação linear é fácil de ver, pois

$$(cf + g)(\alpha_i, \alpha_j) = cf(\alpha_i, \alpha_j) + g(\alpha_i, \alpha_j)$$

para todos i e j. Isto diz simplesmente que

$$[cf + g]_{\mathcal{B}} = c[f]_{\mathcal{B}} + [g]_{\mathcal{B}}.$$

Corolário. *Se* $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ *é uma base ordenada de* V *e* $\mathcal{Q}^* = \{L_1, \ldots, L_n\}$ *é a base dual de* V^*, *então as* n^2 *formas bilineares*

$$f_{ij}(\alpha, \beta) = L_i(\alpha)L_j(\beta), \quad 1 \leq i \leq n, \quad 1 \leq j \leq n$$

formam uma base do espaço L(V, V, F). *Em particular, a dimensão de* L(V, V, F) *é* n^2.

Demonstração. A base dual $\{L_1, \ldots, L_n\}$ é definida essencialmente pelo fato de que $L_i(\alpha)$ é a i-ésima coordenada de α em relação à base ordenada $\mathfrak{B}$ (para todo α em V). Ora, as funções f_{ij} definidas por

$$f_{ij}(\alpha, \beta) = L_i(\alpha)L_j(\beta)$$

são formas bilineares do tipo considerado no Exemplo 1. Se

$$\alpha = x_1\alpha_1 + \dots + x_n\alpha_n \text{ e } \beta = y_1\alpha_1 + \dots + y_n\alpha_n,$$

então

$$f_{ij}(\alpha, \beta) = x_i y_j.$$

Seja f uma forma bilinear arbitrária sôbre V e seja A a matriz de f em relação à base ordenada $\mathfrak{B}$. Então

$$f(\alpha, \beta) = \sum_{i,j} A_{ij} x_i y_j$$

o que diz simplesmente que

$$f = \sum_{i,j} A_{ij} f_{ij}.$$

Agora é evidente que as n^2 formas f_{ij} formam uma base de $L(V, V, F)$.

Pode-se reformular a demonstração do corolário como segue. A matriz da forma bilinear f_{ij} em relação à base ordenada $\mathfrak{B}$ é a matriz "unitária" $E^{i,j}$, cujo único elemento não-nulo é um 1 na linha i e coluna j. Como estas matrizes $E^{i,j}$ constituem uma base do espaço das $n \times n$ matrizes, as formas f_{ij} constituem uma base do espaço das formas bilineares.

O conceito de matriz de uma forma bilinear em relação a uma base ordenada é semelhante ao conceito de matriz de um operador linear em relação a uma base ordenada. Do mesmo modo que para operadores lineares, estaremos interessados no que acontece à matriz que representa uma forma bilinear, ao passarmos de uma base ordenada a outra. Então, suponhamos que $\mathfrak{B} = \{\alpha_1, \dots, \alpha_n\}$ e $\mathfrak{B}' = \{\alpha_1', \dots, \alpha_n'\}$ sejam duas bases ordenadas de V e que f seja uma forma bilinear sôbre V. Como se relacionam as matrizes $[f]_{\mathfrak{B}}$ e $[f]_{\mathfrak{B}'}$? Bem, seja P a $n \times n$ matriz (inversível) tal que

$$[\alpha]_{\mathfrak{B}} = P[\alpha]_{\mathfrak{B}'}$$

para todo α em V. Em outras palavras, definimos P por

$$\alpha_j' \sum_{i=1}^{n} P_{ij}\alpha_i,$$

Para vetores arbitrários α, β em V

$$\begin{aligned} f(\alpha, \beta) &= [\alpha]^t_{\mathfrak{B}}[f]_{\mathfrak{B}}[\beta]_{\mathfrak{B}} \\ &= (P[\alpha]_{\mathfrak{B}'})^t[f]_{\mathfrak{B}}P[\beta]_{\mathfrak{B}}, \\ &= [\alpha]^t_{\mathfrak{B}'}(P^t[f]_{\mathfrak{B}}P)[\beta]_{\mathfrak{B}'}. \end{aligned}$$

Pela definição e unicidade da matriz que representa f em relação à base ordenada $\mathcal{B}'$, devemos ter

(9-4) $$[f]_{\mathcal{B}'} = P^t[f]_{\mathcal{B}}P.$$

Exemplo 4. Seja V o espaço vetorial R^2. Seja f a forma bilinear definida sôbre $\alpha = (x_1, x_2)$ e $\beta = (y_1, y_2)$ por

$$f(\alpha, \beta) = x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2.$$

Ora,

$$f(\alpha, \beta) = [x_1, x_2]\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

e então a matriz de f em relação à base ordenada canônica $\mathcal{B} = \{\epsilon_1, \epsilon_2\}$ é

$$[f]_{\mathcal{B}} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Seja $\mathcal{B}' = \{\epsilon_1', \epsilon_2'\}$ a base ordenada definida por $\epsilon_1' = (1, -1,)$ $\epsilon_2' = (1, 1)$. Neste caso, a matriz P que transforma as coordenadas de $\mathcal{B}'$ para $\mathcal{B}$ é

$$P = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

Assim

$$\begin{aligned} [f]_{\mathcal{B}'} &= [P^t f]_{\mathcal{B}}P \\ &= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}. \end{aligned}$$

O que isto significa é que, se exprimirmos os vetores α e β por meio de suas coordenadas em relação à base $\mathcal{B}'$, digamos

$$\alpha = x_1'\epsilon_1' + x_2'\epsilon_2', \quad \beta = y_1'\epsilon_1' + y_2'\epsilon_2'$$

então

$$f(\alpha, \beta) = 4x_2'y_2'.$$

Uma conseqüência da fórmula (9-4) da mudança de base é a seguinte: Se A e B são $n \times n$ matrizes que representam a mesma forma bilinear sôbre V em relação a bases ordenadas (possìvelmente) diferentes, então A e B têm o mesmo pôsto. De fato, se P é uma $n \times n$ matriz inversível e $B = P^tAP$, é evidente que A e B têm o mesmo pôsto. Isto torna possível definir o pôsto de uma forma bilinear sôbre V, como sendo o pôsto de qualquer matriz que represente a forma em relação a uma base ordenada de V.

É desejável dar uma definição mais intrínseca do pôsto de uma forma bilinear. Isto pode ser feito como segue: Suponhamos que f seja uma forma bilinear sôbre o espaço vetorial V. Fixando um vetor α em V, $f(\alpha, \beta)$ é linear com uma função de β. Desta maneira, cada α fixo determina um funcional linear sôbre V; indique êste funcional linear por $L_f(\alpha)$. Repetindo, se α é um vetor em V, então $L_f(\alpha)$ é o funcional linear sôbre V cujo valor em qualquer vetor β é $f(\alpha, \beta)$. Isto nos dá uma transformação $\alpha \rightarrow L_f(\alpha)$ de V no espaço dual V^*. Como

$$f(c\alpha_1 + \alpha_2, \beta) = cf(\alpha_1, \beta) + f(\alpha_2, \beta)$$

vemos que

$$L_f(c\alpha_1 + \alpha_2) = cL_f(\alpha_1) + L_f(\alpha_2)$$

isto é, L_f é uma transformação linear de V em V^*.

De maneira semelhante, f determina uma transformação linear R_f de V em V^*. Para cada β fixo em V, $f(\alpha, \beta)$ é linear como uma função de α. Definimos $R_f(\beta)$ como sendo o funcional linear sôbre V cujo valor no vetor α é $f(\alpha, \beta)$.

Teorema 2. *Seja* f *uma forma bilinear sôbre o espaço vetorial* V *de dimensão finita. Sejam* L_f *e* R_f *as transformações lineares de* V *em* V* *definidas por* $(L_f\alpha)(\beta) = f(\alpha, \beta) = (R_f\beta)(\alpha)$. *Então pôsto* (L_f) = *pôsto* (R_f).

Demonstração. Pode-se fazer uma demonstração dêste teorema que seja "independente de coordenadas". Tal demonstração é semelhante à demonstração (na Seção 3.7) de que o pôsto-linha de uma matriz é igual ao seu pôsto-coluna. Então, faremos aqui uma demonstração que começa tomando um sistema de coordenadas (base), utilizando depois o teorema "pôsto-linha igual a pôsto-coluna".

Para demonstrar que pôsto (L_f) = pôsto (R_f) bastará demonstrar que L_f e R_f têm a mesma nulidade. Seja $\mathcal{B}$ uma base ordenada de V e seja $A = [f]_{\mathcal{B}}$. Se α e β são vetores em V, com matrizes de coordenadas X e Y em relação à base ordenada $\mathcal{B}$, então $f(\alpha, \beta) = X^tAY$. Ora, $R_f(\beta) = 0$ significa que $f(\alpha, \beta) = 0$ para todo α em

V, isto é, que $X^tAY = 0$ para tôda $n \times 1$ matriz X. A última condição diz simplesmente que $AY = 0$. A nulidade de R_f é portanto igual à dimensão do espaço das soluções de $AY = 0$.

Anàlogamente, $L_f(\alpha) = 0$ se, e sòmente se, $X^tAY = 0$ para tôda $n \times 1$ matriz Y. Assim, α está no núcleo de L_f se e sòmente se $X^tA = 0$, isto é, $A^tX = 0$. A nulidade de L_f é portanto igual à dimensão do espaço das soluções de $A^tA = 0$. Como as matrizes A e A^t têm o mesmo pôsto-coluna, vemos que

$$\text{nulidade } (L_f) = \text{nulidade } (R_f).$$

Definição. *Se* f *é uma forma bilinear sôbre o espaço* V *de dimensão finita, o* **pôsto** *de* f *é o inteiro* r = *pôsto* (L_f) = *pôsto* (R_f).

Corolário 1. *O pôsto de uma forma bilinear é igual ao pôsto da matriz da forma bilinear em relação a qualquer base ordenada.*

Corolário 2. *Se* f *é uma forma bilinear sôbre o espaço vetorial* n-*dimensional* V, *as seguintes afirmações são equivalentes:*

(i) *pôsto* (f) = n.

(ii) *Para cada* α *não-nulo em* V, *existe um* β *em* V *tal que* $f(\alpha, \beta) \neq 0$.

(iii) *Para cada* β *não-nulo em* V, *existe um* α *em* V *tal que* $f(\alpha, \beta) \neq 0$.

Demonstração. A afirmação (ii) diz simplesmente que o núcleo de L_f é o subespaço nulo. A afirmação (iii) diz que o núcleo de R_f é o subespaço nulo. As transformações lineares L_f e R_f têm nulidade 0 se, e sòmente se, elas têm o mesmo pôsto n, isto é, se, e sòmente se, pôsto $(f) = n$.

Definição. *Uma forma bilinear* f *sôbre um espaço vetorial* V *é dita* **não-degenerada** (*ou* **não-singular**) *se satisfaz as condições (ii) e (iii) do Corolário* 2.

Se V é de dimensão finita, então f é não-degenerada, desde que f satisfaça qualquer uma das três condições do Corolário 2. Em particular, f é não-degenerada (não-singular) se, e sòmente se, sua matriz em relação a alguma (tôda) base ordenada de V é uma matriz não-singular.

Exemplo 5. Seja $V = R^n$ e seja f a forma bilinear definida sôbre $\alpha = (x_1, \ldots, x_n)$ e $\beta = (y_1, \ldots, y_n)$ por

$$f(\alpha, \beta) = x_1y_1 + \ldots + x_ny_n.$$

Então f é uma forma bilinear não-degenerada sôbre R^n. A matriz de f em relação à base ordenada canônica é a $n \times n$ matriz unidade:

$$f(X, Y) = X^tY.$$

Esta f é usualmente denominada o produto escalar. O leitor provàvelmente tem familiaridade com esta forma bilinear, pelo menos no caso $n = 3$. Geomètricamente, o número $f(\alpha, \beta)$ é o produto do comprimento de α pelo comprimento de β e pelo cosseno do ângulo entre α e β. Em particular, $f(\alpha, \beta) = 0$ se, e sòmente se, os vetores α e β são ortogonais (perpendiculares).

Exercícios

1. Quais das seguintes funções f, definidas sôbre vetores $\alpha = (x_1, x_2)$ e $\beta = (y_1, y_2)$ em R^2, são formas bilineares?

(a) $f(\alpha, \beta) = 1$;
(b) $f(\alpha, \beta) = (x_1 - y_1)^2 + x_2y_2$;
(c) $f(\alpha, \beta) = (x_1 + y_1)^2 - (x_1 - y_1)^2$;
(d) $f(\alpha, \beta) = x_1y_2 - x_2y_1$.

2. Seja f a forma bilinear sôbre R^2 definida por

$$f((x_1, y_1), (x_2, y_2)) = x_1y_1 + x_2y_2.$$

Determinar a matriz de f em relação a cada uma das seguintes bases:

$$\{(1, 0), (0, 1)\}, \quad \{1, -1), (1, 1)\}, \quad \{1, 2), (3, 4)\}.$$

3. Seja V o espaço das 2×3 matrizes sôbre R e seja f a forma bilinear sôbre V definida por $f(X, y) = \text{traço } (X^tAY)$, onde

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

Determinar a matriz de f em relação à base ordenada

$$\{E^{11}, E^{12}, E^{13}, E^{21}, E^{22}, E^{23}\}$$

onde E^{ij} é a matriz cujo único elemento não-nulo é um 1 na linha i e coluna j.

4. Descrever explìcitamente tôdas as formas bilineares f sôbre R^3 com a propriedade de que $f(\alpha, \beta) = f(\beta, \alpha)$ para todos α, β.

5. Descrever as formas bilineares sôbre R^3 que satisfazem $f(\alpha, \beta) = -f(\beta, \alpha)$ para todos α, β.

6. Seja n um inteiro positivo e seja V o espaço das $n \times n$ matrizes sôbre o corpo dos números complexos. Mostrar que a equação

$$f(A, B) = n \text{ tr } (AB) - \text{tr } (A) \text{ tr } (B)$$

define uma forma bilinear f sôbre B. É verdade que $f(A, B) = f(B, A)$ para tôdas A, B?

7. Seja f a forma bilinear do Exercício 6. Mostrar que f é degenerada (f não é não-degenerada). Seja V_1 o subespaço de V formado pelas matrizes de traço 0 e seja f_1 a restrição de f a V_1. Mostrar que f_1 é não degenerada.

8. Seja f a forma bilinear definida no Exercício 6 e seja V_2 o subespaço de V formado por tôdas as matrizes A tais que traço $(A) = 0$ e $A^* = -A$.

(A^* é a transposta conjugada de A.) Indiquemos por f_2 a restrição de f a V_2. Mostrar que f_2 é negativa definida, isto é, que $f_2(A, A) < 0$ para tôda A não-nula em V_2.

9. Seja f a forma bilinear definida no Exercício 6. Seja W o conjunto das matrizes A em V tais que $f(A, B) = 0$ para tôda B. Mostrar que W é um subespaço de V. Descrever W explìcitamente e determinar sua dimensão.

10. Seja f uma forma bilinear sôbre um espaço vetorial V de dimensão finita. Seja W o subespaço formado pelos β tais que $f(\alpha, \beta) = 0$ para todo α. Mostrar que

$$\text{pôsto}\,(f) = \dim V - \dim W.$$

Usar êste resultado e o resultado do Exercício 9 para calcular o pôsto da forma bilinear definida no Exercício 6.

11. Seja f uma forma bilinear sôbre um espaço vetorial V de dimensão finita. Suponhamos que V_1 seja um subespaço de V com a propriedade de que a restrição de f a V_1 seja não-degenerada. Mostrar que pôsto $(f) \geq \dim V_1$.

12. Sejam f e g formas bilineares sôbre um espaço vetorial V de dimensão finita. Suponhamos que g seja não-singular. Mostrar que existem operadores lineares T_1, T_2 sôbre V, únicos, tais que

$$f(\alpha, \beta) = g(T_1\alpha, \beta) \qquad g(\alpha, T_2\beta)$$

para todos α, β.

13. Mostrar que o resultado do Exercício 12 não é necessàriamente válido se g é singular.

14. Seja f uma forma bilinear sôbre um espaço vetorial V de dimensão finita. Mostrar que f pode ser expresso como um produto de dois funcionais lineares (isto é, $f(\alpha, \beta) = L_1(\alpha)L_2(\beta)$ para L_1, L_2 em V^*) se, e sòmente se, f tem pôsto 1.

9.2 Formas Bilineares Simétricas

O propósito principal desta seção é responder à seguinte pergunta: Se f é uma forma bilinear sôbre o espaço vetorial V de dimensão finita, quando é que existe uma base ordenada $\mathcal{B}$ de V, em relação à qual f é representada por uma matriz diagonal? Vamos demonstrar que isto é possível se, e sòmente se, f é uma forma bilinear simétrica, ou seja, $f(\alpha, \beta) = f(\beta, \alpha)$. O teorema será demonstrado apenas para o caso do corpo de escalares ser um subcorpo dos números complexos. O leitor informado verá que basta o corpo F de escalares ter característica zero, isto é, que, se n é um inteiro positivo, a soma $1 + \ldots + 1$ (n vêzes) em F não é 0.

Definição. *Seja* f *uma forma bilinear sôbre o espaço vetorial* V. *Dizemos que* f *é* **simétrica** *se* $f(\alpha, \beta) = f(\beta, \alpha)$ *para todos os vetores* α, β *em* V.

Se V é de dimensão finita, a forma bilinear f é simétrica se, e sòmente se, sua matriz A em relação a alguma ou (tôda) base ordenada é simétrica, isto é, $A^t = A$. Para ver isto, perguntamos quando é que a forma bilinear

$$f(X, Y) = X^tAY$$

é simétrica. Isto acontece se, e sòmente se, $X^tAY = Y^tAX$ para tôdas matrizes-colunas X e Y. Como X^tAY é uma 1×1 matriz, temos $X^tAY = Y^tA^tX$. Assim, f é simétrica se, e sòmente se, $Y^tA^tX = Y^tAX$ para tôdas X, Y. Evidentemente, isto significa apenas que $A = A^t$. Em particular, deve-se notar que se existir uma base ordenada de V em relação à qual f seja representada por uma matriz diagonal, então f é simétrica, pois qualquer matriz diagonal é uma matriz simétrica.

Se f é uma forma bilinear simétrica, a **forma quadrática associada a** f é a função q de V em F definida por

$$q(\alpha) = f(\alpha, \alpha).$$

Se F é um subcorpo do corpo dos números complexos, a forma bilinear simétrica f é completamente determinada por sua forma quadrática associada, de acôrdo com a **identidade de polarização**.

$$(9\text{-}5) \qquad f(\alpha, \beta) = \tfrac{1}{4}q(\alpha + \beta) - \tfrac{1}{4}q(\alpha - \beta).$$

A demonstração de (9-5) requer apenas cálculos de rotina, que omitiremos. Se f é a forma bilinear do Exemplo 5, ou seja, o produto escalar, então a forma quadrática associada é

$$q(x_1, \ldots, x_n) = x_1^2 + \ldots + x_n^2.$$

Em outras palavras, $q(\alpha)$ é o quadrado do comprimento de α. Para a forma bilinear $f_A(X, Y) = X^tAY$, a forma quadrática associada é

$$q_A(X) = X^tAX = \sum_{i,j} A_{ij}x_ix_j.$$

Uma classe importante de formas bilineares simétricas consiste dos produtos internos sôbre espaços vetoriais reais, discutidos no Capítulo 8. Se V é um espaço vetorial *real*, um **produto interno** sôbre V é uma forma bilinear simétrica f sôbre V que satisfaz

$$(9\text{-}6) \qquad f(\alpha, \alpha) > 0 \quad \text{se} \quad \alpha \neq 0.$$

Uma forma bilinear que satisfaz (9-6) é dita **positiva definida**. Assim, um produto interno sôbre um espaço vetorial real é uma forma bilinear simétrica positiva definida sôbre aquêle espaço. Notemos que um produto interno é não-degenerado. Dois vetores α, β são ditos **ortogonais** em relação ao produto interno f se $f(\alpha, \beta) = 0$. A forma quadrática $q(\alpha) = f(\alpha, \alpha)$ toma apenas valores não-negativos e $q(\alpha)$ é usualmente considerado como sendo o quadrado do comprimento de α. Evidentemente, êstes conceitos de comprimento e ortogonalidade se originam do exemplo mais importante de produto interno — o produto escalar do Exemplo 5.

Se f é uma forma bilinear simétrica sôbre um espaço vetorial V, é conveniente aplicar um pouco da terminologia de produtos internos a f. É particularmente conveniente dizer que α e β são ortogonais em relação a f se $f(\alpha, \beta) = 0$. Não é aconselhável considerar $f(\alpha, \alpha)$ como sendo o quadrado do comprimento de α; por exemplo, se V é um espaço vetorial complexo, podemos ter $f(\alpha, \alpha) = \sqrt{-1}$, ou num espaço vetorial real $f(\alpha, \alpha) = -2$.

Passamos agora ao teorema fundamental desta seção. Ao ler a demonstração, o leitor deverá achar útil pensar no caso particular em que V é um espaço vetorial real e f é um produto interno sôbre V.

Teorema 3. *Seja* V *um espaço vetorial de dimensão finita sôbre um subcorpo do corpo dos números complexos e seja* f *uma forma bilinear simétrica sôbre* V. *Então existe uma base ordenada de* V *em relação à qual* f *é representada por uma matriz diagonal.*

Demonstração. O que precisamos encontrar é uma base ordenada

$$\mathfrak{B} = \{\alpha_1, \ldots, \alpha_n\}$$

tal que $f(\alpha_i, \alpha_j) = 0$ para $i \neq j$. Se $f = 0$ ou $n = 1$, o teorema é òbviamente verdadeiro. Assim, podemos supor $f \neq 0$ e $n > 1$. Se $f(\alpha, \alpha) = 0$ para todo α em V, a forma quadrática associada q é idênticamente 0 e a identidade de polarização (9-5) mostra que $f = 0$. Assim, existe um vetor α em V tal que $f(\alpha, \alpha) = q(\alpha) \neq 0$. Seja W o subespaço unidimensional de V que é gerado por α e seja $W^\perp$ o conjunto dos vetores β em V tais que $f(\alpha, \beta) = 0$. Afirmamos agora que $V = W \oplus W^\perp$. Certamente os subespaços W e $W^\perp$ são disjuntos. Um vetor típico em W é $c\alpha$, onde c é um escalar. Se $c\alpha$ está também em $W^\perp$, então $f(c\alpha, c\alpha) = c^2 f(\alpha, \alpha) = 0$. Mas $f(\alpha, \alpha) \neq 0$, logo $c = 0$. Além disso, todo vetor em V é a soma de um vetor em W e um vetor em $W^\perp$. De fato, seja γ um vetor arbitrário em V e coloquemos

$$\beta = \gamma - \frac{f(\gamma, \alpha)}{f(\alpha, \alpha)}\alpha.$$

Então

$$f(\alpha, \beta) = f(\alpha, \gamma) = \frac{f(\gamma, \alpha)}{f(\alpha, \alpha)} f(\alpha, \alpha)$$

e como f é simétrica, $f(\alpha, \beta) = 0$. Portanto, β está no subespaço $W^{\perp}$. A expressão

$$\gamma = \frac{f(\gamma, \alpha)}{f(\alpha, \alpha)} \alpha + \beta$$

nos mostra que $V = W + W^{\perp}$.

A restrição de f a $W^{\perp}$ é uma forma bilinear simétrica sôbre $W^{\perp}$ Como $W_{\perp}$ tem dimensão $(n - 1)$, podemos supor, por indução, que $W^{\perp}$ possua uma base $\{\alpha_2, \ldots, \alpha_n\}$ tal que

$$f(\alpha_i, \alpha_j) = 0, \quad i \neq j \quad (i \geq 2, j \geq 2).$$

Colocando $\alpha_1 = \alpha$, obtemos uma base $\{\alpha_1, \ldots, \alpha_n\}$ de V tal que $(\alpha_i, \alpha_j) = 0$ para $i \neq j$.

f **Corolário.** *Seja* F *um subcorpo do corpo dos números complexos e seja* A *uma* n $\times$ n *matriz simétrica sôbre* F. *Então existe um* n $\times$ n *matriz inversível* P *sôbre tal que* P^tAP *seja diagonal.*

No caso de F ser o corpo dos números reais, a matriz inversível P neste corolário pode ser escolhida de modo a ser uma matriz *ortogonal*, isto é, $P^t = P^{-1}$. Em outras palavras, se A é uma $n \times n$ matriz simétrica real, existe uma matriz ortogonal real P tal que P^tAP seja diagonal; contudo, isto não é nada evidente a partir do que fizemos acima (ver Capítulo 8).

Teorema 4. *Seja* V *um espaço vetorial de dimensão finita sôbre o corpo dos números complexos. Seja* f *uma forma bilinear simétrica sôbre* V *que tenha pôsto* r. *Então existe uma base ordenada* $\mathfrak{B} = \{\beta_1, \ldots, \beta_n\}$ *de* V *tal que*

(i) *a matriz de* f *em relação à base ordenada* B *é diagonal*;

(ii) $f(\beta_j, \beta_j) = \begin{cases} 1, & j = 1, \ldots, r \\ 0, & j > r. \end{cases}$

Demonstração. Pelo Teorema 3, existe uma base ordenada $\{\alpha_1, \ldots, \alpha_n\}$ de V tal que

$$f(\alpha_i, \alpha_j) = 0 \quad \text{para } i \neq j.$$

Como f tem pôsto r, sua matriz em relação à base ordenada $\{\alpha_1, \ldots, \alpha_n\}$ também o tem. Assim, precisamos ter $f(\alpha_j, \alpha_j) \neq 0$

para exatamente r valores de j. Reordenando os vetores α_j, podemos admitir que

$$f(\alpha_j, \alpha_j) \neq 0, \quad j = 1, \ldots, r.$$

Usemos, agora o fato de que o corpo de escalares é o corpo dos números complexos. Se $\sqrt{f(\alpha_j, \alpha_j)}$ indica uma raiz quadrada complexa qualquer de $f(\alpha_j, \alpha_j)$ e se colocarmos

$$\beta_j = \begin{cases} \dfrac{1}{\sqrt{f(\alpha_j, \alpha_j)}}\alpha_j, & j = 1, \ldots, r \\ \alpha_j, & j > r \end{cases}$$

a base $\{\beta_1, \ldots, \beta_n\}$ satisfará as condições (i) e (ii).

Evidentemente, o Teorema 4 é válido se o corpo de escalares é um subcorpo qualquer do corpo dos números complexos, no qual todo elemento possua uma raiz quadrada. Não é válido, por exemplo, quando o corpo de escalares é o corpo dos números reais. Sôbre o corpo dos números reais, temos o seguinte substituto para o Teorema 4.

Teorema 5. *Seja* V *um espaço vetorial* n-*dimensional sôbre o corpo dos números reais e seja* f *uma forma bilinear simétrica sôbre* V *que tenha pôsto* r. *Então existe uma base ordenada* $\{\beta_1, \ldots, \beta_n\}$ *de* V *em relação à qual a matriz de* f *é diagonal e tal que*

$$f(\beta_j, \beta_j) = \pm 1, \quad j = 1, \ldots, r.$$

Além disso, o número de vetores β_j *da base para os quais* $f(\beta_j, \beta_j) = 1$ *é independente da escolha da base.*

Demonstração. Existe uma base $\{\alpha_1, \ldots, \alpha_n\}$ de V tal que

$$\begin{aligned} f(\alpha_i, \alpha_j) &= 0, & i &\neq j. \\ f(\alpha_j, \alpha_j) &\neq 0, & 1 &\leq j \leq r \\ f(\alpha_j, \alpha_j) &= 0, & j &> r. \end{aligned}$$

Seja

$$\begin{aligned} \beta_j &= |f(\alpha_j, \alpha_j)|^{-1/2}\alpha_j, & 1 &\leq j \leq r \\ \beta_j &= \alpha_j, & j &> r \end{aligned}$$

Então $\{\beta_1, \ldots, \beta_n\}$ é uma base que tem as propriedades enunciadas.

Seja p o númerode vetores β_j da base para os quais $f(\beta_j, \beta_j) = 1$; precisamos mostrar que o número p é independente da particular base que tomemos satisfazendo as condições acima. Seja V^+ o subespaço de V gerado pelo vetores β_j da base para os quais $f(\beta_j, \beta_j) = 1$, e seja V^- o subespaço gerado pelos vetores β_j da base tais que $f(\beta_j, \beta_j) = -1$. Ora, $p = \dim V^+$, logo o que precisamos demonstrar é a unicidade da dimensão de V^+. É fácil ver que se α é um vetor não-nulo em V^+, então $f(\alpha, \alpha) > 0$; em outras palavras, f é positiva definida sôbre o subespaço V^+. Anàlogamente, se α é um vetor não-nulo em V^-, então $f(\alpha, \alpha) < 0$, isto é, f é negativa definida sôbre o subespaço V^-. Seja agora $V^\perp$ o subespaço gerado pelos vetores β_j da base para os quais $f(\beta_j, \beta_j) = 0$. Se α está em $V^\perp$, então $f(\alpha, \beta) = 0$ para todo β em V.

Como $\{\beta_1, \ldots, \beta_n\}$ é uma base de V, temos

$$V = V^+ \oplus V^- \oplus V^\perp.$$

Além disso, afirmamos que se W é um subespaço arbitrário de V sôbre o qual f seja positiva definida, então os subespaços W, V^- e $V^\perp$ são independentes. De fato, suponhamos que α esteja em W, β em V^-, γ em $V^\perp$ e que $\alpha + \beta + \gamma = 0$. Então

$$0 = f(\alpha, \alpha + \beta + \gamma) = f(\alpha, \alpha) + f(\alpha, \beta) + f(\alpha, \gamma)$$
$$0 = f(\beta, \alpha + \beta + \gamma) = f(\beta, \alpha) + f(\beta, \beta) + f(\beta, \gamma).$$

Como γ está em $V^\perp$, $f(\alpha, \gamma) = f(\beta, \gamma) = 0$; como f é simétrica, obtemos

$$0 = f(\alpha, \alpha) + f(\alpha, \beta)$$
$$0 = f(\beta, \beta) + f(\alpha, \beta)$$

logo $f(\alpha, \alpha) = f(\beta, \beta)$. Como $f(\alpha, \beta) \geq 0$ e $f(\beta, \beta) \leq 0$, segue que

$$f(\alpha, \alpha) = f(\beta, \beta) = 0.$$

Mas f é positiva definida sôbre W e negativa definida sôbre V^-. Concluímos que $\alpha = \beta = 0$, e portanto que $\gamma = 0$ também.

Como

$$V = V^+ \oplus V^- \oplus V^\perp$$

e W, V^-, $V^\perp$ são independentes, vemos que $\dim W \leq \dim V^+$. Isto é, se W é um subespaço arbitrário de V sôbre o qual f é positiva definida, a dimensão de W não pode exeder a dimensão de V^+. Se $\mathcal{B}_1$ é uma outra base ordenada de V que satisfaz as condições do teorema, teremos subespaços correspondentes V_1^+, V_1^- e $V_1^\perp$; o argumento acima mostra que $\dim V_1^+ \leq \dim V^+$. Invertendo o argumento, obtemos $\dim V^+ \leq \dim V_1^+$ e, conseqüentemente,

$$\dim V^+ = \dim V_1^+.$$

Existem diversos comentários que devem ser feitos acêrca da base $\{\beta_1, \ldots, \beta_n\}$ do Teorema 5 e dos subespaços associados

V^+, V^- e $V^\perp$. Primeiro, notemos que $V^\perp$ é exatamente o subespaço dos vetores que são "ortogonais" a todo espaço V. Observamos acima que $V^\perp$ está contido neste subespaço; mas,

$$\dim V^\perp = \dim V - (\dim V^+ + \dim V^-) = \dim V - \text{pôsto}\,(f)$$

portanto, todo vetor α tal que $f(\alpha, \beta) = 0$ para todo β deve estar em $V^\perp$. Assim, o subespaço $V^\perp$ é único. Os subespaços V^+ e V^- não são únicos; contudo, suas dimensões são únicas. A demonstração do Teorema 5 nos mostra que $\dim V^+$ é a máxima dimensão possível para qualquer subespaço sôbre o qual f seja positiva definida. Anàlogamente, $\dim V^-$ é a máxima dimensão de qualquer subespaço sôbre o qual f seja negativa definida. É claro que

$$\dim V^+ + \dim V^- = \text{pôsto}\,(f).$$

O número

$$\dim V^+ - \dim V^-$$

freqüentemente é denominado a **assinatura** de f. Ela é introduzida porque as dimensões de V^+ e V^- são fàcilmente determinadas a partir do pôsto de f e da assinatura de f.

Talvez devamos fazer um comentário final a respeito da relação entre formas bilineares simétricas sôbre espaços vetoriais reais e produtos internos. Suponhamos que V seja um espaço vetorial real de dimensão finita e que V_1, V_2, V_3 sejam subespaços de V tais que

$$V = V_1 \oplus V_2 \oplus V_3.$$

Suponhamos que f_1 seja um produto interno sôbre V_1 e f_2 seja um produto interno sôbre V_2. Podemos então definir uma forma bilinear simétrica f sôbre V como segue: Se α, β são vetores em V, então podemos escrever

$$\alpha = \alpha_1 + \alpha_2 + \alpha_3 \text{ e } \beta = \beta_1 + \beta_2 + \beta_3$$

com α_j e β_j em V_j. Seja

$$f(\alpha, \beta) = f_1(\alpha_1, \beta_1) - f_2(\alpha_2, \beta\ ).$$

O subespaço $V^\perp$ para f será V_3, V_1 é um V^+ conveniente para f e V_2 é um V^- conveniente. Uma parte do enunciado do Teorema 5 é que tôda forma bilinear simétrica sôbre V surge desta maneira. O conteúdo adicional do teorema é que um produto interno é representado em relação a alguma base ordenada pela matriz unidade.

Exercícios

1. As seguintes expressões definem formas quadráticas q sôbre R^2. Determinar a forma bilinear simétrica f correspondente a cada q,

(a) ax_1^2;
(b) bx_1x_2
(c) cx_2^2;
(d) $2x_1^2 - \frac{1}{3}x_1x_2$;
(e) $x_1^2 + 9x_2^2$;
(f) $3x_1x_2 - x_2^2$;
(g) $4x_1^2 + 6x_1x_2 - 3x_2^2$.

2. Determinar a matriz, em relação à base ordenada canônica, e o pôsto de cada uma das formas bilineares do Exercício 1. Indicar quais formas são não-degeneradas.

3. Seja $q(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2$ a forma quadrática associada a uma forma bilinear simétrica f sôbre R^2. Mostrar que f é não-degenerada se, e sòmente se, $b^2 = 4ac \neq 0$.

4. Seja V um espaço vetorial de dimensão finita sôbre um subcorpo F do corpo dos números complexos e seja S o conjunto das formas bilineares simétricas sôbre V.

(a) Mostrar que S é um subespaço de $L(V, V, F)$.
(b) Determinar dim S.

Seja Q o conjunto de tôdas as formas quadráticas sôbre V.

(c) Mostrar que Q é um subespaço do espaço de tôdas as funções de V em F.
(d) Descrever explìcitamente um isomorfismo T de Q em S, sem referência a qualquer base.
(e) Seja U um operador linear sôbre V e q um elemento de Q. Mostrar que a equação $(U^tq)(\alpha) = q(U\alpha)$ define uma forma quadrática U^tq sôbre V.
(f) Se U é um operador linear sôbre V, mostrar que a função U^t definida na parte (e) é um operador linear sôbre Q. Mostrar que U^t é inversível se, e sòmente se, U é inversível.

5. Seja q a forma quadrática sôbre R^2 dada por

$$q(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2, \quad a \neq 0.$$

Determinar um operador linear U sôbre R^2 tal que

$$(U^tq(x_1, x_2) = ax_1^2 + \left(c - \frac{b^2}{a}\right)x_2^2.$$

(*Sugestão.* Para determinar U^{-1} (e portanto U), completar o quadrado. Para a definição de U^t, ver a parte (e) do Exercício 4.)

6. Seja q a forma quadrática sôbre R^2 dada por

$$q(x_1, x_2) = 2bx_1x_2.$$

Determinar um operador linear inversível U sôbre R^2 tal que

$$(U^tq)(x_1, x_2) = 2bx_1^2 - 2bx_2^2.$$

7. Seja q a forma quadrática sôbre R^3 dada por

$$q(x_1, x_2, x_3) = x_1x_2 + 2x_1x_3 + x_3^2.$$

Determinar um operador linear inversível U sôbre R^3 tal que

$$(U^tq)(x_1, x_2, x_3) = x_1^2 - x_2^2 + x_3^2.$$

(*Sugestão.* Exprimir U como um produto de operadores semelhantes àqueles usados nos Exercícios 5 e 6.)

8. Seja A uma $n \times n$ matriz simétrica sôbre R e seja q a forma quadrática sôbre R^n dada por

$$q(x_1, \ldots, x_n) = \sum_{i,j} A_{ij}x_ix_j.$$

Generalizar o método usado no Exercício 7 para mostrar que existe um operador linear inversível U sôbre R^n tal que

$$(U^tq)(x_1, \ldots, x_n) = \sum_{i=1}^{n} c_ix_i^2$$

onde c_i é 1, —1 ou 0, $i = 1, \ldots, n$.

9. Seja f uma forma bilinear simétrica sôbre R^n. Usar o resultado do Exercício 8 para demonstrar a existência de uma base ordenada $\mathfrak{B}$ tal que $[f]_{\mathfrak{B}}$ seja diagonal.

10. Seja V o espaço vetorial real das 2×2 matrizes hermitianas (complexas), isto é, 2×2 matrizes complexas A que satisfazem $A_{ij} = \overline{A_{ji}}$.

(a) Mostrar que a equação $q(A) = \det A$ define uma forma quadrática q sôbre V.

(b) Seja W o subespaço de V formado pelas matrizes de traço 0. Mostrar que a forma bilinear f determinada por q é negativa definida sôbre o subespaço W.

11. Seja V um espaço vetorial de dimensão finita e f uma forma bilinear simétrica não-degenerada sôbre V. Mostrar que para cada operador linear T sôbre V existe um único operador T' sôbre V tal que $f(T\alpha, \beta) = f(\alpha, T'\beta)$ para todos α, β em V. Mostrar também que

$$\begin{aligned} (T_1T_2)' &= T_2'T_1' \\ (c_1T_1 + c_2T_2)' &= c_1T_1' + c_2T_2' \\ (T')' &= T. \end{aligned}$$

Quanto disto acima continua válido sem a hipótese de que f é não-degenerada?

12. Seja F um corpo e V o espaço das $n \times 1$ matrizes sôbre F. Suponhamos que A seja uma $n \times n$ matriz fixa sôbre F e f seja a forma bilinear sôbre V definida por $f(X, Y) = X^tAY$. Suponhamos que f seja simétrica e não-degenerada. Seja B uma $n \times n$ matriz sôbre F e T o operador linear sôbre V que leva X em BX. Determinar o operador T' do Exercício 11.

13. Seja V um espaço vetorial de dimensão finita e f uma forma bilinear simétrica não-degenerada sôbre V. Associado a f existe um isomorfismo "natural" de V no espaço dual V^*, sendo êste isomorfismo a transformação L_f da Seção 9.1. Usando L_f, mostrar que para cada base $\mathfrak{B} = \{\alpha_1, \ldots, \alpha_n\}$

de V existe uma única base $\mathcal{B}' = \{\alpha_1', \ldots, \alpha_n'\}$ de V tal que $f(\alpha_i, \alpha_j') = \delta_{ij}$. Mostrar então que para todo vetor α em V temos

$$\alpha = \sum_i f(\alpha, \alpha_i')\alpha_i = \sum_i f(\alpha_i, \alpha)\alpha_i'.$$

14. Sejam V, f, $\mathcal{B}$ e $\mathcal{B}'$ como no Exercício 13. Suponhamos que T seja um operador linear sôbre V e que T' seja o operador que f associa a T, como no Exercício 11. Mostrar que

(a) $[T']_{\mathcal{B}'} = [T]_{\mathcal{B}}^t$
(b) traço (T) = traço $(T') = \sum_i f(T\alpha_i, \alpha_i')$.

15. Sejam V, f, $\mathcal{B}$ e $\mathcal{B}'$ como no Exercício 13. Suponhamos que $[f]_{\mathcal{B}} = A$. Mostrar que

$$\alpha_i' = \sum_j (A^{-1})_{ij}\alpha_j = \sum_j (A^{-1})_{ji}\alpha_j.$$

16. Seja F um corpo e V o espaço das $n \times 1$ matrizes sôbre F. Suponhamos que A seja uma $n \times n$ matriz simétrica inversível sôbre F e que f seja a forma bilinear sôbre V definida por $f(X, Y) = X^tAY$. Seja P uma $n \times n$ matriz inversível sôbre F e $\mathcal{B}$ a base de V formada pelas colunas de P. Mostrar que a base $\mathcal{B}'$ do Exercício 13 consiste das colunas da matriz $A^{-1}(P^t)^{-1}$.

17. Seja V um espaço vetorial de dimensão finita sôbre um corpo F e f uma forma bilinear simétrica sôbre V. Para cada subespaço W de V, seja $W^\perp$ o conjunto dos vetores α em V tais que $f(\alpha, \beta) = 0$ para todo β em W. Mostrar que

(a) $W^\perp$ é um subespaço;
(b) $V = \{0\}^\perp$;
(c) $V^\perp = \{0\}$ se, e sòmente se, f é não-degenerada;
(d) pôsto $(f) = \dim V^\perp$;
(e) se $\dim V = n$ e $\dim W = m$, então $W^\perp \geq n - m$

(*Sugestão.* Seja $\{\beta_1, \ldots, \beta_n\}$ uma base de W e consideremos a aplicação

$$\alpha \to (f(\alpha, \beta_1), \ldots, f(\alpha, \beta_m))$$

de V em F^m);

(f) a restrição de f a W é não-degenerada se, e sòmente se,

$$W \cap W^\perp = \{0\};$$

(g) $V = W \oplus W^\perp$ se, e sòmente se, a restrição de f a W é não-degenerada.

18. Seja V um espaço vetorial de dimensão finita sôbre C e f uma forma bilinear simétrica não-degenerada sôbre V. Demonstrar que existe uma base $\mathcal{B}$ de V tal que $\mathcal{B}' = \mathcal{B}$. (Ver o Exercício 13 para uma definição de $\mathcal{B}'$.)

9.3 Formas Bilineares Anti-Simétricas

Em tôda esta seção V será um espaço vetorial sôbre um subcorpo F do corpo dos números complexos. Uma forma bilinear f sôbre V

é dita **anti-simétrica** se $f(\alpha, \beta) = -f(\beta, \alpha)$ para todos os vetores α, β em V. Demonstraremos um teorema concernente à simplificação da matriz de uma forma bilinear anti-simétrica sôbre um espaço V de dimensão finita. Primeiro faremos algumas observações gerais.

Suponhamos que f seja uma forma bilinear *arbitrária* sôbre V. Se fizermos

$$g(\alpha, \beta) = \tfrac{1}{2}[f(\alpha, \beta) + f(\beta, \alpha)]$$
$$h(\alpha, \beta) = \tfrac{1}{2}[f(\alpha, \beta) - f(\beta, \alpha)]$$

então é fácil verificar que g é uma forma bilinear simétrica sôbre V e h é uma forma bilinear anti-simétrica sôbre V. Além disso $f = g + h$. Ainda mais, esta expressão de f como a soma de uma forma bilinear simétrica e uma anti-simétrica é única. Assim, o espaço $L(V, V, F)$ é a soma direta do subespaço das formas simétricas e o subespaço das formas anti-simétricas.

Se V é de dimensão finita, a forma bilinear f é anti-simétrica se, e sòmente se, sua matriz A em relação a alguma (ou tôda) base ordenada é anti-simétrica, $A^t = -A$. Isto é demonstrado da mesma maneira como se demonstra o fato correspondente sôbre formas bilineares simétricas. Quando f é anti-simétrica, a matriz de f em relação a qualquer base ordenada terá todos os seus elementos diagonais nulos. Isto corresponde exatamente à observação de que $f(\alpha,\alpha) = 0$ para todo α em V, uma vez que $f(\alpha, \alpha) = -f(\alpha, \alpha)$.

Suponhamos que f seja uma forma bilinear anti-simétrica não-nula sôbre V. Como $f \neq 0$, existem vetores α, β em V tais que $f(\alpha, \beta) \neq 0$. Multiplicando α por um escalar conveniente, podemos supor que $f(\alpha, \beta) = 1$. Seja γ um vetor arbitrário no subespaço gerado por α e β, digamos, $\gamma = c\alpha + d\beta$. Então

$$f(\gamma, \alpha) = f(c\alpha + d\beta, \alpha) = df(\beta, \alpha) = -d$$
$$f(\gamma, \beta) = f(c\alpha + d\beta, \beta) = cf(\alpha, \beta) = c$$

e então

$$\gamma = f(\gamma, \beta)\alpha - f(\gamma, \alpha)\beta. \tag{9-7}$$

Em particular, notemos que α e β são, necessàriamente, linearmente independentes; de fato, se $\gamma = 0$, então $f(\alpha, \alpha) = f(\gamma, \beta) = 0$.

Seja W o subespaço bidimensional gerado por α e β. Seja $W^{\perp}$ o conjunto dos vetores δ em V tais que $f(\delta, \alpha) = f(\delta, \beta) = 0$, isto é, o conjunto dos δ tais que $f(\delta, \gamma)$ para todo γ no subespaço W. Afirmamos que $V = W \oplus W^{\perp}$. De fato, seja ϵ um vetor arbitrário em V e

$$\gamma = f(\epsilon, \beta)\alpha - f(\epsilon, \alpha)\beta$$
$$\delta = \epsilon - \gamma.$$

Então γ está em W e δ está em $W^\perp$, pois

$$\begin{aligned} f(\delta, \alpha) &= f(\epsilon - f(\epsilon, \beta)\alpha + f(\epsilon, \alpha)\beta, \alpha) \\ &= f(\epsilon, \alpha) + f(\epsilon, \alpha)f(\beta, \alpha) \\ &= 0 \end{aligned}$$

e, anàlogamente, $f(\delta, \beta) = 0$. Assim, todo ϵ em V é da forma $\epsilon = \gamma + \delta$, com γ em W e δ em $W^\perp$. De (9-7) é evidente que $W \cap W^\perp = \{0\}$, portanto $V = W \oplus W^\perp$.

Ora, a restrição de f a $W^\perp$ é uma forma bilinear anti-simétrica sôbre $W^\perp$. Esta restrição pode ser a forma nula. Se não o fôr, existirão dois vetores α' e β' em $W^\perp$ tais que $f(\alpha', \beta) = 1$. Se indicarmos por W' o subespaço bidimensional gerado por α' e β', teremos

$$V = W \oplus W' \oplus W_0$$

onde W_0 é o conjunto dos vetores δ em $W^\perp$ tais que $f(\alpha', \delta) = f(\beta', \delta) = 0$. Se a restrição de f a W_0 não é a forma nula, podemos selecionar vetores α'', β'' em W_0 tais que $f(\alpha'', \beta'') = 1$, e então continuar.

No caso de dimensão finita, deveria estar evidente que obtemos uma seqüência finita de pares de vetores,

$$(\alpha_1, \beta_1),\ (\alpha_2, \beta_2),\ \ldots,\ (\alpha_k, \beta_k)$$

com as seguintes propriedades:

(i) $f(\alpha_j, \beta_j) = 1$, $j = 1, \ldots, k$.

(ii) $f(\alpha_i, \alpha_j) = f(\beta_i, \beta_j) = f(\alpha_i, \beta_j) = 0$, $i \neq j$.

(iii) Se W_j é o subespaço bidimensional gerado por α_j e β_j, então

$$V = W_1 \oplus \ldots \oplus W_k \oplus W_0$$

onde todo vetor em W_0 é 'ortogonal' a todos α_j e β_j, e a restrição de f a W_0 é a forma nula.

Teorema 6. *Seja* V *um espaço vetorial* n-*dimensional sôbre um subcorpo do corpo dos números complexos e seja* f *uma forma bilinear anti-simétrica sôbre* V. *Então o pôsto* r *de* f *é par e se* r = 2k, *existe uma base ordenada de* V *em relação à qual a matriz de* f *é a soma direta da* (n — r) × (n — r) *matriz nula e* k *cópias da* 2 × 2 *matriz*

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Demonstração. Sejam $\alpha_1, \beta_1, \ldots, \alpha_k, \beta_k$ vetores que satisfaçam as condições (i), (ii) e (iii) acima. Seja $\{\alpha_1, \ldots, \gamma_s\}$ uma base ordenada arbitrária do subespaço W_0. Então

$$\mathcal{B} = \{\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_k, \beta_k, \gamma_1, \ldots, \gamma_s\}$$

é uma base ordenada de V. De (i), (ii) e (iii) é evidente que a matriz de f em relação à base ordenada $\mathfrak{B}$ é a soma direta da $(n - 2k) \times \times (n - 2k)$ matriz nula e k cópias da 2×2 matriz

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{9-8}$$

Além disso, é evidente que o pôsto desta matriz, e portanto o pôsto de f, é $2k$.

Uma conseqüência disto acima é que se f é uma forma bilinear anti-simétrica não-degenerada sôbre V, então a dimensão de V deve ser par. Se $\dim V = 2k$, existe uma base ordenada $\{\alpha_1, \beta_1, \ldots, \ldots, \alpha_k, \beta_k\}$ de V tal que

$$f(\alpha_i, \beta_j) = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$
$$f(\alpha_i, \alpha_j) = f(\beta_i, \beta_j) = 0.$$

A matriz de f em relação a esta base ordenada é a soma direta de k cópias da 2×2 matriz anti-simétrica (9-8). Obtém-se uma outra forma canônica para a matriz de uma forma anti-simétrica não-degenerada se, ao invés da base ordenada acima, considera-se a base ordenada

$$\{\alpha_1, \ldots, \alpha_k, \beta_k, \ldots, \beta_1\}.$$

O leitor deverá achar fácil verificar que a matriz de f em relação à última base ordenada é da forma em blocos

$$\begin{bmatrix} 0 & J \\ -J & 0 \end{bmatrix}$$

onde J é a $k \times k$ matriz

$$\begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \end{bmatrix}.$$

Exercícios

1. Seja V um espaço vetorial sôbre um corpo F. Mostrar que o conjunto das formas bilineares anti-simétricas sôbre V é um subespaço de $L(V, V, F)$.
2. Determinar tôdas as formas bilineares anti-simétricas sôbre R^3.

3. Determinar uma base do espaço das formas bilineares anti-simétricas sôbre R^n.

4. Seja f uma forma bilinear simétrica C^n e g uma forma bilinear anti-simétrica sôbre C^n. Suponhamos que $f + g = 0$. Mostrar que $f = g = 0$.

5. Seja V um espaço vetorial n-dimensional sôbre um subcorpo F de C. Demonstrar o seguinte:

(a) A equação $(Pf)(\alpha, \beta) = \frac{1}{2}f(\alpha, \beta) - \frac{1}{2}f(\beta, \alpha)$ define um operador linear P sôbre $L(V, V, F)$.

(c) pôsto $(P) = \dfrac{n(n-1)}{2}$; nulidade $(P) = \dfrac{n(n+1)}{2}$.

(d) Se U é um operador linear sôbre V, a equação $(U^t f)(\alpha, \beta) = f(U\alpha, U\beta)$ define um operador U^t sôbre $L(V, V, F)$.

(e) Para todo operador linear U, a projeção P comuta com U^t.

6. Demonstrar um análogo do Exercício 11 na Seção 9.2 para formas bilineares anti-simétricas não-degeneradas.

7. Seja f uma forma bilinear sôbre um espaço vetorial V. Sejam L_f e R_f as aplicações de V em V^* associadas a f na Seção 9.1. Demonstrar que f é anti-simétrica se, e sòmente se, $L_f = -R_f$.

8. Demonstrar um análogo do Exercício 17 na Seção 9.2 para formas anti-simétricas.

9. Seja V um espaço vetorial de dimensão finita e L_1, L_2 funcionais lineares sôbre V. Mostrar que a equação

$$f(\alpha, \beta) = L_1(\alpha)L_2(\beta) - {}_1(\beta)L_2(\alpha)$$

define uma forma bilinear anti-simétrica sôbre V. Mostrar que $f = 0$ se, e sòmente se, L_1 e L_2 são linearmente independentes.

10. Seja V um espaço vetorial de dimensão finita sôbre um subcorpo do corpo dos números complexos e f uma forma bilinear anti-simétrica sôbre V. Mostrar que f tem pôsto 2 se, e sòmente se, existem funcionais lineares linearmente independentes L_1, L_2, sôbre V tais que

$$f(\alpha, \beta) = L_1(\alpha)L_2(\beta) - L_1(\beta)L_2(\alpha).$$

11. Seja f uma forma bilinear anti-simétrica arbitrária sôbre R^3. Demonstrar que existem funcionais lineares L_1, L_2 tais que

$$f(\alpha, \beta) = L_1(\alpha)L_2(\beta) - L_1(\beta)L_2(\alpha),$$

12. Seja V um espaço vetorial de dimensão finita sôbre um subcorpo do corpo dos números complexos e sejam f, g formas bilineares anti-simétricas sôbre V. Mostrar que existe um operador linear *inversível* T sôbre V tal que $f(T\alpha, T\beta) = g(\alpha, \beta)$ para todos α, β se, e sòmente se, f e g têm o mesmo pôsto.

13. Mostrar que o resultado do Exercício 12 é válido para formas bilineares simétricas sôbre um espaço vetorial complexo, mas não é válido para formas bilineares simétricas sôbre um espaço vetorial real.

9.4 Grupos que Conservam Formas Bilineares

Seja f uma forma bilinear sôbre o espaço vetorial V e seja T um operador linear sôbre V. Dizemos que T **conserva** f se $f(T\alpha, T\beta) =$

$=f(\alpha, \beta)$ para todos α, β em V. Para quaisquer T e f, a função g definida por $g(\alpha, \beta) = f(T\alpha, T\beta)$, como se vê fàcilmente, é uma forma bilinear sôbre V. Dizer que T conserva f é simplesmente dizer que $g = f$. O operador idêntico conserva tôda forma bilinear. Se S e T são operadores lineares que conservam f, o produto ST também conserva f; de fato, $f(ST\alpha, ST\beta) = f(T\alpha, T\beta) = f(\alpha, \beta)$.

Em outras palavras, a coleção de operadores lineares que conservam uma dada forma bilinear é fechada em relação à formação de produtos (de operadores). Em geral, não se pode dizer muito mais acêrca desta coleção de operadores; nò entanto, se f é não-degenerada, temos o que segue.

Teorema 7. *Seja* f *uma forma bilinear não-degenerada sôbre um espaço vetorial* V *de dimensão finita. O conjunto dos operadores lineares sôbre* V *que conservam* f *é um corpo em relação à operação de composição.*

Demonstração. Seja G o conjunto dos operadores lineares que conservam f. Observamos que o operador idêntico está em G e que, sempre que S e T estão em G, o composto ST também está em G. A partir do fato de que f é não-degenerada, demonstraremos que todo operador T em G é inversível e que T^{-1} também está em G. Suponhamos que T conserve f. Seja α um vetor no núcleo de T. Então, para todo β em V, temos

$$f(\alpha, \beta) = f(T\alpha, T\beta) = f(0, T\beta) = 0.$$

Como f é não-degenerada, $\alpha = 0$. Assim, T é inversível. Evidentemente T^{-1} também conserva f, pois

$$f(T^{-1}\alpha, T^{-1}\beta) = f(TT^{-1}\alpha, TT^{-1}\beta) = f(\alpha, \beta).$$

Se f é uma forma bilinear não-degenerada sôbre o espaço V de dimensão finita, então cada base ordenada $\mathcal{B}$ de V determina um grupo de matrizes que "conservam" f. O conjunto de tôdas as matrizes $[T]_{\mathcal{B}}$, onde T é um operador linear que conserva f, será um grupo em relação à multiplicação de matrizes. Existe uma descrição alternativa dêste grupo de matrizes, como segue. Seja $A = [f]_{\mathcal{B}}$, de modo que se α e β são vetores em V com respectivas matrizes de coordenadas X e Y em relação a $\mathcal{B}$, teremos

$$f(\alpha, \beta) = X^tAY.$$

Seja T um operador linear arbitrário sôbre V e $M = [T]_{\mathcal{B}}$. Então

$$\begin{aligned} f(T\alpha, T\beta) &= (MX)^tA(MY) \\ &= X^t(M^tAM)Y. \end{aligned}$$

Conseqüentemente, T conserva f se, e sòmente se, $M^tAM = A$. Em linguagem matricial, o Teorema 7 diz o seguinte: Se A é uma $n \times n$ matriz inversível, o conjunto das $n \times n$ matrizes M tais que $M^tAM = A$ é um grupo em relação à multiplicação matricial. Se $A = [f]_{\mathcal{B}}$, então M está neste grupo de matrizes se, e sòmente se, $M = [T]_{\mathcal{B}}$ onde T é um operador linear que conserva f.

Antes de passarmos a alguns exemplos, façamos mais uma observação. Suponhamos que f seja uma forma bilinear que seja simétrica. Um operador linear T conserva f se, e sòmente se, T conserva a forma quadrática

$$q(\alpha) = f(\alpha, \alpha)$$

associada a f. Se T conserva f, certamente temos

$$q(T\alpha) = f(T\alpha, T\alpha) = f(\alpha, \alpha) = q(\alpha)$$

para todo α em V. Recìprocamente, como f é simétrica, a identidade de polarização

$$f(\alpha, \beta) = \tfrac{1}{4}q(\alpha + \beta) - \tfrac{1}{4}q(\alpha - \beta)$$

nos mostra que T conserva f se $q(T\gamma) = q(\gamma)$ para cada γ em V. (Estamos supondo aqui que o corpo de escalares seja um subcorpo do corpo dos números complexos.)

Exemplo 6. Seja V o espaço R^n ou o espaço C^n. Seja f a forma bilinear

$$f(\alpha, \beta) = \sum_{j=1}^{n} x_j y_j$$

onde $\alpha = (x_1, \ldots, x_n)$ e $\beta = (y_1, \ldots, y_n)$. O grupo que conserva f é denominado o **grupo ortogonal** (real ou complexo) n-dimensional. O nome 'grupo ortogonal' é mais comumente aplicado ao grupo associado de matrizes em relação à base ordenada canônica. Como a matriz de f em relação à base canônica é I, êste grupo consiste das matrizes M que satisfazem $M^tM = I$. Uma tal matriz M é dita uma $n \times n$ **matriz ortogonal** (real ou complexa). Os dois $n \times n$ grupos ortogonais são usualmente indicados por $O(n, R)$ e $O(n, C)$. Evidentemente, o grupo ortogonal é também o grupo que conserva a forma quadrática

$$q(x_1, \ldots, x_n) = x_1^2 + \ldots + x_n^2.$$

Exemplo 7. Seja f a forma bilinear simétrica sôbre R^n com forma quadrática

$$q(x_1, \ldots, x_n) = \sum_{j=1}^{p} x_j^2 - \sum_{j=p+1}^{n} x_j^2.$$

Esta f é não-degenerada e tem assinatura $2p - n$. O grupo das matrizes que conservam uma forma dêste tipo é denominado um **grupo pseudo-ortogonal**. Quando $p = n$, obtemos o grupo ortogonal $O(n, R)$ como um tipo particular de grupo pseudo-ortogonal. Para cada um dos $n + 1$ valores $p = 0, 1, 2, \ldots, n$, obtemos uma forma bilinear diferente f; contudo, para $p = k$ e $p = n - k$ as formas são uma a oposta da outra e têm, portanto, o mesmo grupo associado. Assim, quando n é ímpar, temos $(n + 1)/2$ grupos pseudo-ortogonais de $n \times n$ matrizes e quando n é par, temos $(n + 2)/2$ dêsses grupos.

Teorema 8. *Seja* V *um espaço vetorial* n-*dimensional sôbre o corpo dos números complexos e seja* f *uma forma bilinear simétrica não-degenerada sôbre* V. *Então o grupo que conserva* f *é isomorfo ao grutpo ortogonal complexo* O(n, C).

Demonstração. Òbviamente, por um isomorfismo entre grupos, queremos dizer uma correspondência bijetora entre seus elementos que 'conserva' a operação de grupo. Seja G o grupo dos operadores lineares sôbre V que conservam a forma bilinear f. Como f é simétrica e não-degenerada, o Teorema 4 nos diz que existe uma base ordenada $\mathcal{B}$ de V em relação à qual f é representada pela $n \times n$ matriz unidade. Portanto, um operador linear T conserva f se, e sòmente se, sua matriz em relação à base ordenada $\mathcal{B}$ é uma matriz ortogonal complexa. Logo

$$T \to [T]_{\mathcal{B}}$$

é um isomorfismo de G em $O(n, C)$.

Teorema 9. *Seja* V *um espaço vetorial* n-*dimensional sôbre o corpo dos números reais e seja* f *uma forma bilinear simétrica não-degenerada sôbre* V. *Então, o grupo que conserva* f *é isomorfo a um* n $\times$ n *grupo pseudo-ortogonal.*

Demonstração. Repetir a demonstração do Teorema 8, usando o Teorema 5 em vez do Teorema 4.

Exemplo 8. Seja f a forma bilinear simétrica sôbre R^4 com forma quadrática

$$q(x, y, z, t) = t^2 - x^2 - y^2 - z^2.$$

Um operador linear T sôbre R^4 que conserva esta forma bilinear (ou quadrática) particular é denominado uma **transformação de Lorentz** e o grupo que conserva f é dito o **grupo de Lorentz.** Gostaríamos de dar um método para a descrição de algumas transformações de Lorentz.

Seja H o espaço vetorial real das 2×2 matrizes complexas A que sejam hermitianas, $A = A^*$. É fácil verificar que

$$\Phi(x, y, z, t) = \begin{bmatrix} t + x & y + iz \\ y - iz & t - x \end{bmatrix}$$

define um isomorfismo Φ de R^4 no espaço H. Por meio dêste isomorfismo, a forma quadrática q é levada sôbre a função determinante, isto é,

$$q(x, y, z, t) = \det \begin{bmatrix} t + x & y + iz \\ y - iz & t - z \end{bmatrix}$$

ou

$$q(\alpha) = \det \Phi\,(\alpha).$$

Isto sugere que podemos estudar as transformações de Lorentz sôbre R^4 estudando operadores lineares sôbre R que conservam determinantes.

Seja M uma 2×2 matriz complexa arbitrária e, para uma matriz hermitiana A, definamos

$$U_M(A) = MAM^*.$$

Ora, MAM^* também é hermitiana. A partir disto, é fácil ver que U_M é um operador linear (real) sôbre H. Perguntamos quando é que U_M 'conserva' determinantes, isto é, $\det[U_M(A)] = \det A$ para cada A em H. Como o determinante de M^* é o complexo conjugado do determinante de M, vemos que

$$\det\,[U_M(A)] = |\det M|^2 \det A.$$

Assim, U_M conserva determinantes exatamente quando $\det M$ tem valor absoluto 1.

Selecionemos então uma 2×2 matriz complexa arbitrária M para a qual $|\det M| = 1$. Então U_M é um operador linear sôbre H que conserva determinantes. Definamos

$$T_M = \Phi^{-1} U_M \Phi.$$

Como Φ é um isomorfismo, T_M é um operador linear sôbre R^4. Além disso, T_M é uma transformação de Lorentz, pois

$$\begin{aligned} q(T_M\alpha) &= q(\Phi^{-1}\, U_M\Phi\alpha) \\ &= \det\,(\Phi\Phi^{-1}U_M\Phi\alpha) \\ &= \det\,(U_M\Phi\alpha) \\ &= \det\,(\Phi\alpha) \\ &= q(\alpha) \end{aligned}$$

e portanto T_M conserva a forma quadrática q.

Usando 2×2 matrizes particulares M, pode-se usar o método acima para calcular transformações de Lorentz particulares. Dois comentários devem ser feitos neste ponto; êles não são difíceis de serem verificados.

(i) Se M_1 e M_2 são 2×2 matrizes inversíveis com elementos complexos, então $U_{M_1} = M_{M_2}$ se, e sòmente se, M_2 é um múltiplo escalar de M_1. Assim, tôdas as tránsformações de Lorentz acima exibidas podem ser obtidas a partir de matrizes unimodulares M, isto é, a partir de matrizes M que satisfazem $\det M = 1$. Se M_1 e M_2 são matrizes unimodulares tais que $M_1 \neq M_2$ e $M_1 \neq -M_2$ então $T_{M_1} \neq T_{M_2}$.

(ii) Nem tôda transformação de Lorentz pode ser obtida pelo método acima.

Exercícios

1. Seja M um membro do grupo ortogonal complexo, $O(n, C)$. Mostrar que M^t, $\overline{M}$ e $M^* = \overline{M}^t$ também pertencem a $O(n, C)$.
2. Suponhamos que M pertença a $O(n, C)$ e que M' seja semelhante a M. M' também pertence a $O(n, C)$?
3. Seja

$$y_j = \sum_{k=1}^{n} M_{jk}x_k$$

onde M é um membro de $O(n, C)$. Mostrar que

$$\sum_j y_j^2 = \sum_j x_j^2.$$

4. Seja M uma $n \times n$ matriz sôbre C com colunas $M_1, M_2, \ldots, M_n$. Mostar que M pertence a $O(n, C)$ se, e sòmente se,

$$M_j^t M_k = \delta_{jk}.$$

5. Seja X uma $n \times 1$ matriz sôbre C. Em que condições $O(n, C)$ contém uma matriz M cuja primeira coluna seja X?
6. Determinar uma matriz em $O(3, C)$ cuja primeira linha seja $(2i, 2i, 3)$.
7. Seja V o espaço das $n \times 1$ matrizes sôbre C e f a forma bilinear sôbre V dada por $f(X, Y) = X^tY$. Seja M pertencente a $O(n, C)$. Qual é a matriz de f em relação à base de V formada pelas colunas $M_1, M_2, \ldots, M_n$ de M?

8. Seja X uma $n \times 1$ matriz sôbre C tal que $X^tX = 1$ e seja I_j a j-ésima coluna da matriz unidade. Mostrar que existe uma matriz M em $O(n, C)$ tal que $MX = I_j$. Se X tem elementos reais mostrar que existe uma M em $O(n, R)$ com a propriedade de que $MX = I_j$.

9. Seja V o espaço das $n \times 1$ matrizes sôbre C, A uma $n \times n$ matriz sôbre C e f a forma bilinear sôbre V dada por $f(X, Y) = X^tAY$. Mostrar que f é invariante sob $O(n, C)$, isto é, $f(MX, MY) = f(X, Y)$ para tôdas X, Y em V e tôda M em $O(n, C)$, se, e sòmente se, A comuta com cada membro de $O(n, C)$.

10. Seja S um conjunto arbitrário de $n \times n$ matrizes sôbre C e S' o conjunto das $n \times n$ matrizes sôbre C que comutam com todo elemento de S. Mostrar que S' é uma álgebra sôbre C.

11. Seja F um subcorpo de C, V um espaço vetorial de dimensão finita sôbre F e f uma forma bilinear não-singular sôbre V. Se T é um operador linear sôbre V que conserva f, demonstrar que $\det T = \pm 1$.

12. Seja F um subcorpo de C, V o espaço das $n \times 1$ matrizes sôbre F, A uma $n \times n$ matriz inversível sôbre F e f a forma bilinear sôbre V dada por $f(X, Y) = X^tAY$. Se M é uma $n \times n$ matriz sôbre F, mostrar que M conserva f se, e sòmente se, $A^{-1}M^tA = M^{-1}$.

13. Seja g uma forma bilinear não-singular sôbre um espaço vetorial V de dimensão finita. Suponhamos que T seja um operador linear inversível sôbre V e que f seja a forma bilinear sôbre V dada por $f(\alpha, \beta) = g(\alpha, T\beta)$. Se U é um operador linear sôbre V, determinar condições necessárias e suficientes para que U conserve f.

14. Seja T um operador linear sôbre C^2 que conserva a forma quadrática $x_1^2 - x_2^2$. Mostrar que

(a) $\det (T) = \pm 1$;

(b) se M é a matriz de T em relação à base canônica, então $M_{22} = \pm M_{11}$, $M_{21} = \pm M_{12}$, $M_{11}^2 - M_{12}^2 = 1$;

(c) se $\det M = 1$, então existe um número complexo não-nulo c tal que

$$M = \frac{1}{2}\begin{bmatrix} c + \dfrac{1}{c} & c - \dfrac{1}{c} \\[2ex] c + \dfrac{1}{c} & c + \dfrac{1}{c} \end{bmatrix}$$

(d) se $\det M = -1$, então existe um número complexo c tal que

$$M = \frac{1}{2}\begin{bmatrix} c + \dfrac{1}{c} & c - \dfrac{1}{c} \\[2ex] -c + \dfrac{1}{c} & -c - \dfrac{1}{c} \end{bmatrix}.$$

15. Seja f a forma bilinear sôbre C^2 definida por

$$f((x_1, x_2), (y_1, y_2)) = x_1y_2 - x_2y_1.$$

Mostrar que

(a) se T é um operador linear sôbre C^2, então $f(T\alpha, T\beta) = (\det T)\, f(\alpha, \beta)$ para todos α, β em C^2;

(b) T conserva f se, e sòmente se, $\det T = +1$.
(c) O que é que (b) diz acêrca do grupo das 2×2 matrizes M tais que $M^tAM = A$ onde

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}?$$

16. Seja n um inteiro positivo, I a $n \times n$ matriz unidade sôbre C e J a $2n \times 2n$ matriz dada por

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}.$$

Seja M a $2n \times 2n$ matriz sôbre C da forma

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

onde A, B, C, D são $n \times n$ matrizes sôbre C. Determinar condições necessárias e suficientes sôbre A, B, C, D para que $M^tJM = J$.

17. Determinar tôdas as formas bilineares sôbre o espaço das $n \times 1$ matrizes sôbre R que sejam invariantes sob $O(n, R)$.

18. Determinar tôdas as formas bilineares sôbre o espaço das $n \times 1$ matrizes sôbre C que sejam invariantes sob $O(n, C)$.

APÊNDICE

Êste Apêndice divide-se de maneira lógica em duas partes. A primeira parte, compreendendo as três primeiras seções, contém certos conceitos fundamentais que ocorrem por todo o livro (na verdade, por tôda a matemática). É mais uma espécie de introdução do livro que apêndice. A segunda parte é mais genuìnamente um apêndice ao texto.

A Seção 1 contém uma discussão sôbre conjuntos, suas reuniões e interseções. A Seção 2 discute o conceito de função e as idéias afins de imagem, domínio, função inversa e a restrição de uma função a um subconjunto do seu domínio. A Seção 3 trata das relações de equivalência. O assunto destas três seções, especialmente o das Seções 1 e 2, é apresentado de uma maneira bem concisa. É tratado mais como um acôrdo sôbre a terminologia que como uma exposição detalhada. Num sentido lógico estrito, esta matéria constitui uma parte dos pré-requisitos para a leitura do livro; contudo, o leitor não deverá se desencorajar se não conseguir aprender completamente o significado das idéias na sua primeira leitura. Estas idéias são importantes, mas o leitor que não tiver muita familiaridade com elas deverá achar mais fácil absorvê-las se rever a discussão de tempos em tempos, à medida que fôr lendo o texto em si.

As Seções 4 e 5 consideram as relações de equivalência no contexto da álgebra linear. A Seção 4 contém uma discussão breve de espaços quocientes. Pode ser lida a qualquer momento após os dois ou três primeiros capítulos do livro. A Seção 5 considera ràpidamente algumas das relações de equivalência que aparecem no livro, tentando indicar como alguns dos resultados do livro poderiam ser interpretados do ponto de vista de relações de equivalência.

A.1 Conjuntos

Usaremos as palavras 'conjunto', 'classe', 'coleção' e 'família' indiferentemente, apesar de darmos preferência a 'conjunto'. Se S é um conjunto e x é um objeto do conjunto S, diremos x é um **mem-**

bro de S, que x é um **elemento de** S, que x **pertence** a S ou simplesmente que x está em S. Se S possui apenas um número finito de membros, $x_1, \ldots, x_n$, freqüentemente descrevemos S exibindo seus elementos dentro de chaves:

$$S = \{x_1, \ldots, x_n\}.$$

Assim, o conjunto S dos inteiros positivos de 1 a 5 seria

$$S = \{1, 2, 3, 4, 5\}.$$

Se S e T são conjuntos, dizemos que S é um **subconjunto de** T, ou que S está **contido em** T, se cada membro de S é um membro de T. Cada conjunto S é um subconjunto de si mesmo. Se S é um subconjunto de T mas S e T não são idênticos, denominamos S um **subconjunto próprio** de T. Em outras palavras, S é um subconjunto próprio de T se S está contido em T mas T não está contido em S.

Se S e T são conjuntos, a **reunião de** S **com** T é o conjunto $S \cup T$, constituído de todos os objetos x que são membros de S ou de T. A **interseção de** S **com** T é o conjunto $S \cap T$, formado por todos os x que são membros de S e de T. Para dois conjuntos arbitrários, S e T, a interseção $S \cap T$ é um subconjunto da reunião $S \cup T$. Isto deve auxiliar a esclarecer o uso da palavra 'ou' que prevalecerá neste livro: Quando dizemos que x está em S ou em T, não excluímos a possibilidade de x estar em ambos S e T.

Para que a interseção de S e T seja sempre um conjunto, é necessário introduzir o **conjunto vazio**, isto é, o conjunto sem elementos. Então $S \cap T$ é o conjunto vazio se, e sòmente se, S e T não têm elementos em comum.

Freqüentemente precisaremos discutir a reunião ou a interseção dos diversos conjuntos. Se $S_1, \ldots, S_n$ são conjuntos, sua **reunião** é o conjunto $\bigcup_{j=1}^{n} S_j$ formado por todos os x que são membros de pelo menos um dos conjuntos $S_1, \ldots, S_n$. Sua **interseção** é o conjunto $\bigcap_{j=1}^{n} S_j$, formado por todos os x que são membros de cada um dos conjuntos $S_1, \ldots, S_n$. Em algumas ocasiões, discutiremos a reunião ou a interseção de uma coleção infinita de conjuntos. Deveria estar evidente a maneira como tais reuniões e interseções são definidas. O exemplo que segue deverá esclarecer estas definições e uma notação para elas.

Exemplo 1. Indiquemos por R o conjunto dos (de todos os) números reais (a reta real). Se t está em R, associamos a t um subconjunto S_t de R, definido como segue: S_t consiste dos números reais x que não são menores que t.

(a) $S_{t_1} \cup S_{t_2} = S_t$ onde t é o menor entre t_1 e t_2.

(b) $S_{t_1} \cap S_{t_2} = S_t$, onde t é o maior entre t_1 e t_2.

(c) Seja I o intervalo unitário, isto é, o conjunto dos t em R que satisfazem $0 \leq t \leq 1$. Então

$$\bigcup_{t \text{ em } I} S_t = S_0$$

$$\bigcap_{t \text{ em } I} S_t = S_1$$

A.2 Funções

Uma **função** consiste do seguinte:

(i) um conjunto X, denominado o domínio da função;

(ii) um conjunto Y, denominado o contradomínio da função;

(iii) uma regra (ou correspondência) f, que associa a cada elemento x de X um único elemento $f(x)$ de Y.

Se (X, Y, f) é umafu nção, diremos que f é **uma função de** X **em** Y. Isto é um tanto confuso, pois não é f que é a função; f é a regra da função. No entanto, êste uso do mesmo símbolo para a função e e sua regra nos fornece uma maneira muito mais maleável de falar sôbre funções. Assim, diremos que f é uma função de X em Y, que X é o domínio de f e que Y é o contradomínio de f—tudo isto significando que (X, Y, f) é uma função como definido acima. Existem várias outras palavras que são comumente usadas no lugar da palavra 'função'. Algumas delas são 'transformação', 'operador' e 'aplicação'. Estas são usadas em contextos onde pareçam ser mais sugestivas na transmissão do papel desempenhado por uma função específica.

Se f é uma função de X em Y, a **imagem** de f é o conjunto dos $f(x)$, x em X. Em outras palavras, a imagem de f consiste dos elementos y em Y tais que $y = f(x)$, para algum x em X. Se a imagem de f é todo o conjunto Y, dizemos que f é uma **função sobrejetora de** X **em** Y ou simplesmente que f é **sobrejetora**. A imagem de f é freqüentemente indicada por $f(X)$.

Exemplo 2. (a) Seja X o conjunto dos números reais e seja $Y = X$. Seja f a função de X em Y definida por $f(x) = x^2$. A imagem de f é o conjunto dos números reais não-negativos. Assim, f não é sobrejetora.

(b) Seja X o plano euclidiano e $Y = X$. Seja f definida como segue: se P é um ponto do plano, então $f(P)$ é o ponto obtido girando-se P de $90°$ (em tôrno da origem, no sentido anti-horário). A imagem de f é todo Y, isto é, todo o plano; portanto, f é sobrejetora.

(c) Novamente, seja X o plano euclidiano. Coloquemos coordenadas em X como em geometria analítica, usando duas retas perpendiculares para identificarmos os pontos de X com pares ordenados de números reais (x_1, x_2). Seja Y o eixo dos x_1, isto é, o conjunto dos pontos (x_1, x_2) com $x_2 = 0$. Se P é um ponto de X, seja $f(P)$ ponto obtido projetando P sôbre o eixo dos x_1, paralelamente ao eixo dos x_2. Em outras palavras, $f((x_1, x_2)) = (x_1, 0)$. A imagem de f é todo Y, portanto f é sobrejetora.

(d) Seja X o conjunto dos números reais e seja Y o conjunto dos números reais positivos. Definamos uma função f de X em Y por $f(x) = e^x$. Então, f é uma função sobrejetora de X em Y.

(e) Seja X o conjunto dos números reais positivos e Y o conjunto dos números reais. Seja f a função logarítmica natural, isto é, função definida por $f(x) = \log x = \ln x$. Novamente, f é sobrejetora, isto é, todo número real é o logaritmo natural de algum número positivo.

Suponhamos que X, Y e Z sejam conjuntos, que f seja uma função de X em Y e que g seja uma função de Y em Z. Existe, associada a f e g, uma função $g \circ f$ de X em Z, conhecida como a **composta** de g e f. É definida por

$$(g \circ f)(x) = g(f(x)).$$

Para um exemplo simples, seja $X = Y = Z$, o conjunto dos números reais, sejam f, g, h as funções de X em X definidas por

$$f(x) = x^2, \; g(x) = e^x, \; h(x) = e^{x^2}$$

e então $h = g \circ f$. A composta $g \circ f$ é freqüentemente indicada por gf; contudo, como mostra o exemplo simples acima, existem ocasiões em que isto pode levar a confusão.

Uma questão de interêsse é a que segue. Suponhamos que f seja uma função de X em Y. Quando é que existe uma função g de Y em X tal que $g(f(x)) = x$ para todo x em X? Indicando por I a **função idêntica** sôbre X, isto é, a função de X em X definida por $I(x) = x$, estamos perguntando: Quando é que existe uma função g de Y em X tal que $g \circ f = I$? A grosso modo, queremos uma função g que 'leva cada elemento de Y de volta ao lugar de onde veio'. Para que uma tal g exista, f precisa ser **injetora**, isto é, f precisa ter a propriedade de que se $x_1 \neq x_2$, então $f(x_1) \neq f(x_2)$. Se f é injetora, existe uma tal função

g. Ela é definida como segue: Seja y um elemento de Y. Se y está na imagem de f, então existe um elemento x em X tal que $y = f(x)$; como f é injetora, existe exatamente um tal x. Definamos $g(y) = x$. Se y não está na imagem de f, definamos $g(y)$ como sendo um elemento qualquer de X. Òbviamente temos $g \circ f = I$.

Seja f uma função de X em Y. Dizemos que f é **inversível** se existe uma função g de Y em X tal que

(i) $g \circ f$ é a função idêntica sôbre X;

(ii) $f \circ g$ é a função idêntica sôbre Y.

Acabamos de ver que se existe uma g satisfazendo (*i*), então f é injetora. Anàlogamente, pode-se ver que se existe uma g satisfazendo (ii), a imagem de f é todo o conjunto Y, isto é, f é sobrejetora. Assim se f é inversível, f é injetora e sobrejetora (bijetora). Recìprocamente, se f é bijetora, existe uma função g de Y em X que satisfaz (i) e (ii). Além disso, esta g é única. Ela é a função de Y em X, definida por esta regra: se y está em Y, então $g(y)$ é o único elemento x em X para o qual $f(x) = y$.

Se f é inversível (bijetora) a **inversa** de f é a única função f^{-1} de Y em X que satisfaz

(i′) $f^{-1}(f(x)) = x$, para cada x em X,

(ii′) $f(f^{-1}(y)) = y$, para cada y em Y.

Exemplo 3. Consideremos as funções do Exemplo 2.

(a) Se $X = Y$, o conjunto dos números reais, e $f(x) = x^2$, então f não é inversível. De fato, f não é injetora nem sobrejetora.

(b) Se $X = Y$, o espaço euclidiano, e f é a 'rotação de 90°', então f é bijetora. A função inversa f^{-1} é a 'rotação de —90°' ou a 'rotação de 270°'.

(c) Se X é o plano, Y o eixo dos x_1 e $f((x_1, x_2)) = (x_1, 0)$, então f não é inversível. De fato, apesar de ser sobrejetora, f não é injetora.

(d) Se X é o conjunto dos números reais, Y o conjunto dos números reais positivos e $f(x) = e^x$, então f é inversível. A função f^{-1} é a função logaritmo natural da parte (e): $\log e^x = x$, $e^{\log y} = y$.

(e) A inversa desta função logarítmica natural é a função exponencial da parte (d).

Seja f uma função de X em Y e seja f_0 uma função de X_0 em Y_0. Dizemos que f_0 é uma **restrição** de f (ou uma restrição de f a X_0) se

(i) X_0 é um subconjunto de X,

(ii) $f_0(x) = f(x)$, para cada x em X_0.

Evidentemente, quando f_0 é uma restrição de f, decorre que Y_0 é um subconjunto de Y. O nome 'restrição' vem do fato de que f e f_0 têm a mesma regra e diferem principalmente porque restringimos o domínio de definição da regra ao subconjunto X_0 de X.

Se nos é dada uma função f e um subconjunto arbitrário X_0 de X, existe uma maneira óbvia de construir uma restrição de f a X_0. Definamos uma função f_0 de X_0 em Y por $f_0(x) = f(x)$ para cada x em X_0. Poder-se-ia perguntar por que não denominamos esta f_0 a restrição de f a X_0. A razão é que, ao discutirmos restrições de f, queremos a liberdade de mudar o contradomínio Y, bem como o domínio X.

Exemplo 4. (a) Seja X o conjunto dos números reais e f a função de X definida por $f(x) = x^2$. Então, f não é uma função inversível mas o será se restringirmos seu domínio aos números reais não-negativos. Seja X_0 o conjunto dos números reais não-negativos e seja f_0 a função de X_0 em X_0 definida por $f_0(x) = x^2$. Então f_0 é uma restrição de f a X_0. Ora, f não é injetora nem sobrejetora, enquanto que f_0 é injetora e sobrejetora. A última afirmação diz simplesmente que cada número não-negativo é o quadrado de exatamente um número não-negativo. A função inversa f_0^{-1} é a função de X_0 em X_0 definida por $f_0^{-1}(x) = \sqrt{x}$.

(b) Seja X o conjunto dos números reais e seja f a função de X em X definida por $f(x) = x^3 + x^2 + 1$. A imagem de f é todo X, portanto f é sobrejetora. A função f certamente não é injetora pois $f(-1) = f(0)$. Mas f é injetora sôbre X_0, o conjunto dos números reais não-negativos, pois a derivada de f é positiva para $x > 0$. Quando x percorre todos os números reais não-negativos, $f(x)$ percorre todos os números reais y tais que $y \geq 1$. Se indicarmos por Y_0 o conjunto dos $y \geq 1$ e por f_0 a função de X_0 em Y_0 definida por $f_0(x) = f(x)$, então f_0 é uma função bijetora de X_0 em Y_0. Conseqüentemente, f_0 possui uma função inversa f_0^{-1} de Y_0 em X_0. Qualquer fórmula para $f_0^{-1}(y)$ é bastante complicada.

(c) Novamente, seja X o conjunto dos números reais e seja f a função seno, isto é, a função de X definida por $f(x) = \text{sen } x$. A imagem de f é o conjunto dos y tais que $-1 \leq y \leq 1$; logo, f não é sobrejetora. Como $f(x + 2\pi) = f(x)$, vemos que f não é injetora. Se indicarmos por X_0 o intervalo $-\pi/2 \leq x \leq -\pi/2$, então f é injetora em X_0. Seja Y_0 o intervalo $-1 \leq y \leq 1$ e seja f_0 a função de X_0 em Y_0 definida por $f_0(x) = \text{sen } x$. Então f_0 é uma restrição de f ao intervalo X_0 e f_0 é bijetora. Esta é apenas uma outra maneira de dizer que, no intervalo de $-\pi/2$ a $\pi/2$, a função seno toma cada valor entre -1 e 1 exatamente uma vez. A função f_0^{-1} é a função arco-seno:

$$f_0^{-1}(y) = \text{sen}^{-1} y = \text{arc sen } y.$$

(d) Êste é um exemplo geral de uma restrição de uma função. É muito mais característico do tipo de restrição que usaremos neste livro do que os exemplos em (b) e (c) acima. O exemplo em (a) é um caso particular dêste. Seja X um conjunto e f uma função de X em si mesmo. Seja X_0 um subconjunto de X. Dizemos que X_0 é **invariante sob** f se para cada x em X_0 o elemento $f(x)$ está em X_0. Se X_0 é invariante sob f, então f induz uma função f_0 em X_0 em si mesmo, restringindo-se o seu domínio de definição a X_0. A importância da invariância é que retringindo f a X_0 podemos obter uma função de X_0 em si mesmo, em vez de simplesmente uma função de X_0 em X.

A.3 Relações de Equivalência

Uma relação de equivalência é um tipo particular de relação entre pares de elementos de um conjunto. Para definir uma relação de equivalência, precisamos primeiro decidir o que é uma 'relação'.

Certamente uma definição formal de 'relação' deve envolver relações familiares tais como '$x = y$', '$x < y$', ' x é a mãe de y' e 'x é mais velho que y'. Se X é um conjunto, o que é necessário para determinar uma relação entre pares de elementos de X? O que se precisa, evidentemente, é de uma regra para determinar se, para dois quaisquer elementos dados x e y em X, x está na relação dada com y ou não. Uma tal regra R, será denominada uma **relação** (binária) sôbre X. Se desejarmos ser ligeiramente mais precisos, poderemos proceder como segue. Indiquemos por $X \times X$ o conjunto dos pares ordenados (x, y) de elementos de X. Uma relação binária sôbre X é uma função R de $X \times X$ no conjunto $\{0, 1\}$. Em outras palavras, R associa a cada par ordenado (x, y) um 1 ou um 0. A idéia é que se $R(x, y) = 1$, então x está na relação dada com y, e se $R(x, y) = 0$, não o está.

Se R é uma relação binária sôbre o conjunto X, é conveniente escrever xRy quando $R(x, y) = 1$. Uma relação binária é dita

(i) **reflexiva**, se xRx para todo x em X;

(ii) **simétrica**, se yRx sempre que xRy;

(iii) **transitiva**, se xRz sempre que xRy e yRz.

Uma *relação de equivalência* sôbre X é uma relação binária sôbre X que é reflexiva, simétrica e transitiva.

Exemplo 5. (a) Sôbre qualquer conjunto, a igualdade é uma relação de equivalência. Em outras palavras, se xRy significa $x = y$, então R é uma relação de equivalência. De fato, $x = x$; se $x = y$ então $y = x$, se $x = y$ e $y = z$ então $x = z$. A relação '$x \neq y$' é simétrica, mas não é reflexiva nem transitiva.

(b) Seja X o conjunto dos números reais e suponhamos que xRy signifique $x < y$. Então R não é uma relação de equivalência. Apesar de ser transitiva, não é nem reflexiva nem simétrica. A relação '$x \leq y$' é reflexiva e transitiva, mas não é simétrica.

(c) Seja E o plano euclidiano e seja X o conjunto dos triângulos no plano E. Então, a congruência é uma relação de equivalência sôbre X, isto é, '$T_1 \cong T_2$' (T_1 é congruente a T_2) é uma relação de equivalência sôbre o conjunto dos triângulos no plano.

(d) Seja X o conjunto dos inteiros:

$$\ldots, -2, -1, 0, 1, 2, \ldots$$

Seja n um inteiro positivo fixo. Definamos uma relação R_n sôbre X por: xR_ny se, e sòmente se, $(x - y)$ é divisível por n. A relação R_n é denominada **congruência módulo** n. Em vez de xR_ny escreve-se em geral,

$$x \equiv y, \text{ mod } n \ (x \text{ é congruente a } y \text{ módulo } n)$$

quando $(x - y)$ é divisível por n. Para cada inteiro positivo n, a congruência módulo n é uma relação de equivalência sôbre o conjunto dos inteiros.

(e) Sejam X e Y conjuntos e f uma função de X em Y. Definamos uma relação R sôbre X por: x_1Rx_2 se, e sòmente se, $f(x_1) = f(x_2)$. É fácil verificar que R é uma relação de equivalência sôbre o conjunto X. Como veremos, êste exemplo engloba, na verdade, tôdas as relações de equivalência.

Suponhamos que R seja uma relação de equivalência sôbre o conjunto X. Se x é um elemento de X, indiquemos por $E(x; R)$ o conjunto dos elementos y em X tais que xRy. Êste conjunto $E(x; R)$ é denominado a **classe de equivalência** de x (segundo a relação de equivalência R). Como R é uma relação de equivalência, as classes de equivalência possuem as seguintes propriedades:

(i) Cada $E(x; R)$ é não-vazio, pois, como xRx, o elemento x pertence a $E(x; R)$.

(ii) Sejam x e y elementos de X. Como R é simétrica, y pertence a $E(x; R)$ se, e sòmente se, x pertence a $E(y; R)$.

(iii) Se x e y são elementos de X, as classes de equivalência $E(x; R)$ e $E(y; R)$ ou são idênticas ou não têm nenhum elemento em comum. Primeiro, suponhamos que xRy. Seja z um elemento arbitrário de $E(x; R)$, isto é, um elemento de X tal que xRz. Como R é simétrica, também temos zRx. Por hipótese, xRy e como R é transitiva, obtemos zRy, ou yRz. Isto mostra que todo membro de $E(x; R)$

é um membro de $E(y; R)$. Pela simetria de R, vemos, anàlogamente, que todo membro de $E(y; R)$ é um membro de $E(x; R)$; logo $E(x; R) = E(y; R)$. Afirmamos agora que se a relação xRy não vale, então $E(x; R) \cap E(y; R)$ é vazia. De fato, se z está em ambas estas classes de equivalência, temos xRz e yRz, portanto xRz e zRy, logo xRy.

Indicando por $\mathfrak{F}$ a família das classes de equivalência segundo uma relação de equivalência R, vemos que: (1) cada conjunto na família $\mathfrak{F}$ é não-vazio; (2) cada elemento x de X pertence a um e sòmente um dos conjuntos na família $\mathfrak{F}$; (3) xRy se, e sòmente se, x e y pertencem ao mesmo conjunto na família 5. Abreviadamente, a relação de equivalência R subdivide X na reunião de uma família de subconjuntos (não-vazios) disjuntos dois a dois. O argumento também vale no outro sentido. Suponhamos que $\mathfrak{F}$ seja uma família arbitrária de subconjuntos de X que satisfaça as condições (1) e (2) acima. Se definirmos uma relação R por (3), então R será uma relação de equivalência sôbre X e $\mathfrak{F}$ a família das classes de equivalência segundo R.

Exemplo 6. Vejamos quais são as classes de equivalência segundo as relações de equivalência do Exemplo 5.

(a) Se R é a igualdade sôbre o conjunto X, então a classe de equivalência do elemento x é simplesmente o conjunto $\{x\}$, cujo único elemento é x.

(b) Se X é o conjunto dos triângulos num plano e R é a relação de congruência, pràticamente tudo o que se pode dizer, de início, é que a classe de equivalência do triângulo T consiste de todos os triângulos que são congruentes a T. Uma das tarefas da geometria plana é dar outras descrições destas classes de equivalência.

(c) Se X é o conjunto dos inteiros e R_n é a relação 'congruência módulo n', então existem precisamente n classes de equivalência. Cada inteiro x pode ser expresso de um único modo sob a forma $x = qn + r$, onde q e r são inteiros e $0 \leq r \leq n - 1$. Isto mostra que cada x é congruente módulo n a exatamente um dos n inteiros $0, 1, 2, \ldots, n-1$. As classes de equivalência são

$$\begin{aligned}
E_0 &= \{\ldots, -2n, -n, 0, n, 2n, \ldots\} \\
E_1 &= \{\ldots, 1-2n, 1-n, 1, 1+n, 1+2n, \ldots\} \\
&\;\;\vdots \\
E_{n-1} &= \{\ldots, n-1-2n, n-1-n, n-1, n-1+n, \\
&\qquad n-1+2n, \ldots\{.
\end{aligned}$$

(d) Suponhamos que X e Y sejam conjuntos, f uma função de X em Y e R a relação de equivalência definida por: x_1Rx_2 se, e sòmente se, $f(x_1) = f(x_2)$. As classes de equivalência para R são exatamente os maiores subconjuntos de X sôbre os quais f é 'constante'. Outra descrição das classes de equivalência é a que segue. Elas estão em correspondência bijetora com os elementos na imagem de f. Se y está na imagem de f, o conjunto dos x em X tais que $f(x) = y$ é uma classe de equivalência para R; isto define uma correspondência bijetora entre os membros da imagem de f e as classes de equivalência R.

Façamos mais um comentário acêrca de relações de equivalência. Dada uma relação de equivalência R sôbre X, seja $\mathfrak{F}$ a família das classes de equivalência segundo R. Associando ao elemento x a classe de equivalência $E(x; R)$, definimos uma função f de X em $\mathfrak{F}$ (na verdade, sôbre $\mathfrak{F}$):

$$f(x) = E(x; R).$$

Isto mostra que R é a relação de equivalência associada a uma função cujo domínio é X, como no Exemplo 5 (e). O que isto nos diz é que tôda relação de equivalência sôbre o conjunto X é determinada como segue. Temos uma regra (função) f que associa a cada elemento x de X um objeto $f(x)$, e xRy se, e sòmente se, $f(x) = f(y)$. Agora deve-se considerar $f(x)$ como uma propriedade de x, de modo que o que a relação de equivalência faz (a grosso modo) é reunir todos os elementos de X que têm esta propriedade em comum. Se o objeto $f(x)$ é a classe de equivalência de x, então tudo o que se disse é que a propriedade comum dos membros de uma classe de equivalência é que êles pertencem à mesma classe de equivalência. É claro que isto não diz muito. Em geral, existem muitas funções distintas f que determinam a dada relação de equivalência como acima, e um objetivo no estudo das relações de equivalência é determinar uma tal f que dê uma descrição significativa e elementar da relação de equivalência. Na Seção A.5 veremos como isto é conseguido para algumas relações de equivalência particulares que surgem em álgebra linear.

A.4 Espaços Quocientes

Seja V um espaço vetorial sôbre o corpo F e seja W um subespaço de V. Existem, em geral, muitos subespaços W' que são suplementares de W, isto é, subespaços com a propriedade de que $V = W \oplus W'$. Se temos um produto interno sôbre V e se W é de dimensão finita, existe um subespaço particular que provàvelmente se denominaria o subespaço suplementar 'natural' de W. Êle é o suplementar ortogonal de W. Mas se V não possui nenhuma estrutura além de sua estrutura de espaço vetorial, não existe nenhuma maneira de selecionar um subespaço W' que se pudesse denominar o subespaço suplementar natural de W. Contudo, pode-se construir, a partir de V e W, um espaço vetorial V/W, conhecido como o 'quociente' de V por W, que desempenhará o papel do suplementar natural de W. Êste espaço quociente não é um subespaço de V, portanto não pode ser realmente um subespaço suplementar de W; no entanto, êle é um espaço vetorial definido apenas em têrmos de V e W que tem a propriedade de ser isomorfo a todo subespaço W' que seja suplementar de W.

Seja W um subespaço do espaço vetorial V. Se α e β são vetores em V, dizemos que α é **congruente** a β **módulo** W se o vetor $(\alpha - \beta)$ está no subespaço W. Se α é congruente a β módulo W, escrevemos

$$\alpha \equiv \beta, \quad \text{mod } W.$$

Ora, a congruência módulo W é uma relação de equivalência sôbre V.

(i) $\alpha \equiv \alpha$, mod W, porque $\alpha - \alpha = 0$ está em W.

(ii) Se $\alpha \equiv \beta$, mod W, então $\beta \equiv \alpha$, mod W. De fato, como W é um subespaço de V, o vetor $(\alpha - \beta)$ está em W se, e sòmente se, $(\beta - \alpha)$ está em W.

(iii) Se $\alpha \equiv \beta$, mod W, e $\beta \equiv \gamma$, mod W, então $\alpha \equiv \gamma$, mod W. De fato, se $(\alpha - \beta)$ e $(\beta - \gamma)$ estão em W, então $\alpha - \gamma = (\alpha - \beta) + (\beta - \gamma)$ está em W.

As classes de equivalência desta relação de equivalência são conhecidas como as **classes laterais** de W. Qual é a classe de equivalência (classe lateral) de um vetor α? Ela consiste dos vetores β em V tais que $(\beta - \alpha)$ está em W, isto é, os vetores da forma $\beta = \alpha + \gamma$, com γ em W. Por esta razão, a classe lateral do vetor α é indicada por

$$\alpha + W.$$

É conveniente pensar na classe lateral de α em relação a W como sendo o conjunto dos vetores obtidos por translação do subespaço W pelo vetor α. Para visualizar estas classes laterais, o leitor pode pensar no seguinte caso particular: Seja V o espaço R^2 e seja W um subespaço unidimensional de V. Se imaginarmos V como sendo o plano euclidiano, W será uma reta passando pela origem. Se $\alpha = (x_1, x_2)$ é um vetor em V, a classe lateral $\alpha + W$ é a reta que passa pelo ponto (x_1, x_2) e é paralela a W.

A coleção de tôdas as classes laterais de W será indicada por V/W. Definamos agora uma adição de vetores e uma multiplicação escalar sôbre V/W como segue:

$$(\alpha + W) + (\beta + W) = (\alpha + \beta) + W$$
$$c(\alpha + W) = (c\alpha) + W.$$

Em outras palavras, a soma da classe lateral de α com a classe lateral de β é a classe lateral de $(\alpha + \beta)$, e o produto do escalar c pela classe lateral de α é a classe lateral do vetor $c\alpha$. Ora, muitos vetores distintos em V terão a mesma classe lateral em relação a W, portanto precisamos verificar que a soma e o produto acima dependem sòmente das classes laterais envolvidas. O que isto significa é que precisamos mostrar o seguinte:

(1) Se $\alpha \equiv \alpha'$, mod W, e $\beta \equiv \beta'$, mod W, então
$\alpha + \beta \equiv \alpha' + \beta'$, mod W.

(2) Se $\alpha \equiv \alpha'$, mod W, então $c\,\alpha \equiv c\alpha'$, mod W.

Êstes fatos são fáceis de verificar. (1) Se $\alpha - \alpha'$ está em W e $\beta - \beta'$ está em W, então como $(\alpha + \beta) - (\alpha' + \beta') = (\alpha - \alpha') + (\beta - \beta')$ vemos que $\alpha + \beta$ é congruente a $\alpha' + \beta'$ módulo W. (2) Se $\alpha - \alpha'$ está em W e c é um escalar arbitrário, então $c\alpha - c\alpha' = c(\alpha - \alpha')$ está em W.

Agora é fácil verificar que V/W, com a adição de vetores e a multiplicação escalar acima definidas, é um espaço vetorial sôbre o corpo F. Deve-se verificar diretamente cada um dos axiomas para um espaço vetorial. Cada uma das propriedades da adição de vetores e da multiplicação escalar decorre da propriedade correspondente das operações em V. Um comentário deve ser feito. O vetor nulo em V/W será a classe lateral do vetor nulo em V. Em outras palavras, W é o vetor nulo em V/W.

O espaço vetorial V/W é denominado o **quociente** (ou diferença) de V por W. Existe uma transformação linear natural Q de V sôbre V/W. É definida por $Q(\alpha) = \alpha + W$. Deve-se ver que definimos as operações em V/W exatamente de modo que esta transformação Q viesse a ser linear. Notemos que o núcleo de Q é exatamente o subespaço W. Denominamos Q a **transformação quociente** (ou **aplicação quociente**) de V sôbre V/W.

A relação entre o espaço quociente V/W e subespaços de V que são suplementares de W pode agora ser enunciada como segue.

Teorema. *Seja* W *um subespaço do espaço vetorial* V *e seja* Q *a aplicação quociente de* V *sôbre* V/W. *Suponhamos que* W *seja um* *subespaço de* V. *Então* V = W ⊕ W' *se, e sòmente se, a restrição de* Q *a* W' *é um isomorfismo de* W' *em* V/W.

Demonstração. Suponhamos que $V = W \oplus W'$. Isto significa que cada vetor α em V pode ser expresso de um único modo sob a forma $\alpha = \gamma + \gamma'$, com γ em W e α' em W'. Então, $Q\alpha = Q\gamma + Q\gamma$, isto é, $\alpha + W = \gamma' + W$. Isto mostra que Q leva W' sôbre V/W, ou seja, que $Q(W') = V/W$. Além disso, Q é injetora em W'; de fato, suponhamos que γ_1' e γ_2' sejam vetores em W' e que $Q\gamma_1' = Q\gamma_2'$. Então, $Q(\gamma_1' - \gamma_2') = 0$ de modo que $\gamma_1' - \gamma_2'$ está em W. Êste vetor também está em W', que é disjunto de W; logo $\gamma_1' - \gamma_2' = 0$. A restrição de Q a W' é portanto uma transformação linear bijetora de W' em V/W.

Suponhamos que W' seja um subespaço de V tal que Q seja injetora em W' e que $Q(W') = V/W$. Seja α um vetor em V. Então existe um vetor γ' em W' tal que $Q\gamma' = Q\alpha$, isto é, $\gamma' + W = \alpha + W$. Isto significa que $\alpha = \gamma + \gamma'$ para algum vetor γ em W. Portanto $V = W + W'$. Para ver que W e W' são disjuntos, suponhamos que γ esteja em W e em W'. Como γ está em W, temos $Q\gamma = 0$. Mas Q é injetora em W', logo devemos ter que $\gamma = 0$. Assim, temos que $V = W \oplus W'$.

O que êste teorema realmente diz é que W' é um suplementar de W se, e sòmente se, W' é um subespaço que contém exatamente um elemento de cada classe lateral de W. Êle mostra que, quando $V = W \oplus W'$, a aplicação quociente Q 'identifica' W' com V/W. Abreviadamente, $(W \oplus W')/W$ é isomorfo a W' de uma maneira 'natural'.

Um fato bastante óbvio deve ser notado. Se W é um subespaço do espaço vetorial V de dimensão finita, então

$$\dim W + \dim (V/W) = \dim V.$$

Pode-se ver isto a partir do teorema acima. Talvez seja mais fácil observar que esta fórmula sôbre dimensões diz:

$$\text{nulidade } (Q) + \text{pôsto } (Q) = \dim V.$$

Não é nosso objetivo aqui fazer um tratamento detalhado dos espaços quocientes. Contudo, existe um resultado fundamental que devemos demonstrar.

Teorema. *Sejam* V *e* Z *espaços vetoriais sobre o corpo* F. *Suponhamos que* T *seja uma transformação linear de* V *sôbre* Z. *Se* W *é o núcleo de* T, *então* Z *é isomorfo a* V/W.

Demonstração. Definamos uma transformação U de V/W em Z por $U(\alpha + W) = T\alpha$. Precisamos verificar que U está bem definida, isto é, que se $\alpha + W = \beta + W$ então $T\alpha = T\beta$. Isto decorre do fato de que W é o núcleo de T; de fato, $\alpha + W = \beta + W$ significa que $\alpha - \beta$ em W e isto ocorre se, e sòmente se, $T(\alpha - \beta) = 0$. Isto mostra não só que U está bem definida, mas também que U é injetora.

Agora é fácil verificar que U é linear e leva V/W sôbre Z, pois T é uma transformação linear de V sôbre Z.

A.5 Relações de Equivalência em Álgebra Linear

Vamos considerar algumas das relações de equivalência que surgem no texto dêste livro. Esta é apenas uma amostra de tais relações.

(1) Sejam m e n inteiros positivos e F um corpo. Seja X o conjunto das $m \times n$ matrizes sôbre F. Então, a linha-equivalência é uma relação de equivalência sôbre o conjunto X. A afirmação 'A é linha-equivalente a B' significa que A pode ser obtida de B por uma sucessão finita de operações elementares sôbre linhas. Se escrevermos $A \sim B$ para indicar que A é linha-equivalente a B, então não é difícil verificar as propriedades (i) $A \sim A$; (ii) se $A \sim B$, então $B \sim A$; (ii) se $A \sim B$ e $B \sim C$, então $A \sim C$. Que sabemos a respeito desta relação de equivalência? Na realidade, sabemos bastante. Por exemplo, sabemos que $A \sim B$ se, e sòmente se, $A = PB$ para alguma $m \times m$ matriz inversível P; ou, $A \sim B$ se, e sòmente se, os sistemas homogêneos de equações lineares $AX = 0$ e $BX = 0$ têm as mesmas soluções. Também temos informações bem explícitas sôbre as classes de equivalência segundo esta relação. Cada $m \times m$ matriz A é linha-equivalente a uma, e sòmente uma, matriz linha-reduzida à forma em escada. O que isto diz é que cada classe de equivalência segundo esta relação contém precisamente uma matriz R linha-reduzida à forma em escada; a classe de equivalência determinada por R consiste das matrizes $A = PR$ onde P é uma $m \times m$ matriz inversível. Pode-se também considerar esta descrição das classes de equivalência da seguinte maneira: Dada uma $m \times n$ matriz A, temos uma regra (função) f que associa a A a matriz $f(A)$, linha-reduzida à forma em escada, que é linha-equivalente a A. A linha-equivalência é completamente determinada por f. De fato, $A \sim B$ se, e sòmente se, $f(A) = f(B)$, isto é, se, e sòmente se, A e B têm a mesma forma em escada linha-reduzida.

(2) Seja n um inteiro positivo e F um corpo. Seja X o conjunto das $n \times n$ matrizes sôbre F. Então, a semelhança é uma relação de

equivalência sôbre X; cada $n \times n$ matriz A é semelhante a si mesma; se A é semelhante a B, então B é semelhante a A; se A é semelhante a B e B é semelhante a C, então A é semelhante a C. Também sabemos muita coisa acêrca desta relação de equivalência. Por exemplo, A é semelhante a B se, e sòmente se, A e B representam o mesmo operador linear sôbre F_n em relação a bases ordenadas (possìvelmente) diferentes. Entretanto, sabemos algo muito mais profundo que isto. Cada $n \times n$ matriz A sôbre F é semelhante (sôbre F) a uma, e sòmente uma, matriz sob a forma racional (Capítulo 7). Em outras palavras, cada classe de equivalência segundo a relação de semelhança contém exatamente uma matriz que está sob a forma racional. Uma matriz sob a forma racional é determinada por uma k-upla $(p_1, \ldots, p_k)$ de polinômios unitários com a propriedade de que p_{j+1} divide p_j, $j = 1, \ldots, k - 1$. Assim, temos uma função f que associa a cada $n \times n$ matriz A uma k-upla $f(A) = (p_1, \ldots, p_k)$ que satisfaz a condição de divisibilidade p_{j+1} divide p_j. Então A e B são semelhantes se, e sòmente se, $f(A) = f(B)$.

(3) Eis um caso particular do Exemplo 2 acima. Seja X o conjunto das 3×3 matrizes sôbre um corpo F. Consideremos a relação de semelhança sôbre X. Se A e B são 3×3 matrizes sôbre F, então A e B são semelhantes se, e sòmente se, possuem o mesmo polinômio característico e o mesmo polinômio minimal. Associado a cada 3×3 matriz A, temos um par (f, p) de polinômios unitários que satisfazem

(a) $\operatorname{gr}(f) = 3$.
(b) p divide f,

sendo f o polinômio característico de A e p o polinômio minimal de A. Dados polinômios unitários f e p sôbre F que satisfaçam (a) e (b), é fácil exibir uma 3×3 matriz sôbre F, cujos polinômios característico e minimal sejam f e p, respectivamente. O que tudo nos diz é o que segue. Se considerarmos a relação de semelhança sôbre o conjunto das 3×3 matrizes sôbre F, as classes de equivalência estarão em correspondência bijetora com os pares ordenados (f, p) de polinômios unitários sôbre F que satisfazem (a) e (b).

BIBLIOGRAFIA

Albert, A. A. *Modern higher algebra*. University of Chicago Press, Chicago, 1937.

Birkhoff, G. and MacLane, S. *A survey of modern algebra*. The Macmillan Co., New York, 1941.

Halmos, P. *Finite-dimensional vector spaces*. D. Van Nostrand Co., Princeton, 1958.

Jacobson, N. *Lectures in abstract algebra*, II. D. Van Nostrand Co., Princeton, 1953.

Schreier, O. and Sperner, E. *Modern algebra and matrix theory*. Chelsea Publishing Co., New York, 1955.

van der Waerden, B. L. *Modern Algebra* (two volumes). Frederick Ungar Publishing Co., New York, 1949.

ÍNDICE ALFABÉTICO

A

Adjunta:
- clássica, 160
- de uma transformação, 253, 255

Álgebra, 115
Algèbricamente fechado, corpo, 135
Alternada, função n-linear, 142
Anel, 138
Anulador:
- de subespaço, 104
- de vetor (T-anulador), 202

Associada, matriz, 204
Auto-adjunta (o):
- matriz, 38, 285
- operador, 258

B

Base, 44
- canônica de F^n, 44
- dual, 99
- ordenada, 51
- ortonormal, 249

Bessel, desigualdade de, 248
Bilinear, forma, 305
- anti-simétrica, 323
- assinatura de, 320
- diagonalização de, 316
- grupo que conserva, 327
- matriz de, 308
- não-degenerada, 312
- positiva definida, 315
- pôsto de, 312
- simétrica, 314

C

Característico:
- poilnômio, 178
- valor, 177
- vetor, 177

Cauchy-Schwarz, desigualdade de, 244
Cayley-Hamilton, teorema de, 179
Cayley, transformada de, 280
Cíclico:
- subespaço, 201
- vetor, 201

Cofator, 159
Completa, matriz, 15
Conjugação, 242
Constante, polinômio, 118
Coordenadas, 50
- matriz de coordenadas, 52
- coordenadas em relação a uma base, 51

Corpo, 2
- algèbricamente fechado, 135
- subcorpo, 2

Cramer, regra de, 163

D

Determinante, função, 143
- existência de, 146
- unicidade de, 153

Diagonalizável (eis):
- operador, 134
- parte diagonalizável de operador, 196
- simultâneamente, 190, 302

Dimensão, 46
Direta, soma, 167, 168
- de matrizes, 172
- de operadores, 171

Disjuntos, subespaços, 166
Divisão com resto, 127
Dual:
- base, 99
- espaço, 97

E

Elementar:
- matriz, 21
- operação elementar sôbre linhas, 7
- operação elementar sôbre colunas, 28

Equivalência, relação de, 341
Escalar, 2

Espectral, teorema, 295
Euclidiano, espaço, 243

F

Finita, de dimensão, 46
Função, 337
inversível, 338
par e ímpar, 43
polinomial, 33

G

Gram-Schmidt, processo de, 246
Grupo, 81
de Lorentz, 331
ortogonal, 329
pseudo-ortogonal, 330
que conserva uma forma, 327
simétrico, 153
unitário, 275

H

Hermitiana, v. Auto-adjunta
Homogêneas, sistema de equações lineares, 3

I

Ideal, 128
ideal principal, 128
Idempotente, transformação, 173
Idêntica, função, 338
Interno, produto, 235, 315
espaço com, 243
forma quadrática de, 238
matriz de, 240
Interseção, 336
de subespaços, 39
Invariante, subespaço, 169
sob função, 341
Inversível:
função, 338
matriz, 24
transformação linear, 77
Irredutível, polinômio, 132
Isomorfismo:
de álgebras, 124
de espaços com produto interno, 271
de espaços vetoriais, 83
de grupos, 330

J

Jordan, forma de, de uma matriz, 221

K

Kronecker, símbolo de, 10

L

Lagrange, fórmula de interpolação de, 123
Linear, álgebra, 115
Linear, combinação:
de equações, 4
de vetores, 34
Lineares, equações (sistemas de), 3
bases de soluções de, 66
homogêneas, 3
regra de Cramer para a resolução de, 163
sistemas equivalentes de, 5
Linear, funcional, 97
Linearmente dependente (independente), 43
Linear, transformação (operador), 67, 76
adjunta (o), 255
auto-adjunta, 258
combinações lineares de, 73
decomposição racional de, 211
diagonalizável, 184
matriz de, 85, 88, 255
não-negativa, 299
não-singular, 78
nilpotente, 196
nulidade de, 71
parte diagonalizável de, 196
positiva, 260
pôsto de, 71
produtos de, 76
semi-simples, 227
transposta de, 110
unitária, 275
Linha-equivalência, 8
resumo de, 57
Linha-reduzida, matriz, 10
matriz linha-reduzida à forma em escada, 12
Linhas, operações sôbre, 6

Lorentz:
grupo de, 331
transformação de, 281, 331

M

Matriz (es), 6
associada, 204
anti-simétrica, 164, 169
auto-adjunta (hermitiana), 38, 285
completa, 15
das coordenadas, 52
de forma bilinear, 308
de produto interno, 240
de transformação linear, 85, 88, 255
de Vandermonde, 123
dos coeficientes, 6
elementar, 21
elementar de Jordan, 221
espaço de, 31
forma de Jordan de, 221
forma racional de, 214
inversível, 24
linha-reduzida, 10
linha-reduzida à forma em escada, 12
normal, 287
ortogonal, 164, 276
positiva, 263
produto de, 18
semelhança de, 94
simétrica, 38, 169, 285
traço de, 98
transposta, 111
triangular, 155
unitária, 164, 275
Máximo divisor comum, 130
Menores principais, 266
Minimal, polinômio, 181

N

n-linear, função, 140
alternada, 142
Não-negativo, operador, 299
Não-singular, transformação linear, 78
Nilpotente, operador, 196
Norma, 239
Normal, operador, 282
resolução espectral de, 296
Núcleo de transformação linear, 70
Nulidade de transformação linear, 71

O

Ordenação, 51
Ordenada, base, 51
Ortogonal (ais),
equivalência de matrizes, 278
grupo, 329
matriz, 164, 276
projeção, 249
suplementar, 247
vetores, 244
Ortonormal,
base, 249
conjunto, 244

P

Paralelogramo, regra de, 241
Permutação (ões), 150
par, ímpar, 151
produto de, 154
sinal de, 151
Polar, decomposição, 301
Polarização, identidades de, 239
Polinomial, função, 33, 126
Polinômio, 117
característico, 178
decomposição de polinômio em fatôres primos, 134
minimal, 181
primo (irredutível), 132
raiz de, 128
Positivo (a),
matriz, 263
operador, 260
Pôsto:
— coluna, 112
de forma bilinear, 312
de matriz, 112
— determinante, 164
de transformação linear, 71
— linha, 57, 112
Primária, teorema de decomposição, 194
Primo, polinômio, 132
Produto:
de matrizes, 18
de permutações, 154
de transformações lineares, 76
Projeção, 173
ortogonal, 249

Q

Quadrática, forma:
- de forma bilinear, 315
- de produto interno, 238

Quociente, espaço, 346

R

Racional de uma matriz, forma, 214
Racional, teorema da decomposição, 211
Reunião, 336
Rígido, movimento, 281
Rotação, 277

S

Semelhantes, matrizes, 94
Semi-simples, operador, 227
Simétrica (o),
- forma bilinear, 314
- grupo, 153
- matriz, 38, 285

Subespaço (s), 37
- anulador de, 104
- cíclico, 201
- disjuntos, 166
- gerado por, 39
- independentes, 167
- soma de, 40
- suplementar, 206
- suplementar ortogonal de, 247
- *T*-admissível, 207

T

Taylor, fórmula de, 230
Traço:
- de matriz, 98
- de transformação linear, 103

Transposta:
- de matriz, 111
- de transformação linear, 110

Triangular, matriz, 155

U

Unidade, matriz, 10
Unitário (a, as),
- equivalência de matrizes, 278
- espaço, 243
- grupo, 275
- matriz, 275
- operador, 275
- polinômio, 118

V

Vandermonde, matriz de, 123
Vetores-linhas, 41
Vetorial, espaço, 31
- base de, 44
- de funções polinomiais, 33
- de *n*-uplas, 31
- de soluções de equações lineares, 38
- dimensão de, 46
- quociente de, 346
- subespaço de, 37

SÍMBOLO S.A. INDÚSTRIAS GRÁFICAS
Rua General Flores, 518 522 525
Telefone 221 5833
São Paulo