QUARTA EDIÇÃO

ÁLGEBRA LINEAR COM APLICAÇÕES

Steven J. Leon
UNIVERSIDADE DE MASSACHUSETTS EM DARTMOUTH

Tradução

VALÉRIA DE MAGALHÃES IORIO
FACULDADE DE INFORMÁTICA DE TERESÓPOLIS
FUNDAÇÃO EDUCACIONAL SERRA DOS ÓRGÃOS

LTC
EDITORA

LTC — Livros Técnicos e Científicos Editora S. A.
Travessa do Ouvidor, 11
Rio de Janeiro, RJ — CEP 20040-040

Para Judith Russ Leon

Prefácio

Este livro é apropriado para disciplinas para alunos do segundo ano ou para disciplinas para alunos do terceiro e quarto anos. O estudante deve estar familiarizado com as noções básicas de cálculo diferencial e integral. Esse pré-requisito pode ser satisfeito por um semestre ou dois trimestres de cálculo elementar.

Se o livro for usado com alunos do segundo ano, então deve-se perder mais tempo, provavelmente, nos primeiros capítulos e omitir muitas das seções nos últimos capítulos. Para disciplinas mais avançadas, poder-se-ia rever rapidamente muitos dos tópicos nos dois primeiros capítulos e depois estudar de uma forma mais completa os capítulos restantes. As explicações no texto são dadas com detalhes suficientes de modo a permitir que alunos iniciantes leiam e entendam o material sem muita dificuldade. Para ajudar ainda mais o estudante, um grande número de exemplos foi resolvido detalhadamente. Além disso, são dados exercícios computacionais ao final de cada capítulo para dar aos alunos oportunidade de fazer experiências numéricas e tentar generalizar os resultados. As aplicações encontram-se espalhadas ao longo do texto. Estas aplicações podem ser usadas para se motivar material novo ou para ilustrar a relevância de material já apresentado.

O livro contém mais material do que o que pode ser dado em uma disciplina de um trimestre ou um semestre. O autor é de opinião de que é mais fácil para um professor pular ou mesmo não dar algum tópico do que suplementar um livro com material externo. Mesmo no caso da omissão de diversos tópicos, o livro ainda deve dar ao aluno uma idéia do escopo geral do assunto. Além disso, muitos dos alunos podem usar o livro mais tarde como referência e, portanto, podem acabar aprendendo muitos dos tópicos omitidos por conta própria.

O ideal seria estudar o livro inteiro em uma seqüência de dois trimestres ou dois semestres. Embora o grupo especial de estudos do currículo de álgebra linear (Linear Algebra Curriculum Study Group, LACSG), financiado pelo NSF*, tenha recomendado dois semestres de álgebra linear, muitas universidades e faculdades ainda não implementaram isso. Em uma seção posterior deste prefácio são dadas algumas sugestões de tópicos a serem cobertos em disciplinas de um semestre, tanto para alunos do segundo quanto do terceiro e quarto anos, e com uma ênfase matricial ou mais teórica. Para ajudar ainda mais o professor na escolha de tópicos, três seções foram designadas como opcionais e marcadas com um asterisco no sumário. Estas seções não são pré-requisito para nenhuma outra no livro. Elas podem ser omitidas sem qualquer quebra de continuidade.

Edições anteriores deste livro foram usadas em diversas universidades e faculdades para uma grande variedade de disciplinas de álgebra linear. Graças ao apoio e entusiasmo de muitos de seus usuários, o livro está agora em sua quarta edição. Ele continua a evoluir a cada nova edição. Embora o sucesso das edições anteriores indique que não há necessidade de se fazerem mudanças fundamentais, sempre existem seções e tópicos que podem ser aperfeiçoados e tornados mais claros. O autor leciona duas ou três disciplinas de álgebra linear por ano e está constantemente procurando maneiras de melhorar a apresentação do material. Autores de resenhas e usuários também contribuíram com muitas sugestões úteis. Em conseqüência, esta nova edição, ao mesmo tempo que mantém a essência das edições anteriores, incorpora uma série de melhoras substanciais.

---

*National Science Foundation, órgão do governo americano que financia pesquisas. (*N.T.*)

O QUE HÁ DE NOVO NA QUARTA EDIÇÃO?

1. Conjuntos de Exercícios Computacionais em Cada Capítulo

Os exercícios para serem feitos com o programa MATLAB foram bastante expandidos. A nova edição inclui uma seção de exercícios com o MATLAB ao final de cada capítulo. Os exercícios das seções variam de 2 a 7 páginas, dependendo do tamanho da seção. Os exercícios são todos cuidadosamente projetados para preencher uma série de objetivos pedagógicos. Eles envolvem muito mais do que simples cálculos mecânicos. Eles fazem com que os alunos efetuem cálculos e respondam perguntas sobre os resultados destes cálculos. As perguntas servem para clarear o significado matemático dos cálculos. Os alunos deveriam não apenas obter experiência com cálculos matriciais mas, também, ganhar novas perspectivas sobre o assunto.

2. Mais Motivação Geométrica

As sucessivas edições deste livro têm dado cada vez mais ênfase à geometria. Esta nova edição inclui ainda mais motivação geométrica adicional para alguns tópicos e nove figuras geométricas novas.

3. Nova Aplicação Envolvendo Teoria dos Grafos e Redes

Uma aplicação envolvendo teoria dos grafos e redes foi acrescentada na Seção 3 do Cap. 1 e nos exercícios com o MATLAB. Também foram acrescentados novos problemas no conjunto de exercícios do Cap. 1 e diversos exemplos resolvidos nesse capítulo foram revistos e melhorados.

4. Motivação Adicional para a Definição de Determinantes

Na primeira seção do Cap. 2 acrescentou-se mais material para uma motivação melhor da definição de determinante de uma matriz. Em conseqüência, a maior parte da seção foi reescrita. O determinante é definido agora como um número associado a uma matriz cujo valor indica se a matriz é ou não invertível. Antes de se considerar a definição geral de determinante de uma matriz A $n \times n$, são examinados os casos particulares $n = 1, 2, 3$. Em cada caso encontra-se uma condição para determinar se A é equivalente ou não à matriz identidade, baseada em se uma expressão envolvendo os coeficientes é diferente de zero. A definição geral é apresentada como uma generalização dessas expressões.

Também foram incluídos novos exercícios nas três seções do Cap. 2.

5. A Seção sobre Mudança de Base Foi Transferida para o Cap. 3

Nesta edição a seção "Mudança de Base" foi transferida do Cap. 4 para o Cap. 3. Muita coisa nessa seção foi reescrita. Os alunos devem achar a versão revista muito mais amigável. Além disso, quatro dos seis conjuntos de exercícios no Cap. 3 foram expandidos.

6. Revisões Importantes na Seção sobre Espaços Munidos de Produto Interno

A Seção "Espaços Munidos de Produto Interno" foi extensamente revista. Ela inclui uma demonstração diferente da desigualdade de Cauchy-Schwarz. A nova demonstração deve fazer mais sentido para os alunos do que a das edições anteriores. Na mesma seção, a definição de um produto interno no espaço vetorial $R^{m \times n}$ também é nova. A norma de Frobenius é definida, então, como a norma proveniente desse produto interno.

7. A Seção sobre Normas Matriciais Foi Transferida para o Cap. 7

A norma de Frobenius para matrizes é definida, agora, na Seção 3 do Cap. 5; o restante do material sobre normas matriciais na Seção 4 foi revisto e transferido para o Cap. 7. As normas matriciais estão contidas, agora, na Seção 4 do Cap. 7, que tem um novo título, "Normas Matriciais e Números Condicionais".

8. Nova Aplicação: Aproximação de Funções por Polinômios Trigonométricos

Foram feitas, também, muitas revisões na seção "Conjuntos Ortonormais" no Cap. 5. Foi incluída uma nova subseção mostrando como encontrar a melhor aproximação por mínimos quadráticos, por um polinômio trigonométrico de grau menor ou igual a n, para uma função em $C[a, b]$. Alguns dos exemplos nessa seção também foram revistos e foi acrescentado material novo sobre matrizes de projeção.

9. Revisões no Cap. 6

Foram incluídos novos exercícios na maioria das seções do Cap. 6. Foram colocados, também, novos exemplos.

10. A seção "Métodos Iterativos" foi retirada do Cap. 7 nesta edição. Nas edições anteriores, essa era uma seção opcional e desconfio de que era raramente usada em disciplinas de álgebra linear. Com todos os melhoramentos incluídos nesta edição, era preciso retirar algum material para manter o número de páginas (e o custo para o aluno) menor.

11. Ao preparar a quarta edição, o autor reviu cuidadosamente cada seção do livro. Além das mudanças maiores já listadas, inúmeras pequenas modificações foram feitas ao longo do texto.

EXERCÍCIOS COMPUTACIONAIS

Esta edição contém seções de exercícios computacionais ao final de cada capítulo. Estes exercícios são baseados no programa MATLAB. O apêndice sobre o MATLAB no livro explica o básico para usar o programa. O programa tem a vantagem de ser uma ferramenta poderosa para cálculos matriciais e ser fácil de aprender. Após a leitura do apêndice, os alunos deveriam ser capazes de fazer os exercícios computacionais sem recorrer a outro livro ou manual sobre o programa. Para ajudar os alunos a começar, recomendamos uma demonstração de 50 minutos do programa em sala de aula. Os exercícios podem ser feitos como exercícios para casa ou como parte de laboratórios computacionais com horário marcado.

Embora esse material possa ser ensinado sem referência alguma ao computador, acreditamos que exercícios computacionais podem melhorar muito o aprendizado do aluno e dar uma nova dimensão ao ensino da álgebra linear. Essa visão parece estar ganhando grande apoio na comunidade matemática em geral. O grupo de estudo sobre o currículo de álgebra linear recomendou o uso de tecnologia em uma primeira disciplina de álgebra linear. Nos encontros das três maiores sociedades matemáticas americanas existem, agora, sessões cujo assunto principal é o uso de computadores no ensino de álgebra linear. A National Science Foundation e a Sociedade Internacional de Álgebra Linear estão financiando um projeto chamado ATLAST (Augmenting the Teaching of Linear Algebra through the use of Software Tools*). O objetivo do projeto é estimular e facilitar o uso de programas de computador no ensino de álgebra linear. O ATLAST já fez dez encontros de professores usando o programa MATLAB. Participantes desses encontros estão projetando exercícios computacionais para disciplinas de álgebra linear e contribuindo, com estes exercícios, para o banco de dados do projeto. Os exercícios do banco de dados ATLAST vão ser reunidos em um livro que tem o título provisório *ATLAST Computer Exercises for Linear Algebra*. Os editores desse livro são Steven J. Leon, Richard Faulkenberry e Eugene Herman.

*Melhorando o Ensino de Álgebra Linear através de Ferramentas Computacionais; a sigla utilizada envolve uma brincadeira: *at last*, em inglês, significa finalmente. (*N.T.*)

SUGESTÕES DE ESCOLHA DE TÓPICOS

I. Seqüência de Dois Semestres

Em uma seqüência de dois semestres, é possível cobrir todas as 39 seções do livro. É possível uma flexibilidade adicional omitindo-se qualquer uma das três seções opcionais nos Caps. 2, 5 e 6. Pode-se incluir, também, uma aula extra demonstrando o uso do programa MATLAB.

II. Um Semestre para Alunos do Segundo Ano

A. Disciplina Básica para o Segundo Ano

Cap. 1	Seções 1 a 5		7 aulas
Cap. 2	Seções 1 e 2		2 aulas
Cap. 3	Seções 1 a 6		9 aulas
Cap. 4	Seções 1 a 3		4 aulas
Cap. 5	Seções 1 a 6		9 aulas
Cap. 6	Seções 1 a 3		4 aulas
		Total	35 aulas

B. Disciplina Sugerida por LACSG com Ênfase em Matrizes

A disciplina central recomendada pelo grupo de estudos do currículo de álgebra linear envolve apenas espaços vetoriais euclidianos. Portanto, para uma tal disciplina, deveriam ser omitidas a Seção 1 do Cap. 3 (sobre espaços vetoriais gerais) e todas as referências e exercícios envolvendo espaços de funções nos Caps. 3 a 6. Todos os tópicos na ementa nuclear de LACSG estão incluídos no texto. Não é necessário acrescentar nenhum material suplementar. O LACSG recomenda 28 aulas para ensinar o material central, mas o autor acha que talvez o esquema a seguir, com 35 aulas, seja mais razoável.

Cap. 1	Seções 1 a 5		7 aulas
Cap. 7	Seção 2 (fatoração *LU*)		1 aula
Cap. 2	Seções 1 a 3		3 aulas
Cap. 3	Seções 2 a 6		6 aulas
Cap. 4	Seção 1		1 aula
Cap. 5	Seções 1 a 6		9 aulas
Cap. 6	Seções 1 e 3 a 5		8 aulas
		Total	35 aulas

III. Um Semestre para Alunos de Terceiro ou Quarto Ano

O material para disciplinas de alunos que já terminaram o ciclo básico depende da formação dos alunos. A seguir, duas sugestões com 35 aulas cada.

A. Disciplina 1

Cap. 1	Seções 1 a 5	6 aulas
Cap. 2	Seções 1 e 2	2 aulas
Cap. 3	Seções 1 a 6	7 aulas
Cap. 5	Seções 1 a 6	9 aulas
Cap. 6	Seções 1 a 6	9 aulas
	Seção 7 caso haja tempo	
Cap. 7	Seção 6	2 aulas
	Parte da Seção 8 caso haja tempo	

B. Disciplina 2

Revisão dos Tópicos nos Caps. 1 a 3		5 aulas
Cap. 4	Seções 1 a 3	3 aulas
Cap. 5	Seções 1 a 6	9 aulas
Cap. 6	Seções 1 a 6 Seção 7 caso haja tempo	9 aulas
Cap. 7	Seções 4 a 8 Caso haja tempo, Seções 1 a 3	9 aulas

AGRADECIMENTOS

O autor gostaria de expressar sua gratidão à longa lista de autores de resenhas que tanto contribuíram para todas as quatro edições deste livro. Obrigado, também, aos diversos usuários que enviaram comentários e sugestões. Agradecimentos especiais para Wayne Barrett e Germund Dahlquist por suas sugestões para a segunda e terceira edições.

Um bom número das revisões e exercícios novos nesta última edição são conseqüência direta de comentários e sugestões de autores de resenhas: Timothy Hardy, Universidade de Iowa do Norte; Inessa Levi, Universidade de Louisville; Dennis McLaughlin, Universidade de Princeton; Hiram Paley, Universidade de Illinois em Urbana; Sandra Shields, Faculdades de William e Mary; Ilya Spitkovsky, Faculdades de William e Mary; Mo Tavakoli, Faculdades Comunitárias de Chaffey; e Santiago Tavares, Universidade da Flórida em Gainesville.

O autor gostaria, também, de agradecer a uma série de pessoas que ajudaram a dar forma a esta edição. Agradecemos a Cleve Moler pela sugestão de dois dos exercícios com o MATLAB. Agradecemos, também, a Roger Horn e a Kermit Sigmon por suas sugestões e, especialmente, a Philip Bacon por comentários detalhados em muitas das seções da terceira edição. A comunidade matemática sofreu uma grande perda com a morte de Philip em novembro de 1991. Sua falta tem sido muito sentida por estudantes e professores da instituição onde trabalhava, a Universidade da Flórida, e por seus muitos amigos.

Agradecemos a Judith Russ Leon e a Ann Cox por terem revisto, independentemente, o manuscrito para a quarta edição. Devemos agradecer a Ann Cox também por resolver os exercícios e verificar as respostas no final do livro.

As revisões finais do manuscrito foram feitas enquanto o autor estava em licença da universidade visitando o Instituto Federal de Tecnologia Suíço (ETH) e a Universidade de Stanford. O autor gostaria de agradecer seus anfitriões Walter Gander e Gene Golub por tornarem essas visitas possíveis. Em particular, o autor gostou muito de ter tido a oportunidade de usar estações Sun para preparar o manuscrito para esta edição.

Agradecemos ao editor de matemática, Bob Pirtle, e ao restante da equipe editorial, de produção e de vendas da Macmillan College Publishing Company pelo trabalho em todas as quatro edições.

Finalmente, o autor gostaria de agradecer as contribuições de Gene Golub e Jim Wilkinson. A maior parte da primeira edição deste livro foi escrita em 1977-1978, enquanto o autor era um Professor Visitante na Universidade de Stanford. Durante esse período, o autor assistiu a disciplinas e conferências sobre álgebra linear numérica dadas por Gene Golub e J. H. Wilkinson. Essas conferências muito influenciaram este livro.

S. L.

Sumário

*O asterisco indica seções opcionais. Veja a primeira seção do prefácio para uma explicação.

Álgebra Linear com Aplicações

CAPÍTULO 1

MATRIZES E SISTEMAS DE EQUAÇÕES

Provavelmente o problema mais importante em matemática é resolver um sistema de equações lineares. Mais de 75% de todos os problemas matemáticos encontrados em aplicações científicas e industriais envolvem a resolução de um sistema linear em alguma etapa. Usando os métodos da matemática moderna, muitas vezes é possível reduzir um problema sofisticado a um único sistema de equações lineares. Sistemas lineares aparecem em aplicações em áreas como administração, economia, sociologia, ecologia, demografia, genética, eletrônica, engenharia e física. Parece, portanto, apropriado começar este livro com uma seção sobre sistemas lineares.

1 SISTEMAS DE EQUAÇÕES LINEARES

Uma *equação linear em n incógnitas* é uma equação da forma

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

onde $a_1, a_2, \ldots, a_n$ e b são números reais e $x_1, x_2, \ldots, x_n$ são as variáveis. Um *sistema linear* de m equações em n incógnitas é, então, um sistema da forma

$$\begin{aligned}
a_{11}x_1 &+ a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\
a_{21}x_1 &+ a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\
&\vdots \\
a_{m1}x_1 &+ a_{m2}x_2 + \cdots + a_{mn}x_n = b_m
\end{aligned}$$

onde os a_{ij} e os b_i são números reais. Vamos nos referir a sistemas da forma (1) como sistemas lineares $m \times n$. Damos, a seguir, alguns exemplos de sistemas lineares:

(a) $\begin{aligned} x_1 + 2x_2 &= 5 \\ 2x_1 + 3x_2 &= 8 \end{aligned}$ (b) $\begin{aligned} x_1 - x_2 + x_3 &= 2 \\ 2x_1 + x_2 - x_3 &= 4 \end{aligned}$ (c) $\begin{aligned} x_1 + x_2 &= 2 \\ x_1 - x_2 &= 1 \\ x_1 \quad\; &= 4 \end{aligned}$

O sistema (a) é um sistema 2×2, (b) é um sistema 2×3 e (c) é um sistema 3×2.

Entendemos por solução de um sistema $m \times n$ uma n-upla ordenada de números $(x_1, x_2, \ldots, x_n)$ que satisfaz todas as equações do sistema. Por exemplo, o par ordenado $(1, 2)$ é uma solução do sistema (a), já que

$$1 \cdot (1) + 2 \cdot (2) = 5$$
$$2 \cdot (1) + 3 \cdot (2) = 8$$

A tripla ordenada (2, 0, 0) é uma solução do sistema (b), pois

$$1 \cdot (2) - 1 \cdot (0) + 1 \cdot (0) = 2$$
$$2 \cdot (2) + 1 \cdot (0) - 1 \cdot (0) = 4$$

De fato, o sistema (b) tem muitas soluções. Se α é um número real qualquer, é fácil ver que a tripla ordenada $(2, \alpha, \alpha)$ é uma solução. Entretanto, o sistema (c) não tem nenhuma solução. A partir da terceira equação, vemos que a primeira coordenada de qualquer solução tem que ser 4. Usando $x_1 = 4$ nas duas primeiras equações, vemos que a segunda coordenada tem que satisfazer

$$4 + x_2 = 2$$
$$4 - x_2 = 1$$

Como não existe número real satisfazendo ambas as equações, o sistema não tem solução. Se um sistema linear não tem solução, dizemos que ele é *incompatível* ou *impossível.* Logo, o sistema (c) é incompatível, enquanto os sistemas (a) e (b) são ambos possíveis (compatíveis).

O conjunto de todas as soluções de um sistema linear é chamado de *conjunto solução* do sistema. Se um sistema é impossível, seu conjunto solução é vazio. Um sistema compatível vai ter um conjunto solução não-vazio. Para resolver um sistema possível, é preciso encontrar seu conjunto solução.

SISTEMAS 2 × 2

Vamos examinar, do ponto de vista geométrico, um sistema da forma

$$a_{11}x_1 + a_{12}x_2 = b_1$$
$$a_{21}x_1 + a_{22}x_2 = b_2$$

Cada uma dessas equações pode ser representada graficamente por uma reta no plano. O par ordenado (x_1, x_2) vai ser uma solução do sistema se e somente se pertencer a ambas as retas. Por exemplo, considere os três sistemas a seguir:

$$\text{(i)}\ \begin{aligned} x_1 + x_2 &= 2 \\ x_1 - x_2 &= 2 \end{aligned} \qquad \text{(ii)}\ \begin{aligned} x_1 + x_2 &= 2 \\ x_1 + x_2 &= 1 \end{aligned} \qquad \text{(iii)}\ \begin{aligned} x_1 + x_2 &= 2 \\ -x_1 - x_2 &= -2 \end{aligned}$$

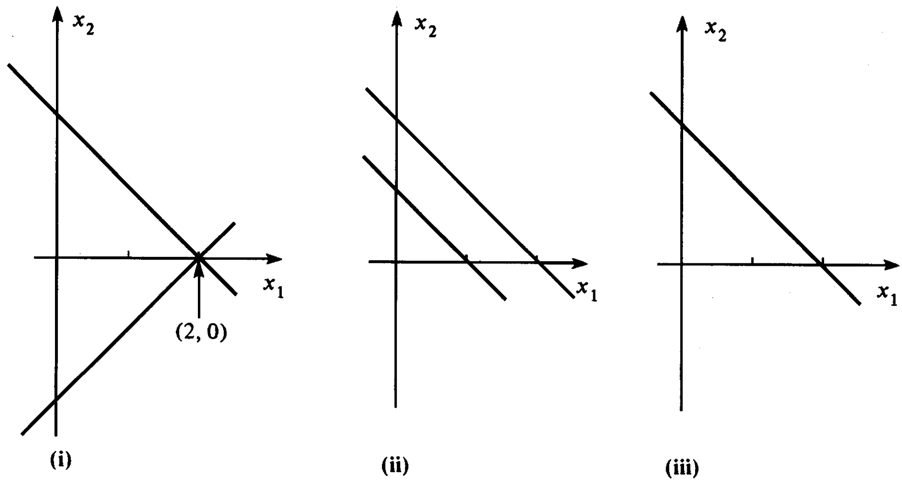


FIG. 1.1.1

As duas retas no sistema (i) se interceptam no ponto (2, 0). Logo, {(2, 0)} é o conjunto solução de (i). No sistema (ii), as duas retas são paralelas, logo, o sistema é incompatível e seu conjunto solução é vazio. As duas equações no sistema (iii) representam a mesma reta; qualquer ponto nessa reta vai ser uma solução do sistema (ver Fig. 1.1.1).

Em geral, existem três possibilidades: as retas se interceptam em um ponto, são paralelas, ou ambas as equações representam a mesma reta. Então, o conjunto solução contém, respectivamente, um zero ou um número infinito de pontos.

A situação é semelhante para sistemas $m \times n$. Um sistema $m \times n$ pode ou não ser compatível. Se for compatível, ele tem que ter exatamente uma solução ou um número infinito de soluções. Essas são as únicas possibilidades. Vamos ver por que isso é assim na Seção 2, quando estudarmos a forma escada. De interesse mais imediato é encontrar todas as soluções de um dado sistema. Para atacar esse problema, vamos definir a noção de *sistemas equivalentes*.

SISTEMAS EQUIVALENTES

Considere os dois sistemas

$$\text{(a)}\quad \begin{aligned} 3x_1 + 2x_2 - x_3 &= -2 \\ x_2 \phantom{{}- x_3} &= 3 \\ 2x_3 &= 4 \end{aligned} \qquad\qquad \text{(b)}\quad \begin{aligned} 3x_1 + 2x_2 - x_3 &= -2 \\ -3x_1 - x_2 + x_3 &= 5 \\ 3x_1 + 2x_2 + x_3 &= 2 \end{aligned}$$

O sistema (a) é fácil de resolver, já que é claro, das duas últimas equações, que $x_2 = 3$ e $x_3 = 2$. Usando esses valores na primeira equação, obtemos

$$\begin{aligned} 3x_1 + 2 \cdot 3 - 2 &= -2 \\ x_1 &= -2 \end{aligned}$$

Logo, a solução do sistema é $(-2, 3, 2)$. O sistema (b) parece ser mais difícil de resolver. De fato, o sistema (b) tem a mesma solução que o sistema (a). Para ver isto, some as duas primeiras equações do sistema:

$$\begin{array}{r} 3x_1 + 2x_2 - x_3 = -2 \\ \underline{-3x_1 - x_2 + x_3 = 5} \\ x_2 \phantom{{}+ x_3} = 3 \end{array}$$

Se (x_1, x_2, x_3) é uma solução arbitrária de (b), ela tem que satisfazer todas as equações do sistema, logo, tem que satisfazer qualquer equação obtida somando-se duas de suas equações. Portanto, x_2 tem que ser igual a 3. Analogamente, (x_1, x_2, x_3) tem que satisfazer a nova equação obtida subtraindo-se a primeira equação da terceira:

$$\begin{array}{r} 3x_1 + 2x_2 + x_3 = 2 \\ \underline{3x_1 + 2x_2 - x_3 = -2} \\ 2x_3 = 4 \end{array}$$

Então, qualquer solução do sistema (b) tem, também, que ser uma solução do sistema (a). Por um argumento análogo, pode-se mostrar que qualquer solução de (a) é, também, uma solução de (b). Isso pode ser feito subtraindo-se a primeira equação da segunda:

$$\begin{array}{r} x_2 \phantom{{}+ x_3} = 3 \\ \underline{3x_1 + 2x_2 - x_3 = -2} \\ -3x_1 - x_2 + x_3 = 5 \end{array}$$

e somando-se a primeira e terceira equações:

$$\begin{array}{rcrcrcr} 3x_1 & + & 2x_2 & - & x_3 & = & -2 \\ & & & & 2x_3 & = & 4 \\ \hline 3x_1 & + & 2x_2 & + & x_3 & = & 2 \end{array}$$

Logo, (x_1, x_2, x_3) é uma solução do sistema (b) se e somente se é uma solução do sistema (a). Portanto, ambos os sistemas têm o mesmo conjunto solução, $\{(-2, 3, 2)\}$.

Definição. Dois sistemas de equações envolvendo as mesmas variáveis são ditos **equivalentes** se têm o mesmo conjunto solução.

É claro que, se trocarmos a ordem em que escrevemos duas equações de um sistema, isso não vai afetar o conjunto solução. O sistema reordenado será equivalente ao sistema original. Por exemplo, os sistemas

$$\begin{array}{rcrcr} x_1 & + & 2x_2 & = & 4 \\ 3x_1 & - & x_2 & = & 2 \\ 4x_1 & + & x_2 & = & 6 \end{array} \qquad \text{e} \qquad \begin{array}{rcrcr} 4x_1 & + & x_2 & = & 6 \\ 3x_1 & - & x_2 & = & 2 \\ x_1 & + & 2x_2 & = & 4 \end{array}$$

envolvem, ambos, as mesmas equações e, portanto, têm que ter o mesmo conjunto solução.

Se uma das equações de um sistema é multiplicada por um número real não-nulo, isso não afetará o conjunto solução, e o novo sistema será equivalente ao sistema original. Por exemplo, os sistemas

$$\begin{array}{rcrcrcr} x_1 & + & x_2 & + & x_3 & = & 3 \\ -2x_1 & - & x_2 & + & 4x_3 & = & 1 \end{array} \qquad \text{e} \qquad \begin{array}{rcrcrcr} 2x_1 & + & 2x_2 & + & 2x_3 & = & 6 \\ -2x_1 & - & x_2 & + & 4x_3 & = & 1 \end{array}$$

são equivalentes.

Se um múltiplo de uma equação é somado a outra equação, o novo sistema será equivalente ao sistema original. Isso acontece porque a n-upla $(x_1, x_2, \ldots, x_n)$ satisfaz as duas equações

$$a_{i1}x_1 + \cdots + a_{in}x_n = b_i$$
$$a_{j1}x_1 + \cdots + a_{jn}x_n = b_j$$

se e somente se satisfaz as equações

$$a_{i1}x_1 + \cdots + a_{in}x_n = b_i$$
$$(a_{j1} + \alpha a_{i1})x_1 + \cdots + (a_{jn} + \alpha a_{in})x_n = b_j + \alpha b_i$$

Resumindo, existem três operações que podem ser efetuadas em um sistema para se obter um sistema equivalente:

(a) A ordem em que duas equações são escritas pode ser trocada.
(b) Os dois lados de uma equação podem ser multiplicados pelo mesmo número real diferente de zero.
(c) Um múltiplo de uma equação pode ser somado a outro.

Dado um sistema de equações, podemos usar essas operações para obter um sistema equivalente que seja mais fácil de resolver.

SISTEMAS $n \times n$

Vamos nos restringir a sistemas $n \times n$ até o final desta seção. Vamos mostrar que, se um sistema $n \times n$ tem exatamente uma solução, então as operações (a) e (c) podem ser usadas para se obter um sistema equivalente "triangular".

Definição. Um sistema está em **forma triangular** se, na k-ésima equação, os coeficientes das $k-1$ primeiras variáveis são todos nulos e o coeficiente de x_k é diferente de zero ($k = 1, \ldots, n$).

EXEMPLO 1. O sistema

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

está em forma triangular, já que, na segunda equação, os coeficientes são 0, 1, -1, respectivamente, e na terceira os coeficientes são 0, 0, 2, respectivamente. Devido à sua forma triangular, o sistema é fácil de resolver. Da terceira equação, temos que $x_3 = 2$. Usando esse valor na segunda equação, obtemos

$$x_2 - 2 = 2 \qquad \text{ou} \qquad x_2 = 4$$

Usando $x_2 = 4$ e $x_3 = 2$ na primeira equação, terminamos com

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

Logo, a solução do sistema é $(-3, 4, 2)$. □

Qualquer sistema triangular $n \times n$ pode ser resolvido da mesma maneira que o exemplo anterior. Primeiro, resolve-se a n-ésima equação para x_n. Esse valor é usado na $(n-1)$-ésima equação para encontrar x_{n-1}. Os valores de x_n e x_{n-1} são usados na $(n-2)$-ésima equação para encontrar x_{n-2}, e assim por diante. Vamos nos referir a esse método de resolver um sistema triangular como *substituição de baixo para cima* ou, simplesmente, *substituição*.

EXEMPLO 2. Resolva o sistema

$$\begin{aligned} 2x_1 - x_2 + 3x_3 - 2x_4 &= 1 \\ x_2 - 2x_3 + 3x_4 &= 2 \\ 4x_3 + 3x_4 &= 3 \\ 4x_4 &= 4 \end{aligned}$$

SOLUÇÃO. Usando substituição, obtemos

$$\begin{aligned} 4x_4 &= 4 & x_4 &= 1 \\ 4x_3 + 3 \cdot 1 &= 3 & x_3 &= 0 \\ x_2 - 2 \cdot 0 + 3 \cdot 1 &= 2 & x_2 &= -1 \\ 2x_1 - (-1) + 3 \cdot 0 - 2 \cdot 1 &= 1 & x_1 &= 1 \end{aligned}$$

Logo, a solução é $(1, -1, 0, 1)$. □

Se o sistema de equações não for triangular, usaremos as operações (a) e (c) para tentar obter um sistema equivalente em forma triangular.

EXEMPLO 3. Resolva o sistema

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

SOLUÇÃO. Subtraindo 3 vezes a primeira linha da segunda, obtemos

$$-7x_2 - 6x_3 = -10$$

Subtraindo 2 vezes a primeira linha da terceira, obtemos

$$-x_2 - x_3 = -2$$

Se trocamos as segunda e terceira equações de nosso sistema, respectivamente, por essas novas equações, obtemos o sistema equivalente

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 3 \\ -7x_2 - 6x_3 &= -10 \\ -x_2 - x_3 &= -2 \end{aligned}$$

Se a terceira equação desse sistema é trocada por sua soma com $-1/7$ da segunda, acabamos com o seguinte sistema triangular:

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 3 \\ -7x_2 - 6x_3 &= -10 \\ -\tfrac{1}{7}x_3 &= -\tfrac{4}{7} \end{aligned}$$

Usando substituição, obtemos

$$x_3 = 4, \qquad x_2 = -2, \qquad x_1 = 3$$ □

Vamos olhar de novo, para o sistema de equações no último exemplo. Podemos associar àquele sistema um arranjo 3×3 de números cujos elementos são os coeficientes das incógnitas x_i.

$$\begin{pmatrix} 1 & 2 & 1 \\ 3 & -1 & -3 \\ 2 & 3 & 1 \end{pmatrix}$$

Vamos nos referir a esse arranjo como a *matriz de coeficientes* do sistema. O termo *matriz* significa, simplesmente, um arranjo retangular de números. Uma matriz com m linhas e n colunas é uma matriz $m \times n$.

Se agregarmos à matriz de coeficientes uma coluna adicional cujos elementos são os números que aparecem do lado direito dos sinais de igualdade no sistema, obteremos a nova matriz

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 3 & -1 & -3 & -1 \\ 2 & 3 & 1 & 4 \end{array}\right)$$

Essa nova matriz será chamada de *matriz aumentada*. Em geral, quando uma matriz B $m \times r$ é agregada a uma matriz A $m \times n$, a matriz obtida é denotada por $(A|B)$. Por exemplo, se

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1r} \\ b_{21} & b_{22} & \cdots & b_{2r} \\ \vdots & & & \\ b_{m1} & b_{m2} & \cdots & b_{mr} \end{pmatrix}$$

então

$$(A|B) = \left(\begin{array}{ccc|ccc} a_{11} & \cdots & a_{1n} & b_{11} & \cdots & b_{1r} \\ \vdots & & & \vdots & & \\ a_{m1} & \cdots & a_{mn} & b_{m1} & \cdots & b_{mr} \end{array}\right)$$

A cada sistema de equações podemos associar uma matriz aumentada da forma

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right)$$

O sistema pode ser resolvido efetuando-se as operações na matriz aumentada. Os nomes das variáveis, x_i, podem ser omitidos até o final dos cálculos. As operações a seguir, efetuadas nas linhas da matriz aumentada, correspondem às três operações usadas para se obter um sistema equivalente.

Operações Elementares sobre as Linhas

I. Trocar duas linhas.
II. Multiplicar uma linha por um número real não-nulo.
III. Substituir uma linha por sua soma com um múltiplo de outra linha.

Voltando ao exemplo, vemos que a primeira linha é usada para anular os elementos na primeira coluna das linhas restantes. Vamos nos referir a essa primeira linha como *linha do pivô*,* e ao elemento 1 com um círculo em volta na primeira linha como *pivô*.

$$\begin{array}{r} \text{pivô} \rightarrow \\ \left.\begin{array}{r}\text{elementos a serem}\\ \text{eliminados}\end{array}\right\} \rightarrow \end{array} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ \boxed{3} & -1 & -3 & -1 \\ \boxed{2} & 3 & 1 & 4 \end{array}\right) \leftarrow \text{linha do pivô}$$

Usando a operação elementar III, subtraímos 3 vezes a segunda linha da primeira e subtraímos 2 vezes a primeira linha da segunda. Ao final, obtemos a matriz

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & -7 & -6 & -10 \\ 0 & \boxed{-1} & -1 & -2 \end{array}\right) \leftarrow \text{linha do pivô}$$

Vamos escolher agora a segunda linha como nossa nova linha do pivô e aplicar a operação elementar III para eliminar o último elemento da segunda coluna. Terminamos com a matriz

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & -7 & -6 & -10 \\ 0 & 0 & -\frac{1}{7} & -\frac{4}{7} \end{array}\right)$$

Essa é a matriz aumentada de um sistema triangular equivalente ao sistema original.

EXEMPLO 4. Resolva o sistema

$$\begin{aligned} - x_2 - x_3 + x_4 &= 0 \\ x_1 + x_2 + x_3 + x_4 &= 6 \\ 2x_1 + 4x_2 + x_3 - 2x_4 &= -1 \\ 3x_1 + x_2 - 2x_3 + 2x_4 &= 3 \end{aligned}$$

SOLUÇÃO. A matriz aumentada desse sistema é

$$\left(\begin{array}{cccc|c} 0 & -1 & -1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 6 \\ 2 & 4 & 1 & -2 & -1 \\ 3 & 1 & -2 & 2 & 3 \end{array}\right)$$

* Essa terminologia não é padrão. (*N. T.*)

Como não é possível anular qualquer elemento usando 0 como pivô, vamos usar a operação elementar I para trocar as duas primeiras linhas da matriz aumentada. A nova primeira linha será a linha do pivô e o elemento pivô será 1.

$$
\text{elemento pivô} \rightarrow \left(\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 6 \\ 0 & -1 & -1 & 1 & 0 \\ \boxed{2} & 4 & 1 & -2 & -1 \\ \boxed{3} & 1 & -2 & 2 & 3 \end{array}\right) \leftarrow \text{linha do pivô}
$$

Agora usamos duas vezes a operação elementar III para anular os dois elementos não-nulos indicados na primeira coluna.

$$
\left(\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 6 \\ 0 & -1 & -1 & 1 & 0 \\ 0 & \boxed{2} & -1 & -4 & -13 \\ 0 & \boxed{-2} & -5 & -1 & -15 \end{array}\right)
$$

A seguir, a segunda linha é usada como linha do pivô para anular os elementos na segunda coluna abaixo do elemento pivô -1.

$$
\left(\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 6 \\ 0 & -1 & -1 & 1 & 0 \\ 0 & 0 & -3 & -2 & -13 \\ 0 & 0 & \boxed{-3} & -3 & -15 \end{array}\right)
$$

Finalmente, a terceira linha é usada como linha do pivô para anular o último elemento na terceira coluna.

$$
\left(\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 6 \\ 0 & -1 & -1 & 1 & 0 \\ 0 & 0 & -3 & -2 & -13 \\ 0 & 0 & 0 & -1 & -2 \end{array}\right)
$$

Essa matriz aumentada representa um sistema triangular. Resolvendo por substituição, obtemos a solução $(2, -1, 3, 2)$. □

Em geral, se um sistema linear $n \times n$ puder ser reduzido a uma forma triangular, então ele terá uma única solução que pode ser obtida por substituição. Podemos pensar no processo de redução como um algoritmo envolvendo $n - 1$ passos. No primeiro passo, escolhemos um elemento pivô entre os elementos não-nulos da primeira coluna da matriz. A linha que contém o elemento pivô é a *linha do pivô*. Trocamos linhas (se necessário) de modo que a linha do pivô seja a primeira linha. Subtraímos, então, múltiplos da linha do pivô de cada uma das $n - 1$ linhas restantes de modo a obter 0 nas posições $(2, 1)$, $\ldots$, $(n, 1)$. No segundo passo, escolhemos um elemento pivô entre os elementos não-nulos da segunda coluna nas linhas de 2 a n da matriz. A linha contendo o pivô é, então, trocada com a segunda linha da matriz e usada como nova linha do pivô. Subtraímos, então, múltiplos da linha do pivô das $n - 2$ linhas restantes de modo a anular todos os elementos da segunda coluna abaixo do pivô. Repetimos o mesmo procedimento para as colunas de 3 a $n - 1$. Observe que, no segundo passo, a linha 1 e a coluna 1 não são modificadas, no segundo passo as duas primeiras linhas e as duas primeiras colunas não são modificadas e assim por diante. Em cada etapa, as dimensões globais do sistema são reduzidas de 1 (ver Fig. 1.1.2).

Se o processo de redução puder ser feito como descrito acima, chegaremos a um sistema equivalente triangular superior após $n - 1$ passos. No entanto, o procedimento não funciona se, em qualquer etapa, todas as escolhas possíveis para um elemento pivô forem iguais a 0. Quando isso acontecer, vamos ter que reduzir o sistema a um tipo particular de forma escada. Essas formas serão estudadas na próxima seção. Elas também serão usadas para sistemas $m \times n$, onde $m \neq n$.

$n = 4$

$$\text{Passo 1}\quad \left[\begin{array}{cccc|c} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \end{array}\right]$$

$$\text{Passo 2}\quad \left[\begin{array}{cccc|c} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & x & x & x \end{array}\right]$$

$$\text{Passo 3}\quad \left[\begin{array}{cccc|c} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & x & x & x \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & 0 & x & x \end{array}\right]$$

FIG. 1.1.2

EXERCÍCIOS

1. Use substituição para resolver cada um dos sistemas de equações a seguir.

(a)
$$\begin{aligned} x_1 - 3x_2 &= 2 \\ 2x_2 &= 6 \end{aligned}$$

(b)
$$\begin{aligned} x_1 + x_2 + x_3 &= 8 \\ 2x_2 + x_3 &= 5 \\ 3x_3 &= 9 \end{aligned}$$

(c)
$$\begin{aligned} x_1 + 2x_2 + 2x_3 + x_4 &= 5 \\ 3x_2 + x_3 - 2x_4 &= 1 \\ -x_3 + 2x_4 &= -1 \\ 4x_4 &= 4 \end{aligned}$$

(d)
$$\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 &= 5 \\ 2x_2 + x_3 - 2x_4 + x_5 &= 1 \\ 4x_3 + x_4 - 2x_5 &= 1 \\ x_4 - 3x_5 &= 0 \\ 2x_5 &= 2 \end{aligned}$$

2. Escreva a matriz dos coeficientes de cada um dos sistemas no Exercício 1.

3. Para cada um dos sistemas a seguir, interprete cada equação como uma reta no plano, faça o gráfico dessas retas e determine geometricamente o número de soluções.

(a)
$$\begin{aligned} x_1 + x_2 &= 4 \\ x_1 - x_2 &= 2 \end{aligned}$$

(b)
$$\begin{aligned} x_1 + 2x_2 &= 4 \\ -2x_1 - 4x_2 &= 4 \end{aligned}$$

(c)
$$\begin{aligned} 2x_1 - x_2 &= 3 \\ -4x_1 + 2x_2 &= -6 \end{aligned}$$

(d)
$$\begin{aligned} x_1 + x_2 &= 1 \\ x_1 - x_2 &= 1 \\ -x_1 + 3x_2 &= 3 \end{aligned}$$

4. Escreva a matriz aumentada de cada um dos sistemas no Exercício 3.

5. Escreva por extenso o sistema de equações que corresponde a cada uma das matrizes aumentadas a seguir.

(a) $\left(\begin{array}{cc|c} 3 & 2 & 8 \\ 1 & 5 & 7 \end{array}\right)$ (b) $\left(\begin{array}{ccc|c} 5 & -2 & 1 & 3 \\ 2 & 3 & -4 & 0 \end{array}\right)$

(c) $\left(\begin{array}{ccc|c} 2 & 1 & 4 & -1 \\ 4 & -2 & 3 & 4 \\ 5 & 2 & 6 & -1 \end{array}\right)$ (d) $\left(\begin{array}{cccc|c} 4 & -3 & 1 & 2 & 4 \\ 3 & 1 & -5 & 6 & 5 \\ 1 & 1 & 2 & 4 & 8 \\ 5 & 1 & 3 & -2 & 7 \end{array}\right)$

6. Resolva cada um dos sistemas a seguir.

(a) $\begin{aligned} x_1 - 2x_2 &= 5 \\ 3x_1 + x_2 &= 1 \end{aligned}$ (b) $\begin{aligned} 2x_1 + x_2 &= 8 \\ 4x_1 - 3x_2 &= 6 \end{aligned}$

(c) $\begin{aligned} 4x_1 + 3x_2 &= 4 \\ \tfrac{2}{3}x_1 + 4x_2 &= 3 \end{aligned}$ (d) $\begin{aligned} x_1 + 2x_2 - x_3 &= 1 \\ 2x_1 - x_2 + x_3 &= 3 \\ -x_1 + 2x_2 + 3x_3 &= 7 \end{aligned}$

(e) $\begin{aligned} 2x_1 + x_2 + 3x_3 &= 1 \\ 4x_1 + 3x_2 + 5x_3 &= 1 \\ 6x_1 + 5x_2 + 5x_3 &= -3 \end{aligned}$ (f) $\begin{aligned} 3x_1 + 2x_2 + x_3 &= 0 \\ -2x_1 + x_2 - x_3 &= 2 \\ 2x_1 - x_2 + 2x_3 &= -1 \end{aligned}$

(g) $\begin{aligned} \tfrac{1}{3}x_1 + \tfrac{2}{3}x_2 + 2x_3 &= -1 \\ x_1 + 2x_2 + \tfrac{3}{2}x_3 &= \tfrac{3}{2} \\ \tfrac{1}{2}x_1 + 2x_2 + \tfrac{12}{5}x_3 &= \tfrac{1}{10} \end{aligned}$

(h) $\begin{aligned} x_2 + x_3 + x_4 &= 0 \\ 3x_1 + 3x_3 - 4x_4 &= 7 \\ x_1 + x_2 + x_3 + 2x_4 &= 6 \\ 2x_1 + 3x_2 + x_3 + 3x_4 &= 6 \end{aligned}$

7. Os dois sistemas

(a) $\begin{aligned} 2x_1 + x_2 &= 3 \\ 4x_1 + 3x_2 &= 5 \end{aligned}$ (b) $\begin{aligned} 2x_1 + x_2 &= -1 \\ 4x_1 + 3x_2 &= 1 \end{aligned}$

têm a mesma matriz de coeficientes, mas números diferentes à direita dos sinais de igualdade. Resolva ambos os sistemas simultaneamente anulando o elemento (2, 1) da matriz aumentada

$$\left(\begin{array}{cc|cc} 2 & 1 & 3 & -1 \\ 4 & 3 & 5 & 1 \end{array}\right)$$

e depois usando substituição para cada uma das colunas correspondentes aos números à direita dos sinais de igualdade.

8. Resolva os dois sistemas

(a) $\begin{aligned} x_1 + 2x_2 - 2x_3 &= 1 \\ 2x_1 + 5x_2 + x_3 &= 9 \\ x_1 + 3x_2 + 4x_3 &= 9 \end{aligned}$ (b) $\begin{aligned} x_1 + 2x_2 - 2x_3 &= 9 \\ 2x_1 + 5x_2 + x_3 &= 9 \\ x_1 + 3x_2 + 4x_3 &= -2 \end{aligned}$

usando operações elementares em uma matriz aumentada 3×5 e depois usando substituição.

9. Considere um sistema da forma

$$\begin{aligned} -m_1x_1 + x_2 &= b_1 \\ -m_2x_1 + x_2 &= b_2 \end{aligned}$$

onde m_1, m_2, b_1 e b_2 são constantes.
(a) Mostre que o sistema tem uma única solução se $m_1 \neq m_2$.
(b) Se $m_1 = m_2$, mostre que o sistema só é compatível se $b_1 = b_2$.
(c) Interprete geometricamente os itens (a) e (b).

10. Considere um sistema da forma

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= 0 \\ a_{21}x_1 + a_{22}x_2 &= 0 \end{aligned}$$

onde a_{11}, a_{12}, a_{21} e a_{22} são constantes. Explique por que um sistema dessa forma tem que ser compatível.

11. Dê uma interpretação geométrica de uma equação linear em três incógnitas. Descreva geometricamente todos os possíveis conjuntos solução para um sistema linear 3×3.

2 FORMA ESCADA

Na Seção 1 aprendemos um método para reduzir um sistema linear $n \times n$ a uma forma triangular. No entanto, esse método falha se, em qualquer etapa do processo de redução, todas as escolhas possíveis para o elemento pivô em uma dada coluna são nulas.

EXEMPLO 1. Considere o sistema representado pela matriz aumentada

$$\begin{array}{cl} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 & -1 \\ -2 & -2 & 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 1 & 3 & -1 \\ 1 & 1 & 2 & 2 & 4 & 1 \end{array}\right) & \begin{array}{l} \leftarrow \text{linha do pivô} \\ \\ \\ \\ \\ \end{array} \end{array}$$

Se a operação elementar III for usada para anular os últimos quatro elementos da primeira coluna, obteremos a matriz

$$\begin{array}{cl} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 2 & 2 & 5 & 3 \\ 0 & 0 & 1 & 1 & 3 & -1 \\ 0 & 0 & 1 & 1 & 3 & 0 \end{array}\right) & \begin{array}{l} \\ \leftarrow \text{linha do pivô} \\ \\ \\ \\ \end{array} \end{array}$$

Nesse estágio a redução a uma forma triangular não pode continuar. Todas as escolhas possíveis para o elemento pivô na segunda coluna são iguais a 0. Como continuar? Como nosso objetivo é simplificar o sistema ao máximo, parece natural passar para a terceira coluna e anular os três últimos elementos.

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right)$$

Na quarta coluna todas as escolhas possíveis para o pivô são iguais a zero; logo, novamente passamos para a próxima coluna. Usando a terceira linha como linha do pivô, anulamos os dois últimos elementos da quinta coluna.

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 & -3 \end{array}\right)$$

As equações representadas pelas duas últimas linhas são

$$0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 = -4$$
$$0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 = -3$$

Como não existem quíntuplas que possam satisfazer essas equações, o sistema é impossível. Note que terminamos com uma matriz de coeficientes que não está em forma triangular; ela está em forma escada. □

Suponha, agora, que modificamos os números à direita do sinal de igualdade de modo a obter um sistema compatível. Por exemplo, se começarmos com

$$\left(\begin{array}{rrrrr|r} 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 & -1 \\ -2 & -2 & 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 1 & 3 & 3 \\ 1 & 1 & 2 & 2 & 4 & 4 \end{array}\right)$$

então o processo de redução vai resultar na matriz aumentada

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

As duas últimas equações do sistema reduzido são satisfeitas por qualquer quíntupla. Logo, o conjunto solução é o conjunto de todas as quíntuplas que satisfazem as três primeiras equações.

(1)
$$\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 &= 1 \\ x_3 + x_4 + 2x_5 &= 0 \\ x_5 &= 3 \end{aligned}$$

Vamos nos referir às variáveis correspondentes aos dois primeiros elementos não-nulos como *variáveis líderes*.* Então, x_1, x_3 e x_5 são as variáveis líderes. As variáveis restantes, correspondentes às colunas que foram puladas no processo de redução, serão chamadas de *variáveis livres*. Logo, as variáveis livres são x_2 e x_4. Transferindo as variáveis livres para o lado direito em (1), obtemos o sistema

(2)
$$\begin{aligned} x_1 + x_3 + x_5 &= 1 - x_2 - x_4 \\ x_3 + 2x_5 &= -x_4 \\ x_5 &= 3 \end{aligned}$$

O sistema (2) é triangular nas incógnitas x_1, x_3, x_5. Portanto, para cada par de valores dados a x_2 e x_4, existirá uma única solução. Por exemplo, se $x_2 = x_4 = 0$, então $x_5 = 3$, $x_3 = -6$, $x_1 = 4$, logo $(4, 0, -6, 0, 3)$ é uma solução do sistema.

Definição. Uma matriz está em **forma escada** se:

(i) o primeiro elemento não-nulo de cada linha é 1;
(ii) se a linha k não consiste apenas em zeros, o número de zeros no início da linha $k + 1$ é maior do que o número de zeros no início da linha k;
(iii) se existirem linhas com todos os elementos iguais a zero, elas ficam abaixo de todas as linhas não-nulas.

* Essa terminologia não é padrão. (*N. T.*)

EXEMPLO 2. As matrizes a seguir estão em forma escada.

$$\begin{pmatrix} 1 & 4 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 3 & 1 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

□

EXEMPLO 3. As matrizes a seguir não estão em forma escada.

$$\begin{pmatrix} 2 & 4 & 6 \\ 0 & 3 & 5 \\ 0 & 0 & 4 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

A primeira matriz não satisfaz a primeira condição. A segunda matriz não satisfaz a terceira condição e a terceira matriz não satisfaz a segunda condição. □

Definição. O processo de usar as operações I, II e III para transformar um sistema linear em outro cuja matriz aumentada está em forma escada é chamado de **método de Gauss.**

Observe que a operação elementar II é necessária para multiplicar as linhas por escalares de modo que os primeiros coeficientes não-nulos de cada linha sejam iguais a 1. Se a matriz em forma escada contém uma linha da forma

$$\left(0 \quad 0 \quad \cdots \quad 0 \,\middle|\, 1\right)$$

o sistema é incompatível. Caso contrário, o sistema é compatível. Se o sistema é compatível (ou consistente) e as linhas não-nulas da matriz em forma escada representam um sistema triangular, então o sistema tem uma única solução.

SISTEMAS COM MAIS EQUAÇÕES DO QUE INCÓGNITAS

Um sistema linear *tem mais equações do que incógnitas* se $m > n$. Em geral (mas nem sempre), tais sistemas são impossíveis.

EXEMPLO 4.

(a)
$$\begin{aligned} x_1 + x_2 &= 1 \\ x_1 - x_2 &= 3 \\ -x_1 + 2x_2 &= -2 \end{aligned}$$

(b)
$$\begin{aligned} x_1 + 2x_2 + x_3 &= 1 \\ 2x_1 - x_2 + x_3 &= 2 \\ 4x_1 + 3x_2 + 3x_3 &= 4 \\ 2x_1 - x_2 + 3x_3 &= 5 \end{aligned}$$

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

SOLUÇÃO. O leitor já deve estar suficientemente familiarizado com o método de Gauss, de modo que podemos omitir as etapas intermediárias na redução de cada um desses sistemas.

Sistema (a)

$$\left(\begin{array}{rr|r} 1 & 1 & 1 \\ 1 & -1 & 3 \\ -1 & 2 & -2 \end{array}\right) \rightarrow \left(\begin{array}{rr|r} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{array}\right)$$

Pela última linha da matriz reduzida, vemos que o sistema é incompatível. As três equações do sistema (a) representam retas no plano. As duas primeiras se interceptam no ponto $(2, -1)$. No entanto, a

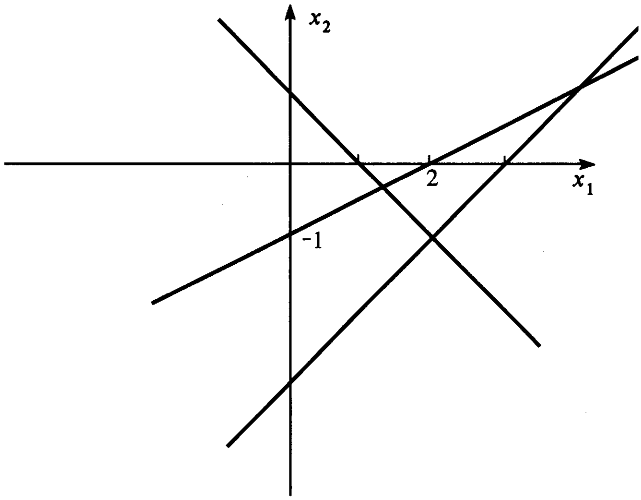


FIG. 1.2.1

terceira reta não contém esse ponto. Logo, não existe nenhum ponto pertencente a todas as três retas (ver Fig. 1.2.1).

Sistema (b)

$$\left(\begin{array}{rrr|r} 1 & 2 & 1 & 1 \\ 2 & -1 & 1 & 2 \\ 4 & 3 & 3 & 4 \\ 2 & -1 & 3 & 5 \end{array}\right) \rightarrow \left(\begin{array}{rrr|r} 1 & 2 & 1 & 1 \\ 0 & 1 & \frac{1}{5} & 0 \\ 0 & 0 & 1 & \frac{3}{2} \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Usando substituição, vemos que o sistema (b) tem exatamente uma solução (0,1, −0,3, 1,5). A solução é única, pois as linhas não-nulas da matriz reduzida formam um sistema triangular.

Sistema (c)

$$\left(\begin{array}{rrr|r} 1 & 2 & 1 & 1 \\ 2 & -1 & 1 & 2 \\ 4 & 3 & 3 & 4 \\ 3 & 1 & 2 & 3 \end{array}\right) \rightarrow \left(\begin{array}{rrr|r} 1 & 2 & 1 & 1 \\ 0 & 1 & \frac{1}{5} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Resolvendo para x_2 e x_1 em termos de x_3, obtemos

$$\begin{aligned} x_2 &= -0{,}2x_3 \\ x_1 &= 1 - 2x_2 - x_3 = 1 - 0{,}6x_3 \end{aligned}$$

Logo, o conjunto solução é o conjunto de todas as triplas ordenadas da forma $(1 - 0,6\alpha, -0,2\alpha, \alpha)$, onde α é um número real. O sistema é possível e indeterminado (tem infinitas soluções) por causa da variável x_3. □

SISTEMAS COM MENOS EQUAÇÕES DO QUE INCÓGNITAS

Um sistema linear tem menos equações do que incógnitas se $m < n$. Embora seja possível para um tal sistema ser incompatível, eles são, em geral, compatíveis e indeterminados. Um tal sistema nunca pode ser possível e determinado (isto é, ter uma única solução). A razão disso é que a forma escada da matriz de coeficientes tem que ter $r \leq m$ linhas não-nulas. Teremos então r variáveis líderes e $n - r$ variáveis livres, onde $n - r \geq n - m > 0$. Se o sistema for compatível, podemos dar valores arbitrários para as variáveis livres e resolver para as outras variáveis. Portanto, um sistema compatível com menos equações do que incógnitas sempre tem infinitas soluções.

EXEMPLO 5.

$$\text{(a)}\quad \begin{aligned} x_1 + 2x_2 + x_3 &= 1 \\ 2x_1 + 4x_2 + 2x_3 &= 3 \end{aligned}$$

$$\text{(b)}\quad \begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 &= 2 \\ x_1 + x_2 + x_3 + 2x_4 + 2x_5 &= 3 \\ x_1 + x_2 + x_3 + 2x_4 + 3x_5 &= 2 \end{aligned}$$

SOLUÇÃO

Sistema (a)

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 4 & 2 & 3 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

É claro que o sistema (a) é incompatível. Podemos pensar nas duas equações do sistema (a) como representando planos no espaço tridimensional. Em geral, dois planos se interceptam ao longo de uma reta; no entanto, nesse caso, os planos são paralelos.

Sistema (b)

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 2 & 2 & 3 \\ 1 & 1 & 1 & 2 & 3 & 2 \end{array}\right) \to \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right)$$

O sistema (b) é compatível, e, como existem duas variáveis livres, o sistema tem infinitas soluções. É muitas vezes conveniente, com sistemas desse tipo, continuar o processo de redução até anular todos os termos acima dos primeiros elementos não-nulos de cada linha. Para o sistema (b), então, vamos continuar e anular os dois primeiros elementos da quinta coluna e, depois, o primeiro elemento da quarta coluna.

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right) \to \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right)$$

$$\to \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right)$$

Colocando as variáveis livres do lado direito do sinal de igualdade, obtemos

$$\begin{aligned} x_1 &= 1 - x_2 - x_3 \\ x_4 &= 2 \\ x_5 &= -1 \end{aligned}$$

Portanto, para quaisquer α e β reais, a quíntupla

$$(1 - \alpha - \beta, \alpha, \beta, 2, -1)$$

é uma solução do sistema. □

FORMA ESCADA REDUZIDA POR LINHAS

Definição. Uma matriz está em **forma escada reduzida por linhas** se:

(i) a matriz está em forma escada;
(ii) o primeiro elemento não-nulo de cada linha é o único elemento diferente de zero na sua coluna.

As seguintes matrizes estão em forma escada reduzida por linhas:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

O processo de usar operações elementares para colocar uma matriz em forma escada reduzida por linhas é conhecido como *método de Gauss-Jordan.*

EXEMPLO 6. Use o método de Gauss-Jordan para resolver o sistema

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

SOLUÇÃO

$$\left(\begin{array}{cccc|c} -1 & 1 & -1 & 3 & 0 \\ \boxed{3} & 1 & -1 & -1 & 0 \\ \boxed{2} & -1 & -2 & -1 & 0 \end{array}\right) \rightarrow \left(\begin{array}{cccc|c} -1 & 1 & -1 & 3 & 0 \\ 0 & 4 & -4 & 8 & 0 \\ 0 & \boxed{1} & -4 & 5 & 0 \end{array}\right)$$

$$\rightarrow \left(\begin{array}{cccc|c} -1 & 1 & -1 & 3 & 0 \\ 0 & 4 & -4 & 8 & 0 \\ 0 & 0 & -3 & 3 & 0 \end{array}\right) \rightarrow \left(\begin{array}{cccc|c} 1 & -1 & \boxed{1} & -3 & 0 \\ 0 & 1 & \boxed{-1} & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{array}\right) \text{forma escada}$$

$$\rightarrow \left(\begin{array}{cccc|c} 1 & \boxed{-1} & 0 & -2 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{array}\right) \rightarrow \left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{array}\right) \text{forma escada reduzida por linhas}$$

Igualando x_4 a um número real arbitrário α, temos $x_1 = \alpha$, $x_2 = -\alpha$ e $x_3 = \alpha$. Portanto, todas as quádruplas da forma $(\alpha, -\alpha, \alpha, \alpha)$ são soluções do sistema. □

APLICAÇÃO 1: FLUXO DE TRÁFEGO

Em uma certa seção do centro de determinada cidade, dois conjuntos de ruas de mão única se cruzam, como ilustra a Fig. 1.2.2. A média do número de veículos por hora que entram e saem dessa seção durante o horário de *rush* é dada no diagrama. Determine a quantidade de veículos entre cada um dos quatro cruzamentos.

SOLUÇÃO. Em cada cruzamento, o número de veículos que entra tem que ser igual ao de veículos que sai. Por exemplo, no cruzamento A, o número de veículos que entra é $x_1 + 450$ e o número de veículos que sai é $x_2 + 610$. Logo,

$$x_1 + 450 = x_2 + 610 \qquad \text{(cruzamento } A\text{)}$$

Analogamente,

$$\begin{aligned} x_2 + 520 &= x_3 + 480 && \text{(cruzamento } B\text{)} \\ x_3 + 390 &= x_4 + 600 && \text{(cruzamento } C\text{)} \\ x_4 + 640 &= x_1 + 310 && \text{(cruzamento } D\text{)} \end{aligned}$$

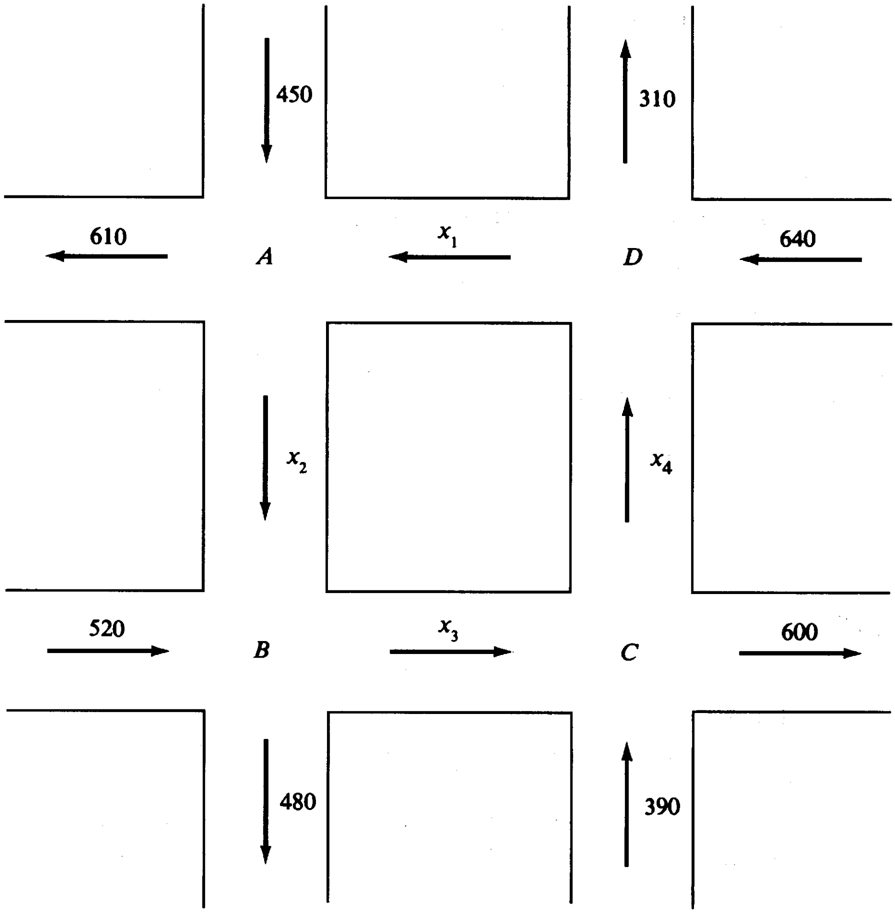


FIG. 1.2.2

A matriz aumentada para esse sistema é

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 160 \\ 0 & 1 & -1 & 0 & -40 \\ 0 & 0 & 1 & -1 & 210 \\ -1 & 0 & 0 & 1 & -330 \end{array}\right)$$

A forma escada reduzida por linhas dessa matriz é

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 330 \\ 0 & 1 & 0 & -1 & 170 \\ 0 & 0 & 1 & -1 & 210 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

O sistema é compatível, e, como tem uma variável livre, existem muitas soluções possíveis. O diagrama de fluxo do tráfego não contém informação suficiente para determinar x_1, x_2, x_3, x_4. Se o número de veículos entre dois dos cruzamentos fosse conhecido, o tráfego nos outros cruzamentos estaria determinado. Por exemplo, se uma média de 200 carros trafega por hora entre os cruzamentos C e D, então $x_4 = 200$. Podemos, então, resolver para x_1, x_2, x_3 em termos de x_4, obtendo

$$x_1 = x_4 + 330 = 530$$

$$x_2 = x_4 + 170 = 370$$

$$x_3 = x_4 + 210 = 410$$

□

SISTEMAS HOMOGÊNEOS

Um sistema de equações lineares é dito *homogêneo* se todas as constantes do lado direito dos sinais de igualdade são nulas. Sistemas homogêneos sempre são compatíveis. É trivial encontrar uma solução: basta fazer todas as variáveis iguais a zero. Portanto, se um sistema homogêneo $m \times n$ tiver uma única solução, ela tem que ser a solução trivial $(0, 0, \ldots, 0)$. O sistema homogêneo no Exemplo 6 tem $m = 3$ equações e $n = 4$ incógnitas. No caso em que $n > m$, sempre vai existir uma variável livre e, portanto, sempre vão existir soluções não-triviais. Esse resultado foi essencialmente mostrado na nossa discussão sobre sistemas com menos equações do que incógnitas, mas, devido à sua importância, vamos enunciá-lo como um teorema.

Teorema 1.2.1. *Um sistema homogêneo $m \times n$ de equações lineares tem uma solução não-trivial se $n > m$.*

Demonstração. Um sistema homogêneo é sempre compatível. A forma escada da matriz pode ter, no máximo, m linhas não-nulas. Logo, existem, no máximo, m variáveis líderes. Como existem $n > m$ variáveis ao todo, sempre vai existir alguma variável livre. As variáveis livres podem assumir valores arbitrários. Então, existe uma solução do sistema para cada conjunto de valores das variáveis livres. □

APLICAÇÃO 2: CIRCUITOS ELÉTRICOS

Em um circuito elétrico é possível determinar a corrente em cada trecho em termos das resistências e das diferenças de potencial. Na Fig. 1.2.3 o símbolo representa uma bateria (medida em volts) que gera uma carga que produz uma corrente. A corrente sai da bateria do lado que contém a reta vertical mais longa, isto é, . O símbolo representa um resistor. As resistências são medidas em ohms. As letras maiúsculas representam os nós, e i representa a corrente entre os nós. As correntes são medidas em ampères. As setas mostram o sentido do fluxo da corrente. Se, no entanto, uma das

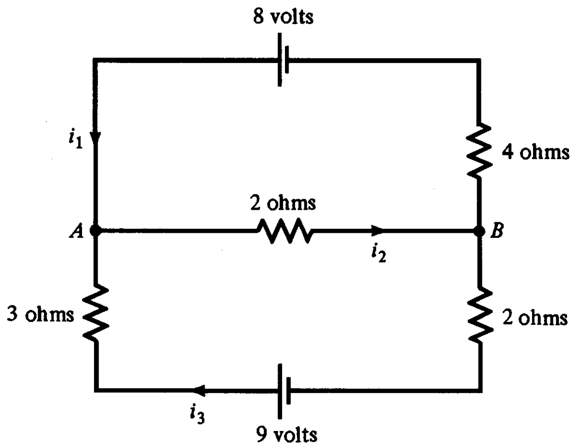


FIG. 1.2.3

correntes, i_2, por exemplo, é negativa, isso significa que a corrente naquele trecho flui no sentido oposto ao da seta.

Para determinar as correntes, são utilizadas as *leis de Kirchhoff*:

1. Em cada nó, a soma das correntes que entram é igual à soma das correntes que saem.
2. Em cada ciclo fechado, a diferença de potencial total é zero.

A diferença de potencial elétrico E em cada resistor é dada pela *lei de Ohm*:

$$E = iR$$

onde i representa a corrente em ampères e R a resistência em ohms.

Vamos encontrar as correntes no circuito ilustrado na Fig. 1.2.3. Da primeira lei, obtemos

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \qquad (\text{nó } A) \\ -i_1 + i_2 - i_3 &= 0 \qquad (\text{nó } B) \end{aligned}$$

Da segunda lei, temos

$$\begin{aligned} 4i_1 + 2i_2 &= 8 \qquad (\text{ciclo superior}) \\ 2i_2 + 5i_3 &= 9 \qquad (\text{ciclo inferior}) \end{aligned}$$

O circuito pode ser representado pela matriz aumentada

$$\left(\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ 4 & 2 & 0 & 8 \\ 0 & 2 & 5 & 9 \end{array}\right)$$

Essa matriz pode ser facilmente reduzida à forma escada

$$\left(\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 0 & 1 & -\frac{2}{3} & \frac{4}{3} \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Resolvendo por substituição, vemos que $i_1 = 1$, $i_2 = 2$ e $i_3 = 1$.

EXERCÍCIOS

1. Quais das matrizes a seguir estão em forma escada? Quais estão em forma escada reduzida por linhas?

(a) $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 3 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ (d) $\begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$

(e) $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix}$ (f) $\begin{pmatrix} 1 & 4 & 6 \\ 0 & 0 & 1 \\ 0 & 1 & 3 \end{pmatrix}$ (g) $\begin{pmatrix} 1 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 2 & 4 \\ 0 & 0 & 1 & 3 & 6 \end{pmatrix}$ (h) $\begin{pmatrix} 0 & 1 & 3 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

2. Cada uma das matrizes aumentadas a seguir está em forma escada. Para cada uma delas, indique se o sistema linear correspondente é compatível ou não. Se o sistema tiver uma única solução, encontre-a.

(a) $\left(\begin{array}{rr|r} 1 & 2 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{array}\right)$ (b) $\left(\begin{array}{rr|r} 1 & 3 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{array}\right)$ (c) $\left(\begin{array}{rrr|r} 1 & -2 & 4 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right)$

(d) $\left(\begin{array}{ccc|c} 1 & -2 & 2 & -2 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 1 & 2 \end{array}\right)$ (e) $\left(\begin{array}{ccc|c} 1 & 3 & 2 & -2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 \end{array}\right)$ (f) $\left(\begin{array}{ccc|c} 1 & -1 & 3 & 8 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right)$

3. Cada uma das matrizes aumentadas a seguir está em forma escada reduzida por linhas. Para cada uma delas, encontre o conjunto solução do sistema linear correspondente.

(a) $\left(\begin{array}{ccc|c} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \end{array}\right)$ (b) $\left(\begin{array}{ccc|c} 1 & 4 & 0 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{array}\right)$ (c) $\left(\begin{array}{ccc|c} 1 & -3 & 0 & 2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right)$

(d) $\left(\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 5 \\ 0 & 0 & 1 & 3 & 4 \end{array}\right)$ (e) $\left(\begin{array}{cccc|c} 1 & 5 & -2 & 0 & 3 \\ 0 & 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$ (f) $\left(\begin{array}{ccc|c} 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right)$

4. Para cada um dos sistemas de equações lineares a seguir, use o método de Gauss para obter um sistema equivalente cuja matriz de coeficientes esteja em forma escada. Indique se o sistema é ou não consistente. Se o sistema for possível e determinado (isto é, sem variáveis livres), use substituição para encontrar a única solução. Se o sistema for possível e indeterminado, coloque-o em forma escada reduzida por linhas e encontre todas as suas soluções.

(a) $\begin{aligned} x_1 - 2x_2 &= 3 \\ 2x_1 - x_2 &= 9 \end{aligned}$

(b) $\begin{aligned} 2x_1 - 3x_2 &= 5 \\ -4x_1 + 6x_2 &= 8 \end{aligned}$

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

(d) $\begin{aligned} 3x_1 + 2x_2 - x_3 &= 4 \\ x_1 - 2x_2 + 2x_3 &= 1 \\ 11x_1 + 2x_2 + x_3 &= 14 \end{aligned}$

(e) $\begin{aligned} 2x_1 + 3x_2 + x_3 &= 1 \\ x_1 + x_2 + x_3 &= 3 \\ 3x_1 + 4x_2 + 2x_3 &= 4 \end{aligned}$

(f) $\begin{aligned} x_1 - x_2 + 2x_3 &= 4 \\ 2x_1 + 3x_2 - x_3 &= 1 \\ 7x_1 + 3x_2 + 4x_3 &= 7 \end{aligned}$

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

(h) $\begin{aligned} x_1 - 2x_2 &= 3 \\ 2x_1 + x_2 &= 1 \\ -5x_1 + 8x_2 &= 4 \end{aligned}$

(i) $\begin{aligned} -x_1 + 2x_2 - x_3 &= 2 \\ -2x_1 + 2x_2 + x_3 &= 4 \\ 3x_1 + 2x_2 + 2x_3 &= 5 \\ -3x_1 + 8x_2 + 5x_3 &= 17 \end{aligned}$

(j) $\begin{aligned} x_1 + 2x_2 - 3x_3 + x_4 &= 1 \\ -x_1 - x_2 + 4x_3 - x_4 &= 6 \\ -2x_1 - 4x_2 + 7x_3 - x_4 &= 1 \end{aligned}$

(k) $\begin{aligned} x_1 + 3x_2 + x_3 + x_4 &= 3 \\ 2x_1 - 2x_2 + x_3 + 2x_4 &= 8 \\ x_1 - 5x_2 + x_4 &= 5 \end{aligned}$

(l) $\begin{aligned} x_1 - 3x_2 + x_3 &= 1 \\ 2x_1 + x_2 - x_3 &= 2 \\ x_1 + 4x_2 - 2x_3 &= 1 \\ 5x_1 - 8x_2 + 2x_3 &= 5 \end{aligned}$

5. Use o método de Gauss-Jordan para resolver cada um dos sistemas a seguir.

(a)
$$\begin{aligned} x_1 + x_2 &= -1 \\ 4x_1 - 3x_2 &= 3 \end{aligned}$$

(b)
$$\begin{aligned} x_1 + 3x_2 + x_3 + x_4 &= 3 \\ 2x_1 - 2x_2 + x_3 + 2x_4 &= 8 \\ 3x_1 + x_2 + 2x_3 - x_4 &= -1 \end{aligned}$$

(c)
$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ x_1 - x_2 - x_3 &= 0 \end{aligned}$$

(d)
$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 0 \\ 2x_1 + x_2 - x_3 + 3x_4 &= 0 \\ x_1 - 2x_2 + x_3 + x_4 &= 0 \end{aligned}$$

6. Considere um sistema linear cuja matriz aumentada tem a forma

$$\left(\begin{array}{rrr|r} 1 & 2 & 1 & 1 \\ -1 & 4 & 3 & 2 \\ 2 & -2 & a & 3 \end{array}\right)$$

Para que valores de a o sistema tem uma única solução?

7. Considere um sistema linear cuja matriz aumentada é da forma

$$\left(\begin{array}{rrr|r} 1 & 2 & 1 & 0 \\ 2 & 5 & 3 & 0 \\ -1 & 1 & \beta & 0 \end{array}\right)$$

(a) O sistema pode ser incompatível? Explique.
(b) Para que valores de β o sistema tem infinitas soluções?

8. Considere um sistema linear cuja matriz aumentada tem a forma

$$\left(\begin{array}{rrr|r} 1 & 1 & 3 & 2 \\ 1 & 2 & 4 & 3 \\ 1 & 3 & a & b \end{array}\right)$$

(a) Para que valores de a e b o sistema tem uma infinidade de soluções?
(b) Para que valores de a e b o sistema é impossível?

9. Dados os sistemas lineares

(a)
$$\begin{aligned} x_1 + 2x_2 &= 2 \\ 3x_1 + 7x_2 &= 8 \end{aligned}$$

(b)
$$\begin{aligned} x_1 + 2x_2 &= 1 \\ 3x_1 + 7x_2 &= 7 \end{aligned}$$

resolva simultaneamente ambos os sistemas incorporando os termos à direita dos sinais de igualdade em uma matriz B 2×2 e colocando em forma escada reduzida por linhas a matriz

$$(A|B) = \left(\begin{array}{rr|rr} 1 & 2 & 2 & 1 \\ 3 & 7 & 8 & 7 \end{array}\right)$$

10. Dados os sistemas lineares

(a)
$$\begin{aligned} x_1 + 2x_2 + x_3 &= 2 \\ -x_1 - x_2 + 2x_3 &= 3 \\ 2x_1 + 3x_2 &= 0 \end{aligned}$$

(b)
$$\begin{aligned} x_1 + 2x_2 + x_3 &= -1 \\ -x_1 - x_2 + 2x_3 &= 2 \\ 2x_1 + 3x_2 &= -2 \end{aligned}$$

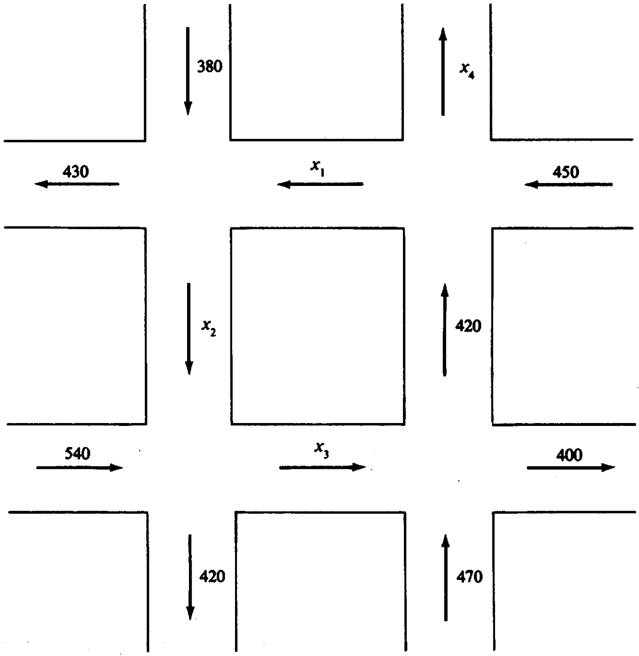
resolva simultaneamente ambos os sistemas colocando em forma escada a matriz aumentada $(A|B)$ e usando substituição duas vezes.

11. Seja (c_1, c_2) um solução do sistema 2×2

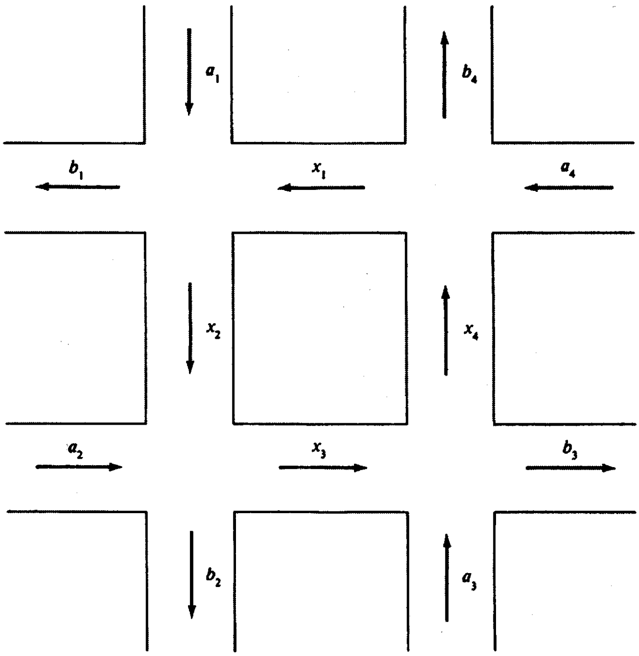
$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= 0 \\ a_{21}x_1 + a_{22}x_2 &= 0 \end{aligned}$$

Mostre que, qualquer que seja o número real α, o par ordenado $(\alpha c_1, \alpha c_2)$ é também uma solução.

12. Determine os valores de x_1, x_2, x_3, x_4 para o seguinte diagrama de fluxo de tráfego:



13. Considere o seguinte diagrama de fluxo de tráfego:



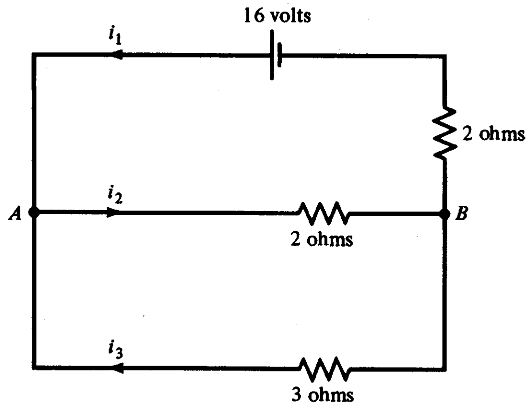
onde a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4 são inteiros positivos fixos. Escreva um sistema linear com as incógnitas x_1, x_2, x_3, x_4 e mostre que o sistema é compatível se e somente se

$$a_1 + a_2 + a_3 + a_4 = b_1 + b_2 + b_3 + b_4$$

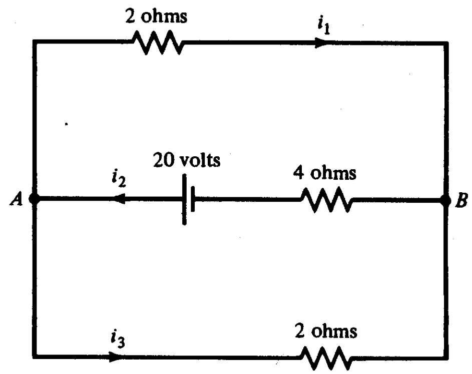
O que você pode concluir sobre o número de veículos que entram e saem da seção ilustrada no diagrama?

14. Determine a corrente em cada um dos trechos dos circuitos ilustrados a seguir.

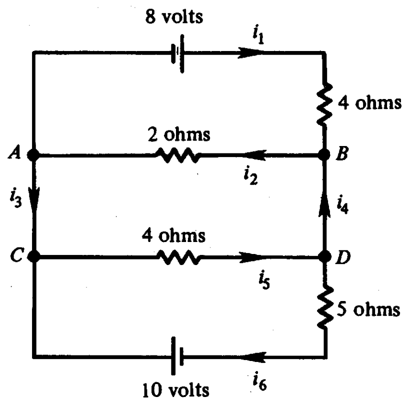
(a)



(b)



(c)



3 ÁLGEBRA MATRICIAL

Nesta seção, vamos definir as operações aritméticas de matrizes e estabelecer algumas de suas propriedades algébricas. Matrizes estão entre as ferramentas mais poderosas da matemática. Para utilizar eficientemente as matrizes, precisamos conhecer a aritmética matricial.

Os elementos de uma matriz são chamados *escalares*. Eles são, em geral, números reais ou complexos. Na maioria das vezes estaremos trabalhando com matrizes cujos elementos são números reais. Ao longo dos cinco primeiros capítulos deste texto, o leitor pode supor que o termo *escalar* se refere a um número real. Entretanto, no Cap. 6, usaremos algumas vezes o conjunto dos números complexos como nosso corpo escalar.

Se quisermos nos referir a matrizes sem escrever especificamente todos os seus elementos, usaremos letras maiúsculas A, B, C e assim por diante. Em geral, a_{ij} denota o elemento da matriz A que fica na i-ésima linha e j-ésima coluna. Então, se A é uma matriz $m \times n$, temos

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Algumas vezes, abreviaremos essa notação escrevendo $A = (a_{ij})$. Analogamente, podemos nos referir à matriz B por (b_{ij}), à matriz C por (c_{ij}) e assim por diante.

IGUALDADE

Definição. Duas matrizes $m \times n$ A e B são ditas **iguais** se $a_{ij} = b_{ij}$ para todos os i e j.

MULTIPLICAÇÃO POR UM ESCALAR

Se A é uma matriz e α é um escalar, então αA é a matriz obtida multiplicando-se cada elemento de A por α. Por exemplo, se

$$A = \begin{pmatrix} 4 & 8 & 2 \\ 6 & 8 & 10 \end{pmatrix}$$

então

$$\tfrac{1}{2}A = \begin{pmatrix} 2 & 4 & 1 \\ 3 & 4 & 5 \end{pmatrix} \quad \text{e} \quad 3A = \begin{pmatrix} 12 & 24 & 6 \\ 18 & 24 & 30 \end{pmatrix}$$

SOMA DE MATRIZES

Se $A = (a_{ij})$ e $B = (b_{ij})$ são ambas matrizes $m \times n$, então a soma $A + B$ é a matriz $m \times n$ cujo elemento (i, j) é $a_{ij} + b_{ij}$, para cada par ordenado (i, j). Por exemplo,

$$\begin{pmatrix} 3 & 2 & 1 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 2 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 5 & 4 & 3 \\ 5 & 7 & 9 \end{pmatrix}$$

$$\begin{pmatrix} 2 \\ 1 \\ 8 \end{pmatrix} + \begin{pmatrix} -8 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} -6 \\ 4 \\ 10 \end{pmatrix}$$

Se definirmos $A - B$ por $A + (-1)B$, então $A - B$ é obtida subtraindo-se de cada elemento de A o elemento correspondente de B. Então,

$$\begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix} - \begin{pmatrix} 4 & 5 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix} + (-1)\begin{pmatrix} 4 & 5 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix} + \begin{pmatrix} -4 & -5 \\ -2 & -3 \end{pmatrix}$$

$$= \begin{pmatrix} 2-4 & 4-5 \\ 3-2 & 1-3 \end{pmatrix} = \begin{pmatrix} -2 & -1 \\ 1 & -2 \end{pmatrix}$$

Se O representa a matriz de mesmo tamanho que A, que tem todos os elementos iguais a 0, então

$$A + O = O + A = A$$

Em outras palavras, a matriz nula age como um elemento neutro em relação à soma no conjunto de todas as matrizes $m \times n$. Além disso, cada matriz A $m \times n$ tem uma inversa aditiva. De fato,

$$A + (-1)A = O = (-1)A + A$$

É comum denotar a inversa aditiva por $-A$. Então

$$-A = (-1)A$$

MULTIPLICAÇÃO DE MATRIZES

Temos ainda que definir a operação mais importante, a multiplicação entre duas matrizes. Grande parte da motivação por trás dessa definição vem de aplicações a sistemas lineares. Se temos um sistema de uma equação linear em uma incógnita, ele pode ser escrito na forma

$$ax = b \tag{1}$$

Pensamos, em geral, em a, x e b como escalares; no entanto, também poderiam ser tratados como matrizes 1×1. Mais geralmente, dado um sistema linear $m \times n$

$$\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\
\vdots & & & & & & & & \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m
\end{array}$$

é desejável escrever o sistema em uma forma semelhante a (1), isto é, como uma equação matricial

$$AX = B$$

onde $A = (a_{ij})$ é conhecida, X é uma matriz $n \times 1$ de incógnitas e B é uma matriz $m \times 1$ que representa os elementos à direita dos sinais de igualdade. Definimos, então,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

e

$$AX = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} \tag{2}$$

Dada uma matriz A $m \times n$ e uma matriz X $n \times 1$, sempre é possível calcular o produto AX por (2). O produto AX é uma matriz $m \times 1$. A regra para determinar o i-ésimo elemento de AX é

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$$

Observe que o i-ésimo elemento é determinado pela i-ésima linha de A. Os elementos daquela linha são multiplicados pelos elementos correspondentes de X e os n produtos são, então, somados. Os leitores familiarizados com produtos internos reconhecerão essa fórmula, simplesmente, como o produto interno da n-upla correspondente à i-ésima linha de A com a n-upla correspondente à matriz X.

$$\begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$$

Para "casar" os elementos correspondentes dessa maneira, o número de colunas de A tem que ser igual ao número de linhas de X. Os elementos de X podem ser escalares ou incógnitas com valores escalares.

EXEMPLO 1.

$$A = \begin{pmatrix} 4 & 2 & 1 \\ 5 & 3 & 7 \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$AX = \begin{pmatrix} 4x_1 + 2x_2 + x_3 \\ 5x_1 + 3x_2 + 7x_3 \end{pmatrix}$$ □

EXEMPLO 2.

$$A = \begin{pmatrix} -3 & 1 \\ 2 & 5 \\ 4 & 2 \end{pmatrix}, \quad X = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

$$AX = \begin{pmatrix} -3\cdot 2 + 1\cdot 4 \\ 2\cdot 2 + 5\cdot 4 \\ 4\cdot 2 + 2\cdot 4 \end{pmatrix} = \begin{pmatrix} -2 \\ 24 \\ 16 \end{pmatrix}$$ □

EXEMPLO 3. Escreva o sistema de equações a seguir como uma equação matricial $AX = B$.

$$\begin{aligned} 3x_1 + 2x_2 + x_3 &= 5 \\ x_1 - 2x_2 + 5x_3 &= -2 \\ 2x_1 + x_2 - 3x_3 &= 1 \end{aligned}$$

SOLUÇÃO

$$\begin{pmatrix} 3 & 2 & 1 \\ 1 & -2 & 5 \\ 2 & 1 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \\ 1 \end{pmatrix}$$ □

Mais geralmente, é possível multiplicar uma matriz A por uma matriz B se o número de colunas de A é igual ao número de linhas de B. A primeira coluna do produto é determinada pela primeira coluna de B, a segunda pela segunda coluna de B e assim por diante. Portanto, para determinar o elemento (i, j) do produto AB, usamos os elementos da i-ésima linha de A e da j-ésima coluna de B.

Definição. Se $A = (a_{ij})$ é uma matriz $m \times n$ e $B = (b_{ij})$ é uma matriz $n \times r$, então o produto $AB = C = (c_{ij})$ é a matriz $m \times r$ cujos elementos são definidos por

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

O que essa definição diz é que, para formar o elemento (i,j) do produto, você tem que pegar a i-ésima linha de A e a j-ésima coluna de B, multiplicar os elementos correspondentes dois a dois e somar os números resultantes.

$$\begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix}$$

Para "casar" os elementos dessa maneira, o número de colunas de A tem que ser igual ao número de linhas de B. Se isso não acontece, a multiplicação é impossível.

EXEMPLO 4. Se

$$A = \begin{pmatrix} -2 & 1 & 3 \\ 4 & 1 & 6 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 3 & -2 \\ 2 & 4 \\ 1 & -3 \end{pmatrix}$$

então

$$AB = \begin{pmatrix} -2\cdot 3 + 1\cdot 2 + 3\cdot 1 & -2\cdot(-2) + 1\cdot 4 + 3\cdot(-3) \\ 4\cdot 3 + 1\cdot 2 + 6\cdot 1 & 4\cdot(-2) + 1\cdot 4 + 6\cdot(-3) \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 20 & -22 \end{pmatrix}$$

e

$$\begin{aligned} BA &= \begin{pmatrix} 3\cdot(-2) - 2\cdot 4 & 3\cdot 1 - 2\cdot 1 & 3\cdot 3 - 2\cdot 6 \\ 2\cdot(-2) + 4\cdot 4 & 2\cdot 1 + 4\cdot 1 & 2\cdot 3 + 4\cdot 6 \\ 1\cdot(-2) - 3\cdot 4 & 1\cdot 1 - 3\cdot 1 & 1\cdot 3 - 3\cdot 6 \end{pmatrix} \\ &= \begin{pmatrix} -14 & 1 & -3 \\ 12 & 6 & 30 \\ -14 & -2 & -15 \end{pmatrix} \end{aligned}$$

□

EXEMPLO 5. Se

$$A = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 1 & 2 \\ 4 & 5 \\ 3 & 6 \end{pmatrix}$$

então é impossível multiplicar A por B, já que o número de colunas de A não é igual ao número de linhas de B. No entanto, é possível multiplicar B por A:

$$BA = \begin{pmatrix} 1 & 2 \\ 4 & 5 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 8 \\ 17 & 26 \\ 15 & 24 \end{pmatrix}$$

□

Se A e B são ambas matrizes $n \times n$, então AB e BA também são matrizes $n \times n$, mas, em geral, elas não são iguais. *A multiplicação de matrizes não é comutativa.*

EXEMPLO 6. Se

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$$

então

$$AB = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 0 & 0 \end{pmatrix}$$

e

$$BA = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$$

e, portanto, $AB \neq BA$. □

APLICAÇÃO 1

João pesa 81 quilos. Ele quer perder peso por meio de um programa de dieta e exercícios. Após consultar a Tabela 1, ele monta o programa de exercícios na Tabela 2. Quantas calorias ele vai queimar por dia se seguir esse programa?

TABELA 1. Calorias Queimadas por Hora

	Atividade esportiva			
Peso	Andar a 3 km/h	Correr a 9 km/h	Andar de bicicleta a 9 km/h	Jogar tênis (moderado)
69	213	651	304	420
73	225	688	321	441
77	237	726	338	468
81	249	764	356	492

TABELA 2. Horas por Dia para Cada Atividade

	Programa de exercícios			
	Andar	Correr	Andar de bicicleta	Jogar tênis
Segunda-feira	1,0	0,0	1,0	0,0
Terça-feira	0,0	0,0	0,0	2,0
Quarta-feira	0,4	0,5	0,0	0,0
Quinta-feira	0,0	0,0	0,5	2,0
Sexta-feira	0,4	0,5	0,0	0,0

SOLUÇÃO. A informação pertinente para João está localizada na quarta linha da Tabela 1. Essa informação pode ser representada por uma matriz X 4×1. A informação na Tabela 2 pode ser representada por uma matriz A 5×4. Para responder a pergunta, simplesmente calculamos AX.

$$\begin{pmatrix} 1,0 & 0,0 & 1,0 & 0,0 \\ 0,0 & 0,0 & 0,0 & 2,0 \\ 0,4 & 0,5 & 0,0 & 0,0 \\ 0,0 & 0,0 & 0,5 & 2,0 \\ 0,4 & 0,5 & 0,0 & 0,0 \end{pmatrix} \begin{pmatrix} 249 \\ 764 \\ 356 \\ 492 \end{pmatrix} = \begin{array}{l} \begin{pmatrix} 605,0 \\ 984,0 \\ 481,6 \\ 1162,0 \\ 481,6 \end{pmatrix} \end{array} \begin{array}{l} \text{Segunda-feira} \\ \text{Terça-feira} \\ \text{Quarta-feira} \\ \text{Quinta-feira} \\ \text{Sexta-feira} \end{array}$$

□

APLICAÇÃO 2

Uma empresa fabrica três produtos. Suas despesas de produção estão divididas em três categorias. Em cada uma dessas categorias, faz-se uma estimativa do custo de produção de um único exemplar de cada produto. Faz-se, também, uma estimativa da quantidade de cada produto a ser fabricado por trimestre. Essas estimativas são dadas nas Tabelas 3 e 4. A empresa gostaria de apresentar a seus acionistas uma

TABELA 3. Custo de Produção por Item (em dólares)

	Produto		
Gastos	A	B	C
Matéria-prima	0,10	0,30	0,15
Pessoal	0,30	0,40	0,25
Despesas gerais	0,10	0,20	0,15

TABELA 4. Quantidade Produzida por Trimestre

	Estação			
Produto	Verão	Outono	Inverno	Primavera
A	4000	4500	4500	4000
B	2000	2600	2400	2200
C	5800	6200	6000	6000

única tabela mostrando o custo total por trimestre de cada uma das três categorias: matéria-prima, pessoal e despesas gerais.

SOLUÇÃO. Vamos considerar o problema em termos de matrizes. Cada uma das duas tabelas pode ser representada por uma matriz.

$$M = \begin{pmatrix} 0{,}10 & 0{,}30 & 0{,}15 \\ 0{,}30 & 0{,}40 & 0{,}25 \\ 0{,}10 & 0{,}20 & 0{,}15 \end{pmatrix}$$

e

$$P = \begin{pmatrix} 4000 & 4500 & 4500 & 4000 \\ 2000 & 2600 & 2400 & 2200 \\ 5800 & 6200 & 6000 & 6000 \end{pmatrix}$$

Se formarmos o produto MP, a primeira coluna de MP vai representar os custos para o verão.

Matéria-prima: $(0{,}10)(4000) + (0{,}30)(2000) + (0{,}15)(5800) = 1870$
Pessoal: $(0{,}30)(4000) + (0{,}40)(2000) + (0{,}25)(5800) = 3450$
Despesas gerais: $(0{,}10)(4000) + (0{,}20)(2000) + (0{,}15)(5800) = 1670$

Os custos para o outono são dados pela segunda coluna de MP.

Matéria-prima: $(0{,}10)(4500) + (0{,}30)(2600) + (0{,}15)(6200) = 2160$
Pessoal: $(0{,}30)(4500) + (0{,}40)(2600) + (0{,}25)(6200) = 3940$
Despesas gerais: $(0{,}10)(4500) + (0{,}20)(2600) + (0{,}15)(6200) = 1900$

As colunas 3 e 4 de MP representam os custos para o inverno e a primavera.

$$MP = \begin{pmatrix} 1870 & 2160 & 2070 & 1960 \\ 3450 & 3940 & 3810 & 3580 \\ 1670 & 1900 & 1830 & 1740 \end{pmatrix}$$

TABELA 5

	Estação				
	Verão	Outono	Inverno	Primavera	Ano
Matéria-prima	1.870	2.160	2.070	1.960	8.060
Pessoal	3.450	3.940	3.810	3.580	14.780
Despesas gerais	1.670	1.900	1.830	1.740	7.140
Custo total de produção	6.990	8.000	7.710	7.280	29.980

Os elementos na primeira linha de MP representam o custo total de matéria-prima para cada um dos quatro trimestres. Os elementos nas linhas 2 e 3 representam o custo total de pessoal e despesas gerais, respectivamente, para cada um dos quatro trimestres. A despesa anual em cada categoria pode ser obtida somando-se os elementos de cada linha. Os números em cada coluna podem ser somados para se obter o custo total de produção para cada trimestre. A Tabela 5 resume o custo total de produção. □

REGRAS DE NOTAÇÃO

Como na álgebra usual, se uma expressão envolve multiplicações e somas, e não existem parênteses para indicar a ordem das operações, as multiplicações são efetuadas antes das somas. Isso é válido tanto para a multiplicação por escalar quanto para a multiplicação matricial. Por exemplo, se

$$A = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} -2 & 1 \\ 3 & 2 \end{pmatrix}$$

então

$$A + BC = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} 7 & 7 \\ -1 & 4 \end{pmatrix} = \begin{pmatrix} 10 & 11 \\ 0 & 6 \end{pmatrix}$$

e

$$3A + B = \begin{pmatrix} 9 & 12 \\ 3 & 6 \end{pmatrix} + \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 10 & 15 \\ 5 & 7 \end{pmatrix}$$

REGRAS ALGÉBRICAS

O teorema a seguir fornece algumas regras úteis para a aritmética matricial.

Teorema 1.3.1. *Cada uma das afirmações a seguir é válida quaisquer que sejam os escalares α e β e quaisquer que sejam as matrizes A, B e C para as quais as operações indicadas estão definidas.*

(1) $A + B = B + A$
(2) $(A + B) + C = A + (B + C)$
(3) $(AB)C = A(BC)$
(4) $A(B + C) = AB + AC$
(5) $(A + B)C = AC + BC$
(6) $(\alpha\beta)A = \alpha(\beta A)$
(7) $\alpha(AB) = (\alpha A)B = A(\alpha B)$
(8) $(\alpha + \beta)A = \alpha A + \beta A$
(9) $\alpha(A + B) = \alpha A + \alpha B$

Vamos demonstrar duas dessas regras e deixar as restantes a cargo do leitor.

Demonstração de (4). Suponha que $A = (a_{ij})$ é uma matriz $m \times n$ e que $B = (b_{ij})$ e $C = (c_{ij})$ são ambas matrizes $n \times r$. Sejam $D = A(B + C)$ e $E = AB + AC$. Então,

$$d_{ij} = \sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj})$$

e

$$e_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj}$$

Mas

$$\sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj}) = \sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj}$$

logo $d_{ij} = e_{ij}$ e, portanto, $A(B + C) = AB + AC$. □

Demonstração de (3). Sejam A uma matriz $m \times n$, B uma matriz $n \times r$ e C uma matriz $r \times s$. Sejam $D = AB$ e $E = BC$. Precisamos mostrar que $DC = AE$. Pela definição de multiplicação de matrizes,

$$d_{il} = \sum_{k=1}^{n} a_{ik}b_{kl} \quad \text{e} \quad e_{kj} = \sum_{l=1}^{r} b_{kl}c_{lj}$$

O elemento (i, j) de DC é

$$\sum_{l=1}^{r} d_{il}c_{lj} = \sum_{l=1}^{r} \left(\sum_{k=1}^{n} a_{ik}b_{kl} \right) c_{lj}$$

e o elemento (i, j) de AE é

$$\sum_{k=1}^{n} a_{ik}e_{kj} = \sum_{k=1}^{n} a_{ik} \left(\sum_{l=1}^{r} b_{kl}c_{lj} \right)$$

Como

$$\sum_{l=1}^{r} \left(\sum_{k=1}^{n} a_{ik}b_{kl} \right) c_{lj} = \sum_{l=1}^{r} \sum_{k=1}^{n} a_{ik}b_{kl}c_{lj} = \sum_{k=1}^{n} a_{ik} \left(\sum_{l=1}^{r} b_{kl}c_{lj} \right)$$

tem-se que

$$(AB)C = DC = AE = A(BC)$$ □

As regras aritméticas dadas no Teorema 1.3.1 parecem bastante naturais, já que são semelhantes às regras que utilizamos para números reais. No entanto, existem algumas diferenças importantes entre as regras para a aritmética matricial e as regras para os números reais. Em particular, a multiplicação de números reais é comutativa; entretanto, vimos, no Exemplo 6, que a multiplicação matricial não é comutativa. Essa diferença merece ser enfatizada.

Cuidado:	Em geral, $AB \neq BA$. A multiplicação de matrizes *não* é comutativa.

Algumas das outras diferenças entre a aritmética matricial e a aritmética para números reais estão ilustradas nos Exercícios 13, 14 e 15.

EXEMPLO 7. Se

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ -3 & 2 \end{pmatrix}, \quad \text{e} \quad C = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$$

verifique que $A(BC) = (AB)C$ e que $A(B + C) = AB + AC$.

SOLUÇÃO

$$A(BC) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 6 & 5 \\ 16 & 11 \end{pmatrix}$$

$$(AB)C = \begin{pmatrix} -4 & 5 \\ -6 & 11 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 5 \\ 16 & 11 \end{pmatrix}$$

Então

$$A(BC) = \begin{pmatrix} 6 & 5 \\ 16 & 11 \end{pmatrix} = (AB)C$$

$$A(B + C) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 7 \\ 5 & 15 \end{pmatrix}$$

$$AB + AC = \begin{pmatrix} -4 & 5 \\ -6 & 11 \end{pmatrix} + \begin{pmatrix} 5 & 2 \\ 11 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 7 \\ 5 & 15 \end{pmatrix}$$

Portanto,

$$A(B + C) = AB + AC$$

□

Notação. Como $(AB)C = A(BC)$, podemos, simplesmente, omitir os parênteses e escrever ABC. O mesmo é verdade para um produto de quatro ou mais matrizes. No caso em que uma matriz $n \times n$ é multiplicada por si mesma um certo número de vezes, é conveniente usar a notação exponencial. Então, se k é um inteiro positivo,

$$A^k = \underbrace{AA \cdots A}_{k \text{ vezes}}$$

EXEMPLO 8. Se

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

então

$$A^2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$$

$$A^3 = AAA = AA^2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix}$$

e, em geral,

$$A^n = \begin{pmatrix} 2^{n-1} & 2^{n-1} \\ 2^{n-1} & 2^{n-1} \end{pmatrix}$$

□

APLICAÇÃO 3

Em uma determinada cidade, 30% das mulheres casadas se divorciam e 20% das mulheres solteiras se casam por ano. Existem 8000 mulheres casadas e 2000 mulheres solteiras. Supondo que a população total de mulheres permanece constante, quantas mulheres estarão casadas e quantas estarão solteiras depois de 1 ano? E depois de 2 anos?

SOLUÇÃO. Forme uma matriz A da maneira descrita a seguir. Os elementos na primeira coluna de A são os percentuais das mulheres casadas e solteiras, respectivamente, que estão casadas 1 ano depois. Então

$$A = \begin{pmatrix} 0{,}70 & 0{,}20 \\ 0{,}30 & 0{,}80 \end{pmatrix}$$

Se $X = \begin{pmatrix} 8000 \\ 2000 \end{pmatrix}$, o número de mulheres casadas e solteiras depois de 1 ano pode ser obtido multiplicando-se A por X.

$$AX = \begin{pmatrix} 0{,}70 & 0{,}20 \\ 0{,}30 & 0{,}80 \end{pmatrix} \begin{pmatrix} 8000 \\ 2000 \end{pmatrix} = \begin{pmatrix} 6000 \\ 4000 \end{pmatrix}$$

Depois de 1 ano, 6000 mulheres estarão casadas e 4000 estarão solteiras. Para encontrar o número de mulheres casadas e solteiras depois de 2 anos, calcule

$$A^2X = A(AX) = \begin{pmatrix} 0{,}70 & 0{,}20 \\ 0{,}30 & 0{,}80 \end{pmatrix} \begin{pmatrix} 6000 \\ 4000 \end{pmatrix} = \begin{pmatrix} 5000 \\ 5000 \end{pmatrix}$$

Depois de 2 anos metade das mulheres estará casada e metade estará solteira. □

A TRANSPOSTA DE UMA MATRIZ

Dada uma matriz A $m \times n$, é muitas vezes útil formar uma nova matriz $n \times m$ cujas colunas são as linhas de A.

Definição. A **transposta** de uma matriz A $m \times n$ é a matriz B $n \times m$ definida por

(3) $$b_{ji} = a_{ij}$$

para $j = 1, \ldots, n$ e $i = 1, \ldots, m$. A transposta de A é denotada por A^T.

Segue de (3) que a j-ésima linha de A^T tem os mesmos elementos, respectivamente, que a j-ésima coluna de A e que a i-ésima coluna de A^T tem os mesmos elementos, respectivamente, que a i-ésima linha de A.

EXEMPLO 9.

(a) Se $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$, então $A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$.

(b) Se $B = \begin{pmatrix} -3 & 2 & 1 \\ 4 & 3 & 2 \\ 1 & 2 & 5 \end{pmatrix}$, então $B^T = \begin{pmatrix} -3 & 4 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 5 \end{pmatrix}$.

(c) Se $C = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$, então $C^T = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$.

A matriz C é igual à sua transposta. □

Existem quatro regras algébricas envolvendo transpostas.

Regras Algébricas para Transpostas

1. $(A^T)^T = A$
2. $(\alpha A)^T = \alpha A^T$
3. $(A + B)^T = A^T + B^T$
4. $(AB)^T = B^T A^T$

Vamos provar a quarta regra e deixar as três primeiras a cargo do leitor. Se A é uma matriz $m \times n$, então, para que a multiplicação seja possível, B tem que ter n linhas. Seja $C = AB$ e denote o elemento (i, j) de A^T, B^T e C^T, respectivamente, por a_{ij}^*, b_{ij}^* e c_{ij}^*. Então

$$c_{ij}^* = c_{ji}, \quad a_{ij}^* = a_{ji}, \quad b_{ij}^* = b_{ji}$$

O elemento (i, j) de B^TA^T é dado por

$$\sum_{k=1}^{n} b_{ik}^* a_{kj}^* = \sum_{k=1}^{n} a_{jk} b_{ki}$$

O elemento (i, j) de $C^T = (AB)^T$ é dado por

$$c_{ij}^* = c_{ji} = \sum_{k=1}^{n} a_{jk} b_{ki}$$

Logo, $(AB)^T = B^TA^T$.

A matriz C no Exemplo 9 é sua própria transposta. Isso acontece muitas vezes com matrizes em aplicações.

Definição. Uma matriz A $n \times n$ é dita **simétrica** se $A^T = A$.

As matrizes a seguir são todas simétricas:

$$\begin{pmatrix} 1 & 0 \\ 0 & -4 \end{pmatrix} \quad \begin{pmatrix} 2 & 3 & 4 \\ 3 & 1 & 5 \\ 4 & 5 & 3 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & -2 \\ 2 & -2 & -3 \end{pmatrix}$$

Um tipo de aplicação que nos leva a matrizes simétricas são problemas envolvendo redes. Esses problemas são muitas vezes resolvidos usando-se técnicas de uma área da matemática chamada *teoria dos grafos*.

APLICAÇÃO 4: REDES E GRAFOS

A teoria dos grafos é uma das áreas importantes da matemática aplicada. É usada para modelar problemas em praticamente todas as ciências aplicadas. A teoria dos grafos é particularmente útil em aplicações envolvendo redes de comunicação.

Um *grafo* é definido como um conjunto de pontos chamados *vértices* junto com um conjunto de pares não-ordenados de vértices chamados de *arestas*. A Fig. 1.3.1 dá uma representação geométrica de um grafo. Podemos pensar nos vértices V_1, V_2, V_3, V_4, V_5 como correspondendo a nós em uma rede de comunicação. Os segmentos de reta unindo os vértices correspondem às arestas: $\{V_1, V_2\}$, $\{V_2, V_5\}$, $\{V_3, V_4\}$, $\{V_3, V_5\}$, $\{V_4, V_5\}$. Cada aresta representa um elo de comunicação direta entre dois nós da rede.

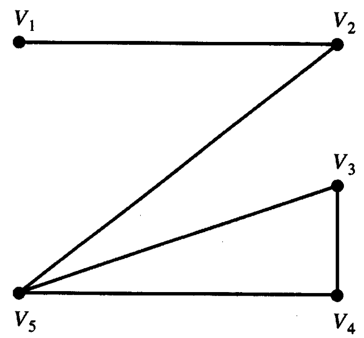


FIG. 1.3.1

Uma rede de comunicação verdadeira pode envolver um grande número de vértices e arestas. De fato, se existem milhões de vértices, uma representação gráfica da rede seria muito confusa. Uma alternativa é usar uma representação matricial para a rede. Se o grafo contém um total de n vértices, podemos definir uma matriz A $n \times n$ por

$$a_{ij} = \begin{cases} 1 & \text{Se } \{V_i, V_j\} \text{ é uma aresta do grafo} \\ 0 & \text{Se não existe aresta conectando } V_i \text{ e } V_j \end{cases}$$

A matriz A é chamada a *matriz de adjacência* do grafo. A matriz de adjacência para o grafo da Fig. 1.3.1 é dada por

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

Observe que a matriz A é simétrica. De fato, qualquer matriz de adjacência tem que ser simétrica, pois se $\{V_i, V_j\}$ é uma aresta do grafo, então $a_{ij} = a_{ji} = 1$, e se não existe aresta conectando V_i a V_j, então $a_{ij} = a_{ji} = 0$. Em qualquer dos casos, $a_{ij} = a_{ji}$.

Podemos pensar em um *caminho* no grafo como uma seqüência de arestas unindo um vértice a outro. Por exemplo, na Fig. 1.3.1, as arestas $\{V_1, V_2\}$, $\{V_2, V_5\}$ representam um caminho do vértice V_1 ao vértice V_5. Dizemos que o comprimento do caminho é 2, já que ele consiste em duas arestas. Um modo simples de descrever o caminho é indicar o movimento entre os vértices usando setas. Então, $V_1 \to V_2 \to V_5$ denota um caminho de comprimento 2 de V_1 a V_5. Analogamente, $V_4 \to V_5 \to V_2 \to V_1$ representa um caminho de comprimento 3 de V_4 a V_1. É possível passar pela mesma aresta mais de uma vez em um caminho. Por exemplo, $V_5 \to V_3 \to V_5 \to V_3$ é um caminho de comprimento 3 de V_5 a V_3. Em geral, calculando as potências da matriz de adjacência, podemos determinar o número de caminhos de qualquer comprimento especificado entre dois vértices.

Teorema 1.3.2. *Se A é uma matriz de adjacência $n \times n$ de um grafo e se $a_{ij}^{(k)}$ representa o elemento (i, j) de A^k, então $a_{ij}^{(k)}$ é igual ao número de caminhos de comprimento k de V_i a V_j.*

Demonstração. A demonstração é por indução matemática. No caso $k = 1$, segue, da própria definição de matriz de adjacência, que a_{ij} representa o número de caminhos de comprimento 1 de V_i a V_j. Suponha que, para algum m, cada elemento de A^m é igual ao número de caminhos de comprimento m entre os vértices correspondentes. Então, $a_{il}^{(m)}$ é o número de caminhos de comprimento m de V_i a V_l. Se existe uma aresta $\{V_l, V_j\}$, então $a_{il}^{(m)}a_{lj} = a_{il}^{(m)}$ é o número de caminhos de comprimento $m + 1$ de V_i a V_j da forma

$$V_i \to \cdots \to V_l \to V_j$$

Temos, então, que o número total de caminhos de comprimento $m + 1$ de V_i a V_j é dado por

$$a_{i1}^{(m)}a_{1j} + a_{i2}^{(m)}a_{2j} + \cdots + a_{in}^{(m)}a_{nj}$$

Mas isso é simplesmente o elemento (i, j) de A^{m+1}. □

EXEMPLO 10. Para determinar o número de caminhos de comprimento 3 entre dois vértices quaisquer do grafo da Fig. 1.3.1, precisamos apenas calcular

$$A^3 = \begin{pmatrix} 0 & 2 & 1 & 1 & 0 \\ 2 & 0 & 1 & 1 & 4 \\ 1 & 1 & 2 & 3 & 4 \\ 1 & 1 & 3 & 2 & 4 \\ 0 & 4 & 4 & 4 & 2 \end{pmatrix}$$

Portanto, o número de caminhos de tamanho 3 de V_3 a V_5 é $a_{35}^{(3)} = 4$. Observe que a matriz A^3 é simétrica. Isso reflete o fato de que existe o mesmo número de caminhos de comprimento 3 de V_i a V_j que de V_j a V_i. □

EXERCÍCIOS

1. Se

$$A = \begin{pmatrix} 3 & 1 & 4 \\ -2 & 0 & 1 \\ 1 & 2 & 2 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 1 & 0 & 2 \\ -3 & 1 & 1 \\ 2 & -4 & 1 \end{pmatrix}$$

calcule:

(a) $2A$ (b) $A + B$ (c) $2A - 3B$
(d) $(2A)^T - (3B)^T$ (e) AB (f) BA
(g) $A^T B^T$ (h) $(BA)^T$

2. Para cada um dos pares de matrizes dados a seguir, determine se é ou não possível efetuar a multiplicação da primeira matriz pela segunda. Se for possível, efetue a multiplicação.

(a) $\begin{pmatrix} 3 & 5 & 1 \\ -2 & 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 3 \\ 4 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 4 & -2 \\ 6 & -4 \\ 8 & -6 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 4 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 1 & 1 \\ 4 & 5 \end{pmatrix}$ (d) $\begin{pmatrix} 4 & 6 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 & 5 \\ 4 & 1 & 6 \end{pmatrix}$

(e) $\begin{pmatrix} 4 & 6 & 1 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 & 5 \\ 4 & 1 & 6 \end{pmatrix}$

3. Para cada um dos pares no Exercício 2 é possível multiplicar a segunda matriz pela primeira. Qual o tamanho da matriz produto?

4. Escreva cada um dos sistemas a seguir como uma equação matricial.

(a) $\begin{aligned} 3x_1 + 2x_2 &= 1 \\ 2x_1 - 3x_2 &= 5 \end{aligned}$ (b) $\begin{aligned} x_1 + x_2 \phantom{{}+2x_3} &= 5 \\ 2x_1 + x_2 - x_3 &= 6 \\ 3x_1 - 2x_2 + 2x_3 &= 7 \end{aligned}$ (c) $\begin{aligned} 2x_1 + x_2 + x_3 &= 4 \\ x_1 - x_2 + 2x_3 &= 2 \\ 3x_1 - 2x_2 - x_3 &= 0 \end{aligned}$

5. Se

$$A = \begin{pmatrix} 3 & 4 \\ 1 & 1 \\ 2 & 7 \end{pmatrix}$$

verifique que:

(a) $5A = 3A + 2A$ (b) $6A = 3(2A)$ (c) $(A^T)^T = A$

6. Se

$$A = \begin{pmatrix} 4 & 1 & 6 \\ 2 & 3 & 5 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 1 & 3 & 0 \\ -2 & 2 & -4 \end{pmatrix}$$

verifique que:

(a) $A + B = B + A$ (b) $3(A + B) = 3A + 3B$ (c) $(A + B)^T = A^T + B^T$

7. Se

$$A = \begin{pmatrix} 2 & 1 \\ 6 & 3 \\ -2 & 4 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 2 & 4 \\ 1 & 6 \end{pmatrix}$$

verifique que

(a) $3(AB) = (3A)B = A(3B)$ (b) $(AB)^T = B^T A^T$

8. Se

$$A = \begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} -2 & 1 \\ 0 & 4 \end{pmatrix}, \quad C = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$$

verifique que:

(a) $(A + B) + C = A + (B + C)$
(b) $(AB)C = A(BC)$
(c) $A(B + C) = AB + AC$
(d) $(A + B)C = AC + BC$

9. Prove a associatividade para a multiplicação de matrizes 2×2, isto é, considere

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}, \quad C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

e mostre que

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

10. Seja

$$A = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Calcule A^2 e A^3. O que deve ser A^n?

11. Seja

$$A = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Calcule A^2 e A^3. O que devem ser A^{2n} e A^{2n+1}?

12. Seja

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Mostre que $A^n = O$ para $n \geq 4$.

13. Encontre matrizes A e B 2×2 diferentes da matriz nula para as quais $AB = O$.

14. Encontre matrizes não-nulas A, B, C tais que

$$AC = BC \quad \text{e} \quad A \neq B$$

15. A matriz

$$A = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$$

tem a propriedade que $A^2 = O$. É possível para uma matriz simétrica 2×2 ter essa propriedade? Prove sua resposta.

16. O produto de duas matrizes simétricas é necessariamente simétrico? Prove sua resposta.

17. Seja A uma matriz $m \times n$.

(a) Explique por que é possível efetuar as multiplicações A^TA e AA^T.
(b) Mostre que A^TA e AA^T são ambas simétricas.

18. Sejam A e B matrizes simétricas $n \times n$. Prove que $AB = BA$ se e somente se AB também é simétrica.

19. Na Aplicação 1, suponha que João perdeu 4 quilos. Se ele continuar com o mesmo programa de exercícios, quantas calorias vai queimar a cada dia?

20. Na Aplicação 3, quantas mulheres estarão casadas e quantas estarão solteiras depois de 3 anos?

21. Considere a matriz

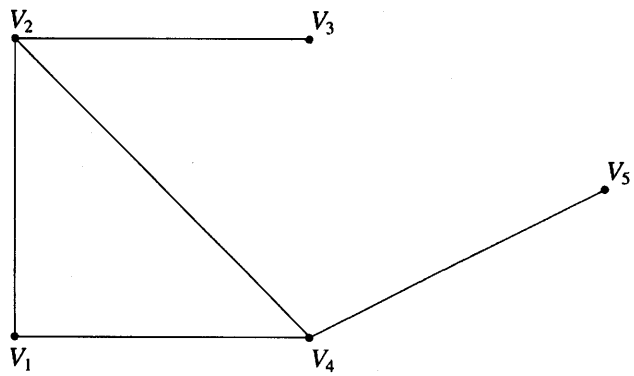
$$A = \begin{pmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix}$$

(a) Desenhe um grafo que tenha A como matriz de adjacência. Não se esqueça de marcar os vértices no gráfico.

(b) Analisando o grafo, determine o número de caminhos de comprimento 2 de V_2 a V_3 e de V_2 a V_5.

(c) Calcule a segunda linha de A^3 e use-a para determinar o número de caminhos de comprimento 3 de V_2 a V_3 e de V_2 a V_5.

22. Considere o grafo



(a) Encontre a matriz de adjacência A do grafo.

(b) Calcule A^2. O que os elementos da primeira linha de A^2 lhe dizem sobre os caminhos de comprimento 2 que começam em V_1?

(c) Calcule A^3. Quantos caminhos de comprimento 3 existem de V_2 a V_4? Quantos caminhos de comprimento menor ou igual a 3 existem de V_2 a V_4?

23. Seja A uma matriz 2×2 com $a_{11} \neq 0$ e seja $a = a_{21}/a_{11}$. Mostre que A pode ser fatorada em um produto da forma

$$\begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ 0 & b \end{pmatrix}$$

Qual o valor de b?

4 TIPOS ESPECIAIS DE MATRIZES

Vamos estudar, nesta seção, tipos especiais de matrizes, como matrizes triangulares, diagonais e elementares. Esses tipos especiais de matrizes têm um papel importante na solução de equações matriciais. Começamos considerando uma matriz especial I que age como a identidade multiplicativa, isto é,

$$IA = AI = A$$

para toda matriz A $n \times n$. Vamos discutir, também, a existência e o cálculo de inversas multiplicativas.

A MATRIZ IDENTIDADE

Uma matriz muito importante é a matriz I $n \times n$ com 1 ao longo da diagonal e 0 fora da diagonal. Então, $I = (\delta_{ij})$, onde

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

Se A é qualquer matriz $n \times n$, $AI = IA = A$. A matriz I age como uma identidade para a multiplicação de matrizes $n \times n$ e conseqüentemente é denominada *matriz identidade*. Por exemplo,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 4 & 1 \\ 2 & 6 & 3 \\ 0 & 1 & 8 \end{pmatrix} = \begin{pmatrix} 3 & 4 & 1 \\ 2 & 6 & 3 \\ 0 & 1 & 8 \end{pmatrix}$$

e

$$\begin{pmatrix} 3 & 4 & 1 \\ 2 & 6 & 3 \\ 0 & 1 & 8 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 4 & 1 \\ 2 & 6 & 3 \\ 0 & 1 & 8 \end{pmatrix}$$

Em geral, se B é uma matriz $m \times n$ qualquer e C é uma matriz $n \times r$ qualquer, então

$$BI = B \quad \text{e} \quad IC = C$$

Notação. O conjunto de todas as n-uplas de números reais é chamado de *espaço euclidiano de dimensão n* e é denotado por R^n. Os elementos de R^n são chamados de *vetores*. Entretanto, observe que a solução da equação matricial $AX = B$ é uma matriz $n \times 1$, e não uma n-upla. Em geral, ao trabalhar com equações matriciais, é mais conveniente pensar em cada elemento de R^n como um vetor coluna (matriz $n \times 1$), em vez de vetor linha (matriz $1 \times n$). A notação padrão para um vetor coluna é uma letra minúscula em negrito:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \mathbf{x}^T = (x_1, \ldots, x_n)$$

vetor coluna vetor linha

Seguindo essa convenção, vamos passar a usar a notação $A\mathbf{x} = \mathbf{b}$, em vez de $AX = B$, para representar um sistema de equações lineares.

Dada uma matriz A $m \times n$, é muitas vezes necessário fazer referência a uma determinada linha ou coluna. Vamos denotar a i-ésima linha de A por $\mathbf{a}(i, :)$ e a j-ésima coluna por $\mathbf{a}(:, j)$. Vamos trabalhar principalmente com colunas. Por essa razão, vamos simplificar a notação usando $\mathbf{a}_j$ no lugar de $\mathbf{a}(:, j)$. Como as referências a vetores linhas são bem menos freqüentes, não simplificaremos a notação para vetores linhas. Resumindo, se A é uma matriz $m \times n$, as linhas de A são dadas por

$$\mathbf{a}(i, :) = (a_{i1}, a_{i2}, \ldots, a_{in}) \qquad i = 1, \ldots, m$$

e as colunas, por

$$\mathbf{a}_j = \mathbf{a}(:, j) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \qquad j = 1, \ldots, n$$

Analogamente, se B é uma matriz $n \times r$, então $B = (\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_r)$. A única exceção é a matriz identidade. A notação padrão para a j-ésima coluna de I é $\mathbf{e}_j$, em vez de $\mathbf{i}_j$. A matriz identidade $n \times n$, então, é escrita na forma

$$I = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$$

MATRIZES DIAGONAIS E TRIANGULARES

Uma matriz A $n \times n$ é dita *triangular superior* se $a_{ij} = 0$ sempre que $i > j$; ela é *triangular inferior* se $a_{ij} = 0$ sempre que $i < j$. Além disso, A é simplesmente *triangular* se for triangular superior ou inferior. Por exemplo, as matrizes

$$\begin{pmatrix} 3 & 2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 5 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} 1 & 0 & 0 \\ 6 & 2 & 0 \\ 1 & 4 & 3 \end{pmatrix}$$

são ambas triangulares. A primeira é triangular superior, e a segunda é triangular inferior.

Uma matriz triangular pode ter 0 na diagonal. No entanto, para um sistema linear $A\mathbf{x} = \mathbf{b}$ estar em forma triangular, a matriz de coeficientes A tem que ser triangular sem elementos nulos na diagonal.

Uma matriz A $n \times n$ é *diagonal* se $a_{ij} = 0$ sempre que $i \neq j$. As matrizes

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

são todas diagonais. Uma matriz diagonal é, ao mesmo tempo, triangular superior e inferior.

INVERSÃO DE MATRIZES

Definição. Uma matriz A $n \times n$ é dita **invertível** ou **não-singular** se existe uma matriz B tal que $AB = BA = I$. A matriz B é uma inversa multiplicativa de A.

Se B e C são ambas inversas multiplicativas de A, então

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

Logo, uma matriz pode ter, no máximo, uma inversa multiplicativa. Vamos chamar a inversa multiplicativa de uma matriz não-singular A simplesmente de *inversa* de A e denotá-la por A^{-1}.

EXEMPLO 1. As matrizes

$$\begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} -\frac{1}{10} & \frac{2}{5} \\ \frac{3}{10} & -\frac{1}{5} \end{pmatrix}$$

são inversas uma da outra, já que

$$\begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} -\frac{1}{10} & \frac{2}{5} \\ \frac{3}{10} & -\frac{1}{5} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

e

$$\begin{pmatrix} -\frac{1}{10} & \frac{2}{5} \\ \frac{3}{10} & -\frac{1}{5} \end{pmatrix} \begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

□

EXEMPLO 2. As matrizes triangulares

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} 1 & -2 & 5 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix}$$

são inversas uma da outra, já que

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -2 & 5 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

e

$$\begin{pmatrix} 1 & -2 & 5 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

□

EXEMPLO 3. A matriz

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

não tem inversa. De fato, se B é qualquer matriz 2×2, então

$$BA = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} b_{11} & 0 \\ b_{21} & 0 \end{pmatrix}$$

Logo, BA não pode ser igual a I. □

Definição. Uma matriz é dita **singular** ou **não-invertível** se ela não tem uma inversa multiplicativa.

SISTEMAS EQUIVALENTES

Dado um sistema linear $n \times n$ $A\mathbf{x} = \mathbf{b}$, podemos obter um sistema equivalente multiplicando ambos os lados da equação matricial por uma matriz invertível M.

(1) $$A\mathbf{x} = \mathbf{b}$$

(2) $$MA\mathbf{x} = M\mathbf{b}$$

É claro que qualquer solução de (1) é também uma solução de (2). Por outro lado, se $\hat{\mathbf{x}}$ é uma solução de (2), então

$$M^{-1}(MA\hat{\mathbf{x}}) = M^{-1}(M\mathbf{b})$$

$$A\hat{\mathbf{x}} = \mathbf{b}$$

de modo que os dois sistemas são equivalentes.

Para chegar a um sistema equivalente que seja mais fácil de resolver, podemos multiplicar os dois lados da equação $A\mathbf{x} = \mathbf{b}$ por uma série de matrizes invertíveis $E_1, \ldots, E_k$, obtendo

$$U\mathbf{x} = \mathbf{c}$$

onde $U = E_k \cdots E_1 A$ e $\mathbf{c} = E_k \cdots E_1 \mathbf{b}$. Esse novo sistema vai ser equivalente ao sistema original desde que $M = E_k \cdots E_1$ seja invertível. Entretanto, M é um produto de matrizes invertíveis. O teorema a seguir mostra que qualquer produto de matrizes invertíveis é invertível.

Teorema 1.4.1. *Se A e B são matrizes invertíveis $n \times n$, então AB é invertível e $(AB)^{-1} = B^{-1}A^{-1}$.*

Demonstração

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

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

□

Por indução, segue que se $E_1, \ldots, E_k$ são todas invertíveis, então o produto $E_1E_2 \cdots E_k$ é invertível e

$$(E_1E_2 \cdots E_k)^{-1} = E_k^{-1} \cdots E_2^{-1}E_1^{-1}$$

Vamos mostrar a seguir que podemos efetuar qualquer uma das operações elementares sobre as linhas de uma matriz A multiplicando A por uma matriz invertível à esquerda.

MATRIZES ELEMENTARES

Uma matriz obtida a partir da matriz identidade I por uma das operações elementares é chamada de matriz *elementar*.

Existem três tipos de matrizes elementares, uma para cada operação elementar.

Tipo I

Uma matriz elementar do tipo I é obtida trocando-se a ordem de duas linhas de I.

EXEMPLO 4. Seja

$$E_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

E_1 é uma matriz elementar do tipo I, já que foi obtida trocando-se as duas primeiras linhas de I. Seja A uma matriz 3×3.

$$E_1A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$AE_1 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{12} & a_{11} & a_{13} \\ a_{22} & a_{21} & a_{23} \\ a_{32} & a_{31} & a_{33} \end{pmatrix}$$

Multiplicando A à esquerda por E_1, trocamos as duas primeiras linhas de A. Multiplicar A à direita por E_1 equivale a efetuar a operação elementar sobre colunas que consiste na troca das duas primeiras colunas de A. □

Tipo II

Uma matriz elementar do tipo II é uma matriz obtida multiplicando-se uma linha de I por uma constante não-nula.

EXEMPLO 5.

$$E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

é uma matriz elementar do tipo II.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 3a_{31} & 3a_{32} & 3a_{33} \end{pmatrix}$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & 3a_{13} \\ a_{21} & a_{22} & 3a_{23} \\ a_{31} & a_{32} & 3a_{33} \end{pmatrix}$$

A multiplicação à esquerda por E_2 efetua a operação elementar sobre as linhas que consiste em multiplicar a terceira linha por 3, enquanto a multiplicação à direita por E_2 efetua a operação elementar sobre as colunas que consiste em multiplicar a terceira coluna por 3. □

Tipo III

Uma matriz elementar de tipo III é uma matriz obtida de I somando-se um múltiplo de uma das linhas à outra linha.

EXEMPLO 6.

$$E_3 = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

é uma matriz elementar do tipo III. Se A é uma matriz 3×3, então

$$E_3A = \begin{pmatrix} a_{11} + 3a_{31} & a_{12} + 3a_{32} & a_{13} + 3a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$AE_3 = \begin{pmatrix} a_{11} & a_{12} & 3a_{11} + a_{13} \\ a_{21} & a_{22} & 3a_{21} + a_{23} \\ a_{31} & a_{32} & 3a_{31} + a_{33} \end{pmatrix}$$

Multiplicação à esquerda por E_3 soma 3 vezes a terceira linha à segunda. Multiplicação à direita por E_3 soma 3 vezes a primeira coluna à terceira. □

Em geral, suponha que E é uma matriz elementar $n \times n$. Podemos pensar em E como sendo obtida de I por uma operação elementar sobre as linhas ou sobre as colunas. Se A é uma matriz $n \times r$, multiplicar A por E à esquerda tem o efeito de efetuar a mesma operação sobre as linhas de A. Se B é uma matriz $m \times n$, multiplicar B por E à direita equivale a efetuar a mesma operação sobre as colunas de B.

Teorema 1.4.2. *Se E é uma matriz elementar, então E é invertível e E^{-1} é uma matriz elementar do mesmo tipo.*

Demonstração. Se E é uma matriz elementar do tipo I obtida trocando-se as i-ésima e j-ésima linhas, então podemos transformar E de volta em I trocando as mesmas linhas novamente. Isso significa que $EE = I$ e, portanto, E é sua própria inversa. Se E é uma matriz elementar do tipo II obtida multiplicando-se a i-ésima linha de A por um escalar não-nulo α, então E pode ser transformada de volta na identidade por uma multiplicação da sua i-ésima linha ou coluna por $1/\alpha$. Então

$$E^{-1} = \begin{array}{cl} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & O & \\ & & 1 & & & & \\ & & & 1/\alpha & & & \\ & & & & 1 & & \\ & O & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} & \begin{array}{l} \\ \\ \\ i\text{-ésima linha} \\ \\ \\ \\ \end{array} \end{array}$$

Por fim, suponha que E é uma matriz elementar do tipo III obtida de I somando-se m vezes a i-ésima linha à j-ésima linha.

$$E = \begin{array}{cl} \begin{pmatrix} 1 & & & & & & \\ \vdots & \ddots & & & & & O \\ 0 & \cdots & 1 & & & & \\ \vdots & & & \ddots & & & \\ 0 & \cdots & m & \cdots & 1 & & \\ \vdots & & & & & \ddots & \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} & \begin{array}{l} \\ \\ i\text{-ésima linha} \\ \\ j\text{-ésima linha} \\ \\ \\ \end{array} \end{array}$$

E pode ser transformada de volta em I subtraindo-se m vezes sua i-ésima linha de sua j-ésima linha, ou subtraindo-se m vezes sua j-ésima coluna de sua i-ésima coluna. Então,

$$E^{-1} = \begin{pmatrix} 1 & & & & & & \\ \vdots & \ddots & & & & O & \\ 0 & \cdots & 1 & & & & \\ \vdots & & & \ddots & & & \\ 0 & \cdots & -m & \cdots & 1 & & \\ \vdots & & & & & \ddots & \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

□

Definição. Uma matriz B é **equivalente por linhas** a A se existe uma seqüência finita de matrizes elementares $E_1, E_2, \ldots, E_k$ tal que

$$B = E_k E_{k-1} \cdots E_1 A$$

Em outras palavras, B é equivalente por linhas a A se B puder ser obtida de A por um número finito de operações elementares. Em particular, duas matrizes aumentadas $(A|\mathbf{b})$ e $(B|\mathbf{c})$ são equivalentes por linhas se e somente se os sistemas $A\mathbf{x} = \mathbf{b}$ e $B\mathbf{x} = \mathbf{c}$ são equivalentes.

As propriedades a seguir, para matrizes equivalentes por linhas, são conseqüências do Teorema 1.4.2.

(i) Se A é equivalente por linhas a B, então B é equivalente por linhas a A.
(ii) Se A é equivalente por linhas a B e B é equivalente por linhas a C, então A é equivalente por linhas a C.

Os detalhes das demonstrações de (i) e (ii) são deixados a cargo do leitor.

Teorema 1.4.3. *Seja A uma matriz $n \times n$. Então, as seguintes afirmações são equivalentes:*

(a) *A é invertível;*
(b) *$A\mathbf{x} = \mathbf{0}$ tem apenas a solução trivial $\mathbf{0}$;*
(c) *A é equivalente por linhas a I.*

Demonstração. Vamos provar primeiro que (a) implica (b). Se A é invertível e $\hat{\mathbf{x}}$ é uma solução de $A\mathbf{x} = \mathbf{0}$, então

$$\hat{\mathbf{x}} = I\hat{\mathbf{x}} = (A^{-1}A)\hat{\mathbf{x}} = A^{-1}(A\hat{\mathbf{x}}) = A^{-1}\mathbf{0} = \mathbf{0}$$

Logo, $A\mathbf{x} = \mathbf{0}$ tem apenas a solução trivial. Vamos mostrar agora que (b) implica (c). Usando operações elementares, o sistema pode ser transformado em um sistema da forma $U\mathbf{x} = \mathbf{0}$, onde U está em forma escada. Se um dos elementos da diagonal de U fosse igual a zero, a última linha de U seria nula. Mas então $A\mathbf{x} = \mathbf{0}$ seria equivalente a um sistema com mais incógnitas que equações e portanto, pelo Teorema 1.2.1, teria uma solução não-trivial. Logo, U tem que ser uma matriz triangular com todos os elementos diagonais iguais a 1. Segue, então, que I é a forma escada reduzida por linha de A e, portanto, A é equivalente por linhas a I.

Finalmente, vamos provar que (c) implica (a). Se A é equivalente por linhas a I, existem matrizes elementares $E_1, E_2, \ldots, E_k$ tais que

$$A = E_k E_{k-1} \cdots E_1 I = E_k E_{k-1} \cdots E_1$$

Como E_i é invertível para $i = 1, \ldots, k$, o produto $E_k E_{k-1} \cdots E_1$ também é invertível. Logo, A é não-singular e

$$A^{-1} = (E_k E_{k-1} \cdots E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_k^{-1}$$

□

Corolário 1.4.4. *O sistema n de equações lineares com n incógnitas $A\mathbf{x} = \mathbf{b}$ tem uma única solução se e somente se A é invertível.*

Demonstração. Se A é invertível, então $A^{-1}\mathbf{b}$ é a única solução de $A\mathbf{x} = \mathbf{b}$. Por outro lado, suponha que $A\mathbf{x} = \mathbf{b}$ tem uma única solução $\hat{\mathbf{x}}$. Se A fosse singular, $A\mathbf{x} = \mathbf{0}$ teria uma solução $\mathbf{z} \neq \mathbf{0}$. Seja $\mathbf{y} = \hat{\mathbf{x}} + \mathbf{z}$. É claro que $\mathbf{y} \neq \hat{\mathbf{x}}$ e

$$A\mathbf{y} = A(\hat{\mathbf{x}} + \mathbf{z}) = A\hat{\mathbf{x}} + A\mathbf{z} = \mathbf{b} + \mathbf{0} = \mathbf{b}$$

Então $\mathbf{y}$ é também uma solução de $A\mathbf{x} = \mathbf{b}$, o que é uma contradição. Portanto, se $A\mathbf{x} = \mathbf{b}$ tem uma única solução, A tem que ser invertível. □

Se A é invertível, então A é equivalente por linhas a I, logo existem matrizes elementares $E_1, \ldots, E_k$ tais que

$$E_k E_{k-1} \cdots E_1 A = I$$

Multiplicando ambos os lados dessa equação à direita por A^{-1}, obtemos

$$E_k E_{k-1} \cdots E_1 I = A^{-1}$$

Logo, a mesma seqüência de operações elementares que transforma uma matriz invertível A em I transforma I em A^{-1}. Isso nos dá um método para calcular A^{-1}. Aumentando a matriz A com I e efetuando as operações elementares que transformam A em I sobre as linhas da matriz aumentada, I vai ser transformada em A^{-1}. Em outras palavras, a forma escada reduzida por linhas da matriz aumentada $(A|I)$ é $(I|A^{-1})$.

EXEMPLO 7. Calcule A^{-1} se

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

SOLUÇÃO

$$\left(\begin{array}{rrr|rrr} 1 & 4 & 3 & 1 & 0 & 0 \\ -1 & -2 & 0 & 0 & 1 & 0 \\ 2 & 2 & 3 & 0 & 0 & 1 \end{array}\right) \rightarrow \left(\begin{array}{rrr|rrr} 1 & 4 & 3 & 1 & 0 & 0 \\ 0 & 2 & 3 & 1 & 1 & 0 \\ 0 & -6 & -3 & -2 & 0 & 1 \end{array}\right)$$

$$\rightarrow \left(\begin{array}{rrr|rrr} 1 & 4 & 3 & 1 & 0 & 0 \\ 0 & 2 & 3 & 1 & 1 & 0 \\ 0 & 0 & 6 & 1 & 3 & 1 \end{array}\right) \rightarrow \left(\begin{array}{rrr|rrr} 1 & 4 & 0 & \frac{1}{2} & -\frac{3}{2} & -\frac{1}{2} \\ 0 & 2 & 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 6 & 1 & 3 & 1 \end{array}\right)$$

$$\rightarrow \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 2 & 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 6 & 1 & 3 & 1 \end{array}\right) \rightarrow \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ 0 & 0 & 1 & \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \end{array}\right)$$

Portanto,

$$A^{-1} = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \end{pmatrix}$$

□

EXEMPLO 8. Resolva o sistema

$$\begin{aligned} x_1 + 4x_2 + 3x_3 &= 12 \\ -x_1 - 2x_2 &= -12 \\ 2x_1 + 2x_2 + 3x_3 &= 8 \end{aligned}$$

A matriz de coeficientes desse sistema é a matriz A do último exemplo. Logo, a solução do sistema é

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \end{pmatrix} \begin{pmatrix} 12 \\ -12 \\ 8 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ -\frac{8}{3} \end{pmatrix}$$

□

EXERCÍCIOS

1. Quais das matrizes a seguir são matrizes elementares? Classifique cada matriz elementar por tipo.

(a) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5 & 0 & 1 \end{pmatrix}$ (d) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

2. Encontre a inversa de cada uma das matrizes do Exemplo 1. Para cada matriz elementar, verifique que sua inversa é uma matriz elementar do mesmo tipo.

3. Para cada par de matrizes dado a seguir, encontre uma matriz elementar E tal que $EA = B$.

(a) $A = \begin{pmatrix} 2 & -1 \\ 5 & 3 \end{pmatrix}$, $B = \begin{pmatrix} -4 & 2 \\ 5 & 3 \end{pmatrix}$

(b) $A = \begin{pmatrix} 2 & 1 & 3 \\ -2 & 4 & 5 \\ 3 & 1 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 1 & 3 \\ 3 & 1 & 4 \\ -2 & 4 & 5 \end{pmatrix}$

(c) $A = \begin{pmatrix} 4 & -2 & 3 \\ 1 & 0 & 2 \\ -2 & 3 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 4 & -2 & 3 \\ 1 & 0 & 2 \\ 0 & 3 & 5 \end{pmatrix}$

4. Para cada par de matrizes dado a seguir, encontre uma matriz elementar E tal que $AE = B$.

(a) $A = \begin{pmatrix} 4 & 1 & 3 \\ 2 & 1 & 4 \\ 1 & 3 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 3 & 1 & 4 \\ 4 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix}$

(b) $A = \begin{pmatrix} 2 & 4 \\ 1 & 6 \end{pmatrix}$, $B = \begin{pmatrix} 2 & -2 \\ 1 & 3 \end{pmatrix}$

(c) $A = \begin{pmatrix} 4 & -2 & 3 \\ -2 & 4 & 2 \\ 6 & 1 & -2 \end{pmatrix}$, $B = \begin{pmatrix} 2 & -2 & 3 \\ -1 & 4 & 2 \\ 3 & 1 & -2 \end{pmatrix}$

5. Considere as matrizes

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 1 & 3 \\ 1 & 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 1 & 3 \\ 2 & 2 & 6 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 & 4 \\ 0 & -1 & -3 \\ 2 & 2 & 6 \end{pmatrix}$$

(a) Encontre uma matriz elementar E tal que $EA = B$.
(b) Encontre uma matriz elementar F tal que $FB = C$.
(c) C é equivalente por linhas a A? Explique.

6. Seja

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 6 & 4 & 5 \\ 4 & 1 & 3 \end{pmatrix}$$

(a) Encontre matrizes elementares E_1, E_2, E_3 tais que

$$E_3E_2E_1A = U$$

onde U é uma matriz triangular superior.

(b) Determine as inversas de E_1, E_2, E_3 e defina $L = E_1^{-1}E_2^{-1}E_3^{-1}$. Que tipo de matriz é L? Verifique que $A = LU$.

7. Seja

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

(a) Verifique que

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

(b) Use A^{-1} para resolver $A\mathbf{x} = \mathbf{b}$ para as seguintes escolhas de $\mathbf{b}$:

(i) $\mathbf{b} = (1, 1, 1)^T$
(ii) $\mathbf{b} = (1, 2, 3)^T$
(iii) $\mathbf{b} = (-2, 1, 0)^T$

8. Encontre a inversa de cada uma das matrizes a seguir.

(a) $\begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}$ (c) $\begin{pmatrix} 2 & 6 \\ 3 & 8 \end{pmatrix}$ (d) $\begin{pmatrix} 3 & 0 \\ 9 & 3 \end{pmatrix}$

(e) $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ (f) $\begin{pmatrix} 2 & 0 & 5 \\ 0 & 3 & 0 \\ 1 & 0 & 3 \end{pmatrix}$ (g) $\begin{pmatrix} -1 & -3 & -3 \\ 2 & 6 & 1 \\ 3 & 8 & 3 \end{pmatrix}$ (h) $\begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ -1 & -2 & -3 \end{pmatrix}$

9. Dadas

$$A = \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

calcule A^{-1} e use-a para:

(a) encontrar uma matriz X 2×2 tal que $AX = B$;

(b) encontrar uma matriz Y 2×2 tal que $YA = B$.

10. Considere as matrizes

$$A = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 6 & 2 \\ 2 & 4 \end{pmatrix}, \quad C = \begin{pmatrix} 4 & -2 \\ -6 & 3 \end{pmatrix}$$

Resolva cada uma das equações matriciais a seguir.

(a) $AX + B = C$
(b) $XA + B = C$
(c) $AX + B = X$
(d) $XA + C = X$

11. Seja

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Mostre que, se $d = a_{11}a_{22} - a_{21}a_{12} \neq 0$, então

$$A^{-1} = \frac{1}{d}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

12. Seja A uma matriz não-singular. Mostre que A^{-1} também é não-singular e que $(A^{-1})^{-1} = A$.

13. Prove que, se A é invertível, então A^T é invertível e

$$(A^T)^{-1} = (A^{-1})^T$$

[***Sugestão:*** $(AB)^T = B^TA^T$.]

14. Seja A uma matriz invertível $n \times n$. Use indução matemática para provar que A^m é invertível e que

$$(A^m)^{-1} = (A^{-1})^m$$

para $m = 1, 2, 3, ...$

15. A transposta de uma matriz elementar é uma matriz elementar do mesmo tipo? O produto de duas matrizes elementares é uma matriz elementar?

16. Seja U e R matrizes triangulares superiores $n \times n$ e seja $T = UR$. Mostre que T também é triangular superior e que $t_{jj} = u_{jj}r_{jj}$ para $j = 1, ..., n$.

17. Sejam A e B matrizes $n \times n$ e seja $C = AB$. Prove que, se B é singular, então C tem que ser singular.

[***Sugestão:*** Use o Teorema 1.4.3.]

18. Seja U uma matriz triangular superior com todos os elementos diagonais diferentes de zero.
(a) Explique por que U tem que ser invertível.
(b) Explique por que U^{-1} tem que ser triangular superior.

19. Sejam A uma matriz invertível $n \times n$ e B uma matriz $n \times r$. Mostre que a forma escada reduzida por linhas de $(A|B)$ é $(I|C)$, onde $C = A^{-1}B$.

20. Em geral, a multiplicação de matrizes não é comutativa (isto é, $AB \neq BA$). No entanto, existem certos casos especiais em que a comutatividade é válida. Mostre que:
(a) se D_1 e D_2 são matrizes diagonais, então $D_1D_2 = D_2D_1$;
(b) se A é uma matriz $n \times n$ e

$$B = a_0I + a_1A + a_2A^2 + \cdots + a_kA^k$$

onde $a_0, a_1, ..., a_k$ são escalares, então $AB = BA$.

21. Mostre que, se A é uma matriz simétrica invertível, então A^{-1} também é simétrica.

22. Prove que, se A é equivalente por linhas a B, então B é equivalente por linhas a A.

23. (a) Prove que, se A é equivalente por linhas a B e se B é equivalente por linhas a C, então A é equivalente por linhas a C.
(b) Prove que duas matrizes invertíveis $n \times n$ quaisquer são equivalentes por linha.

24. Prove que B é equivalente por linhas a A se e somente se existe uma matriz invertível M tal que $B = MA$.

25. Dado um vetor $\mathbf{x} \in R^{n+1}$, a matriz $V\,(n + 1) \times (n + 1)$, definida por

$$v_{ij} = \begin{cases} 1 & \text{se } j = 1 \\ x_i^{j-1} & \text{para } j = 2, \ldots, n+1 \end{cases}$$

é chamada de matriz de Vandermonde.
(a) Mostre que, se

$$V\mathbf{c} = \mathbf{y}$$

e

$$p(x) = c_1 + c_2x + \cdots + c_{n+1}x^n$$

então

$$p(x_i) = y_i, \quad i = 1, 2, \ldots, n+1$$

(b) Suponha que $x_1, x_2, ..., x_{n+1}$ são todos distintos. Mostre que, se $\mathbf{c}$ é uma solução de $V\mathbf{x} = \mathbf{0}$, então os coeficientes $c_1, c_2, ..., c_n$ têm que ser todos nulos e, portanto, V tem que ser invertível.

5 MATRIZES EM BLOCO

Muitas vezes é útil pensar em uma matriz como sendo composta de um número de submatrizes. Uma matriz A pode ser subdividida em matrizes menores desenhando-se retas horizontais entre as linhas e retas verticais entre as colunas. Por exemplo, considere a matriz

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

Se desenharmos retas entre a segunda e terceira linhas e entre a terceira e quarta colunas, dividiremos A em quatro submatrizes A_{11}, A_{12}, A_{21} e A_{22}.

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \left(\begin{array}{ccc|cc} 1 & -2 & 4 & 1 & 3 \\ 2 & 1 & 1 & 1 & 1 \\ \hline 3 & 3 & 2 & -1 & 2 \\ 4 & 6 & 2 & 2 & 4 \end{array}\right)$$

Um modo útil de dividir uma matriz é em colunas. Por exemplo, se

$$B = \begin{pmatrix} -1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 4 & 1 \end{pmatrix}$$

podemos dividir B em três submatrizes colunas:

$$B = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) = \left(\begin{array}{c|c|c} -1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 4 & 1 \end{array}\right)$$

Suponha que temos uma matriz

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

e que queremos calcular a segunda coluna de AB sem calcular o produto inteiro. Como a segunda coluna de AB é determinada por $\mathbf{b}_2$, basta calcular

$$A\mathbf{b}_2 = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 1 & -2 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 15 \\ -1 \end{pmatrix}$$

Em geral, se A é uma matriz $m \times n$ e $B = (\mathbf{b}_1, \ldots, \mathbf{b}_r)$ é uma matriz $n \times r$, então a j-ésima coluna de AB é $A\mathbf{b}_j$. Isso é uma conseqüência direta da definição de multiplicação de matrizes. Se $C = AB$, então

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

e portanto

$$\mathbf{c}_j = \begin{pmatrix} \sum a_{1k} b_{kj} \\ \sum a_{2k} b_{kj} \\ \vdots \\ \sum a_{mk} b_{kj} \end{pmatrix} = A\mathbf{b}_j$$

Logo,

$$\boxed{AB = (A\mathbf{b}_1, A\mathbf{b}_2, \ldots, A\mathbf{b}_r)}$$

Em particular,

$$(\mathbf{a}_1, \ldots, \mathbf{a}_n) = A = AI = (A\mathbf{e}_1, \ldots, A\mathbf{e}_n)$$

Seja A uma matriz $m \times n$. Dividindo A em linhas, obtemos

$$A = \begin{pmatrix} \mathbf{a}(1,:) \\ \mathbf{a}(2,:) \\ \vdots \\ \mathbf{a}(m,:) \end{pmatrix}$$

Se B é uma matriz $n \times r$, a i-ésima linha de AB é obtida multiplicando-se a i-ésima linha de A por B. Logo, a i-ésima linha de AB é $\mathbf{a}(i,:)B$. Em geral, o produto AB pode ser dividido em linhas da seguinte maneira:

$$AB = \begin{pmatrix} \mathbf{a}(1,:)B \\ \mathbf{a}(2,:)B \\ \vdots \\ \mathbf{a}(m,:)B \end{pmatrix}$$

Para ilustrar esse resultado, vamos ver um exemplo. Se

$$A = \begin{pmatrix} 2 & 5 \\ 3 & 4 \\ 1 & 7 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 3 & 2 & -3 \\ -1 & 1 & 1 \end{pmatrix}$$

então

$$\mathbf{a}(1,:)B = \begin{pmatrix} 1 & 9 & -1 \end{pmatrix}$$
$$\mathbf{a}(2,:)B = \begin{pmatrix} 5 & 10 & -5 \end{pmatrix}$$
$$\mathbf{a}(3,:)B = \begin{pmatrix} -4 & 9 & 4 \end{pmatrix}$$

Esses são os vetores linha do produto AB.

$$AB = \begin{pmatrix} \mathbf{a}(1,:)B \\ \mathbf{a}(2,:)B \\ \mathbf{a}(3,:)B \end{pmatrix} = \begin{pmatrix} 1 & 9 & -1 \\ 5 & 10 & -5 \\ -4 & 9 & 4 \end{pmatrix}$$

Vamos ver, agora, como calcular o produto AB em termos de divisões mais gerais de A e B. As submatrizes que dividem A e B são muitas vezes chamadas *blocos*.

MULTIPLICAÇÃO EM BLOCOS

Sejam A uma matriz $m \times n$ e B uma matriz $n \times r$. Muitas vezes é útil dividir A e B e escrever o produto AB em termos das submatrizes que dividem A e B. Vamos considerar quatro casos.

Caso 1

$B = (B_1 \;\; B_2)$, onde B_1 é uma matriz $n \times t$ e B_2 é uma matriz $n \times (r - t)$.

$$\begin{aligned} AB &= A(\mathbf{b}_1, \ldots, \mathbf{b}_t, \mathbf{b}_{t+1}, \ldots, \mathbf{b}_r) \\ &= (A\mathbf{b}_1, \ldots, A\mathbf{b}_t, A\mathbf{b}_{t+1}, \ldots, A\mathbf{b}_r) \\ &= \begin{pmatrix} AB_1 & AB_2 \end{pmatrix} \end{aligned}$$

Então,

$$\boxed{A\begin{pmatrix} B_1 & B_2 \end{pmatrix} = \begin{pmatrix} AB_1 & AB_2 \end{pmatrix}}$$

Caso 2

$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$, onde A_1 é uma matriz $k \times n$ e A_2 é uma matriz $(m-k) \times n$.

$$AB = \begin{pmatrix} \mathbf{a}(1,:) \\ \vdots \\ \mathbf{a}(k,:) \\ \mathbf{a}(k+1,:) \\ \vdots \\ \mathbf{a}(m,:) \end{pmatrix} B = \begin{pmatrix} \mathbf{a}(1,:)B \\ \vdots \\ \mathbf{a}(k,:)B \\ \mathbf{a}(k+1,:)B \\ \vdots \\ \mathbf{a}(m,:)B \end{pmatrix} = \begin{pmatrix} A_1B \\ A_2B \end{pmatrix}$$

Então,

$$\boxed{\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} B = \begin{pmatrix} A_1B \\ A_2B \end{pmatrix}}$$

Caso 3

$A = (A_1 \;\; A_2)$ e $B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$, onde A_1 é uma matriz $m \times s$, A_2 é uma matriz $m \times (n-s)$, B_1 é uma matriz $s \times r$ e B_2 é uma matriz $(n-s) \times r$. Se $C = AB$, então

$$c_{ij} = \sum_{l=1}^{n} a_{il}b_{lj} = \sum_{l=1}^{s} a_{il}b_{lj} + \sum_{l=s+1}^{n} a_{il}b_{lj}$$

Logo, c_{ij} é a soma do elemento (i,j) de A_1B_1 com o elemento (i,j) de A_2B_2. Portanto,

$$AB = C = A_1B_1 + A_2B_2$$

$$\boxed{\begin{pmatrix} A_1 & A_2 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = A_1B_1 + A_2B_2}$$

Caso 4

Sejam A e B matrizes divididas da seguinte forma:

$$A = \begin{array}{cc} \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right) & \begin{array}{l} k \\ m-k \end{array} \\ \begin{array}{cc} s & n-s \end{array} & \end{array}, \qquad B = \begin{array}{cc} \left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right) & \begin{array}{l} s \\ n-s \end{array} \\ \begin{array}{cc} t & r-t \end{array} & \end{array}$$

Sejam

$$A_1 = \begin{pmatrix} A_{11} \\ A_{21} \end{pmatrix} \qquad A_2 = \begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix}$$

$$B_1 = \begin{pmatrix} B_{11} & B_{12} \end{pmatrix} \qquad B_2 = \begin{pmatrix} B_{21} & B_{22} \end{pmatrix}$$

Pelo caso 3, tem-se que

$$AB = \begin{pmatrix} A_1 & A_2 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = A_1B_1 + A_2B_2$$

Dos casos 1 e 2, vê-se que

$$A_1B_1 = \begin{pmatrix} A_{11} \\ A_{21} \end{pmatrix} B_1 = \begin{pmatrix} A_{11}B_1 \\ A_{21}B_1 \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} & A_{11}B_{12} \\ A_{21}B_{11} & A_{21}B_{12} \end{pmatrix}$$

$$A_2B_2 = \begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix} B_2 = \begin{pmatrix} A_{12}B_2 \\ A_{22}B_2 \end{pmatrix} = \begin{pmatrix} A_{12}B_{21} & A_{12}B_{22} \\ A_{22}B_{21} & A_{22}B_{22} \end{pmatrix}$$

Portanto,

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}$$

Em geral, se os blocos têm as dimensões apropriadas, a multiplicação em bloco pode ser efetuada da mesma maneira que a operação usual entre matrizes.

Se

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1t} \\ \vdots & & \\ A_{s1} & \cdots & A_{st} \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} B_{11} & \cdots & B_{1r} \\ \vdots & & \\ B_{t1} & \cdots & B_{tr} \end{pmatrix}$$

então

$$AB = \begin{pmatrix} C_{11} & \cdots & C_{1r} \\ \vdots & & \\ C_{s1} & \cdots & C_{sr} \end{pmatrix}$$

onde

$$C_{ij} = \sum_{k=1}^{t} A_{ik}B_{kj}$$

A multiplicação só pode ser efetuada dessa maneira se o número de colunas de A_{ik} for igual ao número de linhas de B_{kj} para todo k.

EXEMPLO 1. Sejam

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 \\ 3 & 3 & 2 & 2 \end{pmatrix}$$

e

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ \hline 3 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 \end{array}\right)$$

Divida A em quatro blocos e efetue a multiplicação em bloco.

SOLUÇÃO. Como cada B_{kj} tem duas linhas, as matrizes A_{ik} têm que ter duas colunas. Existem duas possibilidades:

$$\text{(i)} \qquad \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ \hline 2 & 2 & 1 & 1 \\ 3 & 3 & 2 & 2 \end{array}\right)$$

e, nesse caso,

$$\left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ \hline 2 & 2 & 1 & 1 \\ 3 & 3 & 2 & 2 \end{array}\right)\left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ \hline 3 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 \end{array}\right) = \left(\begin{array}{cc|cc} 8 & 6 & 4 & 5 \\ \hline 10 & 9 & 6 & 7 \\ 18 & 15 & 10 & 12 \end{array}\right)$$

ou

$$\text{(ii)} \qquad \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 \\ \hline 3 & 3 & 2 & 2 \end{array}\right)$$

e, nesse caso,

$$\left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 \\ \hline 3 & 3 & 2 & 2 \end{array}\right) \left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ \hline 3 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 \end{array}\right) = \left(\begin{array}{cc|cc} 8 & 6 & 4 & 5 \\ 10 & 9 & 6 & 7 \\ \hline 18 & 15 & 10 & 12 \end{array}\right)$$

□

EXEMPLO 2. Seja A uma matriz $n \times n$ da forma

$$\begin{pmatrix} A_{11} & O \\ O & A_{22} \end{pmatrix}$$

onde A_{11} é $k \times k$ $(k < n)$. Mostre que A é invertível se e somente se A_{11} e A_{22} são invertíveis.

SOLUÇÃO. Se A_{11} e A_{22} são invertíveis, então

$$\begin{pmatrix} A_{11}^{-1} & O \\ O & A_{22}^{-1} \end{pmatrix} \begin{pmatrix} A_{11} & O \\ O & A_{22} \end{pmatrix} = \begin{pmatrix} I_k & O \\ O & I_{n-k} \end{pmatrix} = I$$

e

$$\begin{pmatrix} A_{11} & O \\ O & A_{22} \end{pmatrix} \begin{pmatrix} A_{11}^{-1} & O \\ O & A_{22}^{-1} \end{pmatrix} = \begin{pmatrix} I_k & O \\ O & I_{n-k} \end{pmatrix} = I$$

de modo que A é invertível e

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} & O \\ O & A_{22}^{-1} \end{pmatrix}$$

Por outro lado, se A é invertível, seja $B = A^{-1}$ e divida B da mesma maneira que A. Como

$$BA = I = AB$$

tem-se que

$$\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \begin{pmatrix} A_{11} & O \\ O & A_{22} \end{pmatrix} = \begin{pmatrix} I_k & O \\ O & I_{n-k} \end{pmatrix} = \begin{pmatrix} A_{11} & O \\ O & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

$$\begin{pmatrix} B_{11}A_{11} & B_{12}A_{22} \\ B_{21}A_{11} & B_{22}A_{22} \end{pmatrix} = \begin{pmatrix} I_k & O \\ O & I_{n-k} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} & A_{11}B_{12} \\ A_{22}B_{21} & A_{22}B_{22} \end{pmatrix}$$

Portanto,

$$B_{11}A_{11} = I_k = A_{11}B_{11}$$
$$B_{22}A_{22} = I_{n-k} = A_{22}B_{22}$$

logo, A_{11} e A_{22} são invertíveis com inversas B_{11} e B_{22}, respectivamente. □

EXERCÍCIOS

1. Seja A uma matriz invertível $n \times n$. Efetue as multiplicações indicadas.

(a) $A^{-1}(A \quad I)$ (b) $\begin{pmatrix} A \\ I \end{pmatrix} A^{-1}$ (c) $(A \quad I)^T(A \quad I)$ (d) $(A \quad I)(A \quad I)^T$ (e) $\begin{pmatrix} A^{-1} \\ I \end{pmatrix}(A \quad I)$

2. Seja $B = A^TA$. Mostre que $b_{ij} = \mathbf{a}_i^T\mathbf{a}_j$.

3. Sejam

$$A = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$$

(a) Calcule $A\mathbf{b}_1$ e $A\mathbf{b}_2$.

(b) Calcule $\mathbf{a}(1,:)B$ e $\mathbf{a}(2,:)B$.

(c) Calcule AB e verifique que suas colunas são os vetores encontrados em (a) e que suas linhas são os vetores encontrados em (b).

4. Sejam

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad E = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad O = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

e

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ \hline 3 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 \end{array}\right)$$

Efetue cada uma das multiplicações em bloco indicadas.

(a) $\begin{pmatrix} O & I \\ I & O \end{pmatrix}\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ (b) $\begin{pmatrix} C & O \\ O & C \end{pmatrix}\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$

(c) $\begin{pmatrix} D & O \\ O & I \end{pmatrix}\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ (d) $\begin{pmatrix} E & O \\ O & E \end{pmatrix}\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$

5. Efetue cada uma das multiplicações em bloco indicadas a seguir.

(a) $\left(\begin{array}{ccc|c} 1 & 1 & 1 & -1 \\ 2 & 1 & 2 & -1 \end{array}\right)\left(\begin{array}{ccc} 4 & -2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 2 \\ \hline 1 & 2 & 3 \end{array}\right)$ (b) $\left(\begin{array}{cc} 4 & -2 \\ 2 & 3 \\ 1 & 1 \\ \hline 1 & 2 \end{array}\right)\left(\begin{array}{ccc|c} 1 & 1 & 1 & -1 \\ 2 & 1 & 2 & -1 \end{array}\right)$

(c) $\left(\begin{array}{cc|cc} \frac{3}{5} & -\frac{4}{5} & 0 & 0 \\ \frac{4}{5} & \frac{3}{5} & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \end{array}\right)\left(\begin{array}{cc|c} \frac{3}{5} & \frac{4}{5} & 0 \\ -\frac{4}{5} & \frac{3}{5} & 0 \\ \hline 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right)$ (d) $\left(\begin{array}{ccc|cc} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right)\left(\begin{array}{cc} 1 & -1 \\ 2 & -2 \\ 3 & -3 \\ \hline 4 & -4 \\ 5 & -5 \end{array}\right)$

6. Sejam

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \text{e} \quad A^T = \begin{pmatrix} A_{11}^T & A_{21}^T \\ A_{12}^T & A_{22}^T \end{pmatrix}$$

É possível efetuar as multiplicações em bloco AA^T e A^TA? Explique.

7. Considere matrizes A $m \times n$, X $n \times r$ e B $m \times r$. Mostre que

$$AX = B$$

se e somente se

$$A\mathbf{x}_j = \mathbf{b}_j, \quad j = 1, \ldots, r$$

8. Seja

$$A = \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix}$$

onde todos os quatro blocos são matrizes $n \times n$.

(a) Se A_{11} e A_{22} são invertíveis, mostre que A também é invertível e que A^{-1} tem que ser da forma

$$\left(\begin{array}{c|c} A_{11}^{-1} & C \\ \hline O & A_{22}^{-1} \end{array}\right)$$

(b) Encontre C.

9. Seja

$$A = \begin{pmatrix} O & I \\ B & O \end{pmatrix}$$

onde todas as quatro submatrizes são $k \times k$. Encontre A^2 e A^4.

10. Seja I a matriz identidade $n \times n$. Encontre a forma em bloco da inversa de cada uma das matrizes $2n \times 2n$ a seguir.

(a) $\begin{pmatrix} O & I \\ I & O \end{pmatrix}$ (b) $\begin{pmatrix} I & O \\ B & I \end{pmatrix}$

11. Sejam A e B matrizes $m \times n$ e $n \times r$, respectivamente. Defina matrizes S e M $(m + n) \times (m + n)$ por

$$S = \begin{pmatrix} I & A \\ O & I \end{pmatrix}, \qquad M = \begin{pmatrix} AB & O \\ B & O \end{pmatrix}$$

Encontre a forma em bloco de S^{-1} e use-a para calcular a forma em bloco do produto $S^{-1}MS$.

12. Seja

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

onde A_{11} é uma matriz invertível $k \times k$. Mostre que A pode ser fatorada em um produto

$$\begin{pmatrix} I & O \\ B & I \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ O & C \end{pmatrix}$$

onde

$$B = A_{21}A_{11}^{-1} \quad \text{e} \quad C = A_{22} - A_{21}A_{11}^{-1}A_{12}$$

13. Sejam A uma matriz $n \times n$ e $\mathbf{x} \in R^n$.

(a) Um escalar c pode também ser considerado uma matriz 1×1 $C = (c)$ e um vetor $\mathbf{b} \in R^n$ pode ser considerado uma matriz $n \times 1$. Mostre que, embora o produto matricial CB não esteja definido, o produto matricial BC é igual a $c\mathbf{b}$, a multiplicação de $\mathbf{b}$ pelo escalar c.

(b) Divida A em colunas e $\mathbf{x}$ em linhas e efetue a multiplicação em bloco de A por $\mathbf{x}$.

(c) Mostre que

$$A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

14. Mostre que, se A é uma matriz $n \times n$ com a propriedade $A\mathbf{x} = \mathbf{0}$ para todo $\mathbf{x} \in R^n$, então $A = O$.

[***Sugestão:*** faça $\mathbf{x} = \mathbf{e}_j$ para $j = 1, \ldots, n$.]

15. Sejam B e C matrizes $n \times n$ com a propriedade $B\mathbf{x} = C\mathbf{x}$ para todo $\mathbf{x} \in R^n$. Mostre que $B = C$.

16. Seja A uma matriz $m \times n$ e suponha que $A = XY^T$, onde X é uma matriz $m \times k$ e Y é $n \times k$. Divida X em colunas e Y^T em linhas e efetue a multiplicação em bloco XY^T. Expresse A como uma soma de k matrizes, cada uma das quais definidas em termos dos vetores colunas de X e Y. Uma soma dessa forma é chamada de uma *expansão em produtos exteriores* de A. Uma aplicação de tais expansões aparece em imagens digitais (ver Seção 6 do Cap. 7).

17. Considere um sistema da forma

$$\begin{pmatrix} A & \mathbf{a} \\ \mathbf{c}^T & \beta \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ x_{n+1} \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ b_{n+1} \end{pmatrix}$$

onde A é uma matriz invertível $n \times n$ e $\mathbf{a}$, $\mathbf{b}$ e $\mathbf{c}$ são vetores em R^n.

(a) Multiplique ambos os lados do sistema por

$$\begin{pmatrix} A^{-1} & \mathbf{0} \\ -\mathbf{c}^T A^{-1} & 1 \end{pmatrix}$$

para obter um sistema triangular equivalente.

(b) Faça $\mathbf{y} = A^{-1}\mathbf{a}$ e $\mathbf{z} = A^{-1}\mathbf{b}$. Mostre que, se $\beta - \mathbf{c}^T\mathbf{y} \neq 0$, então a solução do sistema é dada por

$$x_{n+1} = \frac{b_{n+1} - \mathbf{c}^T\mathbf{z}}{\beta - \mathbf{c}^T\mathbf{y}}$$

$$\mathbf{x} = \mathbf{z} - x_{n+1}\mathbf{y}$$

EXERCÍCIOS COM O MATLAB PARA O CAPÍTULO 1

Os exercícios a seguir devem ser resolvidos através do computador, usando-se o programa MATLAB, descrito no Apêndice deste livro. Os exercícios incluem também perguntas, que devem ser respondidas, sobre os princípios matemáticos subjacentes, ilustrados nos cálculos. Use o comando `diary` do MATLAB para guardar sua sessão em um arquivo. Após editar e imprimir o arquivo, as respostas das perguntas podem ser escritas diretamente no papel impresso.

O MATLAB tem um programa de ajuda que explica todas as suas operações e comandos. Os comandos do MATLAB estão impressos nestes exercícios na fonte que imita um máquina de escrever. Por exemplo, para obter informação sobre o comando `rand`, basta digitar `help rand`, e para informação sobre o operador "\", digite `help \`. As operações do MATLAB utilizadas nos exercícios para o Cap. 1 são +, −, ∗, ′,\ e os comandos utilizados são `inv`, `round`, `rand`, `flops`, `rref`, `format`, `sum`, `eye`, `triu`, `ones`, `zeros`, `magic`.

1. Use o MATLAB para gerar matrizes aleatórias 4×4 A e B. Em cada item a seguir, calcule $A1$, $A2$, $A3$, $A4$ como indicado e determine quais das matrizes são iguais. Você pode usar o MATLAB para testar se duas matrizes são iguais ou não, calculando sua diferença.

(a) $A1 = A * B$, $A2 = B * A$, $A3 = (A' * B')'$, $A4 = (B' * A')'$

(b) $A1 = A' * B'$, $A2 = (A * B)'$, $A3 = B' * A'$, $A4 = (B * A)'$

(c) $A1 = \texttt{inv}(A * B)$, $A2 = \texttt{inv}(A) * \texttt{inv}(B)$
$A3 = \texttt{inv}(B * A)$, $A4 = \texttt{inv}(B) * \texttt{inv}(A)$

(d) $A1 = \texttt{inv}((A * B)')$, $A2 = \texttt{inv}(A' * B')$
$A3 = \texttt{inv}(A') * \texttt{inv}(B')$, $A4 = (\texttt{inv}(A) * \texttt{inv}(B))'$

2. Gere uma matriz 8×8 e um vetor em R^8, ambos com coeficientes inteiros, digitando

$$A = \texttt{round}(10 * \texttt{rand}(8)) \quad \text{and} \quad \mathbf{b} = \texttt{round}(10 * \texttt{rand}(8, 1))$$

(a) Podemos estimar a quantidade de operações aritméticas envolvidas na resolução do sistema $A\mathbf{x} = \mathbf{b}$ usando a função `flops` do MATLAB. O valor da variável `flops` é a estimativa do MATLAB do número total de operações aritméticas em ponto flutuante efetuadas até então na sessão atual do MATLAB. O valor de `flops` pode ser anulado digitando-se `flops(0)`. Recoloque `flops` igual a 0 e depois encontre a solução $\mathbf{x}$ do sistema $A\mathbf{x} = \mathbf{b}$ usando o operador "\". Depois, digite `flops` para determinar quantas operações aritméticas foram efetuadas, aproximadamente, para resolver o sistema.

(b) A seguir, vamos resolver o sistema usando o método de Gauss-Jordan. Primeiro faça `flops` igual a 0. Calcule a forma escada reduzida por linhas da matriz aumentada $(A \quad \mathbf{b})$. Isso pode ser feito com o comando

$$U = \texttt{rref}([A \quad \mathbf{b}])$$

Em aritmética exata, a última coluna da forma escada reduzida por linhas deve ser a solução do sistema. Por quê? Explique. Faça $\mathbf{y}$ igual à última coluna de U e digite `flops` para estimar o número de operações em ponto flutuante efetuadas para calcular $\mathbf{y}$. Qual dos métodos foi mais eficiente? O que usou o operador "\" ou o de Gauss-Jordan? (Observação: o operador "\" consiste, essencialmente, em redução a uma forma triangular seguida de substituição.)

(c) As soluções $\mathbf{x}$ e $\mathbf{y}$ obtidas pelos dois métodos parecem ser iguais, mas se você examinar mais dígitos dos vetores usando o comando `format long` vai ver que não são idênticas. Quantos dígitos são iguais nos dois vetores? Uma maneira mais fácil de comparar os dois vetores é usar `format short` e examinar a diferença $\mathbf{x} - \mathbf{y}$.

(d) Qual das duas soluções encontradas, $\mathbf{x}$ e $\mathbf{y}$, é mais precisa? Para responder, compare cada um dos produtos $A\mathbf{x}$ e $A\mathbf{y}$ com o lado direito da equação, $\mathbf{b}$. A maneira mais simples de fazer isso é olhar as diferenças $\mathbf{r} = \mathbf{b} - A\mathbf{x}$ e $\mathbf{s} = \mathbf{b} - A\mathbf{y}$. Os vetores $\mathbf{r}$ e $\mathbf{s}$ são chamados de vetores residuais para as soluções calculadas $\mathbf{x}$ e $\mathbf{y}$, respectivamente. Qual das duas soluções tem o menor vetor residual?

3. Faça $A = \texttt{round}(10) * \texttt{rand}(6))$. Por construção, a matriz A tem elementos inteiros. Vamos mudar a sexta coluna de A de modo a torná-la singular. Faça

$$B = A', \qquad A(:, 6) = -\,\texttt{sum}(B(1:5, :))'$$

(a) Faça $\mathbf{x} = \texttt{ones}(6, 1)$ e use o MATLAB para calcular $A * \mathbf{x}$. Por que sabemos que A tem que ser singular? Explique. Verifique que A é singular calculando sua forma escada reduzida por linhas.

(b) Faça

$$B = \mathbf{x} * [1:6]$$

O produto AB deveria ser igual à matriz nula. Por quê? Explique. Verifique que isso acontece calculando AB com a operação $*$ do MATLAB.

(c) Faça

$$C = \texttt{round}(10 * \texttt{rand}(6)) \qquad \text{e} \qquad D = B + C$$

Embora $C \neq D$, os produtos AC e AD deveriam ser iguais. Por quê? Explique. Calcule $A * C$ e $A * D$ e verifique que são, de fato, iguais.

4. Construa uma matriz da maneira descrita a seguir. Faça

$$B = \texttt{eye}(10) - \texttt{triu}(\texttt{ones}(10), 1)$$

Por que sabemos que B tem que ser singular? Faça

$$C = \texttt{inv}(B) \quad \text{e} \quad \mathbf{x} = C(:, 10)$$

Agora mude B ligeiramente, fazendo $B(10, 1) = -1/256$. Use o MATLAB para calcular o produto $B\mathbf{x}$. O que você pode concluir sobre a nova matriz B do resultado desse cálculo? Ela ainda é invertível? Explique. Use o MATLAB para calcular sua forma escada reduzida por linhas.

5. Gere uma matriz A fazendo

$$A = \texttt{round}(10 * \texttt{rand}(6))$$

e gere um vetor $\mathbf{b}$ digitando

$$\mathbf{b} = \texttt{round}(20 * (\texttt{rand}(6, 1) - 0.5))$$

(a) Como A foi gerada aleatoriamente, esperaríamos que ela fosse invertível. O sistema $A\mathbf{x} = \mathbf{b}$ deveria, então, ter uma única solução. Encontre a solução usando o operador "\". Use o MATLAB para encontrar a forma escada reduzida por linhas U de $[A\ \mathbf{b}]$. Compare a última coluna de U com a solução $\mathbf{x}$. Em aritmética exata, deveriam ser iguais. Por quê? Explique. Para comparar as duas, calcule a diferença $U(:, 7) - \mathbf{x}$ ou examine ambas usando `format long`.

(b) Vamos, agora, mudar A de modo a torná-la singular. Faça

$$A(:, 3) = A(:, 1:2) * [4\ 3]'$$

Use o MATLAB para calcular `rref`$([A\ \mathbf{b}])$. Quantas soluções tem o sistema $A\mathbf{x} = \mathbf{b}$? Explique.

(c) Faça

$$\mathbf{y} = \texttt{round}(20 * (\texttt{rand}(6, 1) - 0.5)) \quad \text{e} \quad \mathbf{c} = A * \mathbf{y}$$

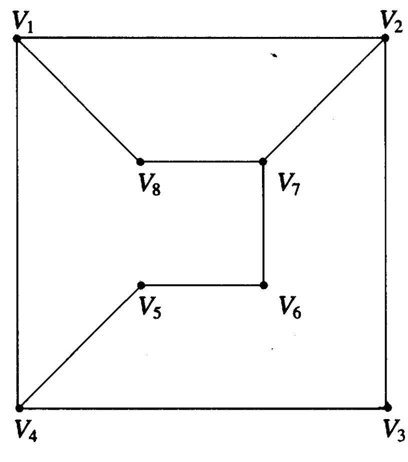
Por que sabemos que o sistema $A\mathbf{x} = \mathbf{c}$ tem que ser compatível? Explique. Calcule a forma escada reduzida por linhas U de $[A\ \mathbf{c}]$. Quantas soluções tem o sistema $A\mathbf{x} = \mathbf{c}$? Explique.

(d) A variável livre determinada pela forma escada deve ser x_3. Examinando o sistema correspondente à matriz U, você deve ser capaz de determinar a solução para $x_3 = 0$. Coloque essa solução no MATLAB como um vetor coluna $\mathbf{w}$. Para verificar que $A\mathbf{w} = \mathbf{c}$, calcule o vetor residual $\mathbf{c} - A\mathbf{w}$.

(e) Faça $U(:, 7) =$ `zeros`$(6, 1)$. A matriz U deve corresponder agora à forma escada reduzida por linhas de $(A|\mathbf{0})$. Use U para determinar a solução do sistema homogêneo quando a variável livre $x_3 = 1$ (faça isso à mão) e coloque seu resultado como um vetor $\mathbf{z}$. Verifique sua resposta calculando $A * \mathbf{z}$.

(f) Faça $\mathbf{v} = \mathbf{w} + 3 * \mathbf{z}$. O vetor $\mathbf{v}$ deve ser uma solução do sistema $A\mathbf{x} = \mathbf{c}$. Por quê? Explique. Verifique que $\mathbf{v}$ é uma solução usando o MATLAB para calcular o vetor residual $\mathbf{c} - A\mathbf{v}$. Qual o valor da variável livre x_3 para essa solução? Como podemos determinar todas as soluções possíveis em termos dos vetores $\mathbf{w}$ e $\mathbf{z}$? Explique.

6. Considere o grafo



(a) Determine a matriz de adjacência A para o grafo e coloque-a no MATLAB.

(b) Calcule A^2 e determine o número de caminhos de comprimento 2 de (i) V_1 a V_7, (ii) V_4 a V_8, (iii) V_5 a V_6, (iv) V_8 a V_3.

(c) Calcule A^4, A^6, A^8 e responda as perguntas em (b) para caminhos de comprimento 4, 6 e 8. Faça uma conjectura sobre quando não haverá caminhos de comprimento par do vértice V_i para o vértice V_j.

(d) Calcule A^3, A^5, A^7 e responda as perguntas em (b) para caminhos de comprimento 3, 5 e 7. Sua conjectura em (c) é válida para caminhos de comprimento ímpar? Explique. Faça uma conjectura sobre a existência ou não de caminhos de comprimento k de V_i a V_j, dependendo se $i + j + k$ é par ou ímpar.

(e) Se adicionarmos as arestas $\{V_3, V_6\}$ e $\{V_5, V_8\}$ ao grafo, a matriz de adjacência B do novo grafo pode ser gerada fazendo-se $B = A$ e depois digitando-se

$$B(3,6) = 1, \qquad B(6,3) = 1, \qquad B(5,8) = 1, \qquad B(8,5) = 1$$

Calcule B^k para $k = 2, 3, 4, 5$. Sua conjectura em (d) ainda é válida para o novo grafo?

(f) Adicione a aresta $\{V_6, V_8\}$ à figura e construa a matriz de adjacência C do grafo resultante. Calcule algumas potências de C para determinar se sua conjectura em (d) ainda é válida para esse novo grafo.

7. Na Aplicação 3 da Seção 3, o número de mulheres casadas e solteiras depois de um e dois anos foi determinado calculando-se os produtos AX e A^2X para as matrizes dadas A e X. Use `format long` e coloque essas matrizes no MATLAB. Calcule A^k e A^kX para $k = 5, 10, 15, 20$. O que acontece com A^k quando k aumenta? Qual é a distribuição final de mulheres casadas e solteiras na cidade?

8. Faça $A = \texttt{magic}(8)$ e depois calcule sua forma escada reduzida por linhas. Os primeiros elementos não-nulos de cada linha, iguais a 1, devem corresponder às três primeiras variáveis x_1, x_2, x_3 e as cinco variáveis restantes são todas livres.

(a) Faça $\mathbf{c} = [1:8]'$ e determine se o sistema $A\mathbf{x} = \mathbf{c}$ é compatível ou não, calculando a forma escada reduzida por linhas de $[A \;\; \mathbf{c}]$. O sistema é compatível? Explique.

(b) Faça

$$\mathbf{b} = [8 \;\; -8 \;\; -8 \;\; 8 \;\; 8 \;\; -8 \;\; -8 \;\; 8]';$$

e considere o sistema $A\mathbf{x} = \mathbf{b}$. Esse sistema deve ser compatível. Verifique que esse é o caso calculando $U = \texttt{rref}([A \;\; \mathbf{b}])$. Deveríamos ser capazes de encontrar uma solução para qualquer escolha das cinco variáveis livres. De fato, faça $\mathbf{x2} = \texttt{round}(10 * \texttt{rand}(5, 1))$. Se $\mathbf{x2}$ representa as cinco últimas coordenadas de uma solução do sistema, então deveríamos ser capazes de determinar $\mathbf{x1} = (x_1, x_2, x_3)^T$ em termos de $\mathbf{x2}$. Para fazer isso, defina $U = \texttt{rref}([A \;\; \mathbf{b}])$. As linhas não-nulas de U correspondem a um sistema linear em bloco,

$$\begin{pmatrix} I & V \end{pmatrix} \begin{pmatrix} \mathbf{x1} \\ \mathbf{x2} \end{pmatrix} = \mathbf{c} \tag{1}$$

Para resolver a equação (1), faça

$$V = U(1:3,\ 4:8), \qquad \mathbf{c} = U(1:3,\ 9)$$

e use o MATLAB para calcular $\mathbf{x1}$ em termos de $\mathbf{x2}$, $\mathbf{c}$ e V. Defina $\mathbf{x} = [\mathbf{x1}; \mathbf{x2}]$ e verifique que $\mathbf{x}$ é uma solução do sistema.

9. Defina

$$B = [-1, -1;\ 1, 1] \quad \text{e} \quad A = [\texttt{zeros}(2), \texttt{eye}(2);\ \texttt{eye}(2), B]$$

e verifique que $B^2 = 0$.

(a) Use o MATLAB para calcular A^2, A^4, A^6 e A^8. Faça uma conjectura sobre a forma em bloco da matriz A^{2k} em termos das submatrizes I, O e B. Use indução matemática para provar sua conjectura para qualquer inteiro positivo k.

(b) Use o MATLAB para calcular A^3, A^5, A^7 e A^9. Faça uma conjectura sobre a forma em bloco da matriz A^{2k-1} em termos das submatrizes I, O e B. Prove sua conjectura.

10. (a) O comando

$$A = \texttt{round}(10 * \texttt{rand}(6)), \qquad B = A' * A$$

do MATLAB gera uma matriz simétrica com elementos inteiros. Por quê? Explique. Calcule B dessa maneira e verifique essa afirmação. A seguir, divida B em quatro submatrizes 3×3. Para determinar as submatrizes usando o MATLAB, faça

$$B11 = B(1:3,\ 1:3), \qquad B12 = B(1:3,\ 4:6)$$

e defina $B21$ e $B22$ de maneira análoga, usando as linhas de 4 a 6 de B.

(b) Defina $C = \texttt{inv}(B11)$. Deveríamos ter $C^T = C$ e $B21^T = B12$. Por quê? Explique. Use o operador do MATLAB para calcular as transpostas e verificar essas afirmações. A seguir, defina

$$E = B21 * C \qquad \text{e} \qquad F = B22 - B21 * C * B21'$$

e use as funções `eye` e `zeros` para construir

$$L = \begin{pmatrix} I & O \\ E & I \end{pmatrix}, \quad D = \begin{pmatrix} B11 & O \\ O & F \end{pmatrix}$$

Calcule $H = L * D * L'$ e compare o resultado com B, calculando $H - B$. Prove que, se todos os cálculos fossem feitos em aritmética exata, então LDL^T seria exatamente igual a B.

CAPÍTULO 2

DETERMINANTES

É possível associar a cada matriz quadrada um número real chamado determinante da matriz. O valor desse número vai dizer se a matriz é invertível ou não.

Na Seção 1, definiremos o determinante de uma matriz. Na Seção 2, estudaremos as propriedades do determinante e desenvolveremos um método de redução para calcular determinantes. Esse método é, em geral, o mais simples para calcular determinantes de matrizes $n \times n$ quando $n > 3$. Na Seção 3, vamos ver como usar determinantes para resolver sistemas lineares $n \times n$ e para calcular a inversa de uma matriz. Também apresentaremos na Seção 3 uma aplicação envolvendo criptografia. Outras aplicações de determinantes serão feitas nos Caps. 3 e 6.

1 O DETERMINANTE DE UMA MATRIZ

É possível associar a cada matriz A $n \times n$ um escalar, $\det(A)$, cujo valor vai nos dizer se a matriz é ou não invertível. Antes de dar a definição geral, vamos considerar alguns casos particulares.

Caso 1. Matrizes 1×1

Se $A = (a)$ é uma matriz 1×1, então A tem uma inversa multiplicativa se e somente se $a \neq 0$. Logo, se definirmos

$$\det(A) = a$$

A será invertível se e somente se $\det(A) \neq 0$.

Caso 2. Matrizes 2×2

Seja

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Pelo Teorema 1.4.3, A é invertível se e somente se é equivalente por linhas a I. Então, se $a_{11} \neq 0$, podemos testar se A é ou não equivalente por linhas a I efetuando as seguintes operações:

1. Multiplique a segunda linha de A por a_{11}

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{11}a_{21} & a_{11}a_{22} \end{pmatrix}$$

2. Subtraia a_{21} vezes a primeira linha da nova segunda linha

$$\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{11}a_{22} - a_{21}a_{12} \end{pmatrix}$$

Como $a_{11} \neq 0$, a matriz resultante é equivalente por linhas a I se e somente se

(1) $$a_{11}a_{22} - a_{21}a_{12} \neq 0$$

Se $a_{11} = 0$, podemos trocar as duas linhas de A. A matriz resultante

$$\begin{pmatrix} a_{21} & a_{22} \\ 0 & a_{12} \end{pmatrix}$$

é equivalente por linhas a I se e somente se $a_{21}a_{12} \neq 0$. Essa condição é equivalente a (1) quando $a_{11} = 0$. Logo, se A é qualquer matriz 2×2 e definirmos

$$\det(A) = a_{11}a_{22} - a_{12}a_{21}$$

então A é invertível se e somente se $\det(A) \neq 0$.

Notação. É costume denotar o determinante de uma matriz particular colocando-se o arranjo de números entre retas verticais. Por exemplo, se

$$A = \begin{pmatrix} 3 & 4 \\ 2 & 1 \end{pmatrix}$$

então

$$\begin{vmatrix} 3 & 4 \\ 2 & 1 \end{vmatrix}$$

representa o determinante de A.

Caso 3. Matrizes 3×3

Podemos testar se uma matriz 3×3 é ou não invertível efetuando operações elementares para verificar se ela é ou não equivalente por linhas à matriz identidade I. Para anular os elementos da primeira coluna de uma matriz arbitrária A 3×3, vamos supor, primeiro, que $a_{11} \neq 0$. Podemos, então, anular os elementos desejados subtraindo a_{21}/a_{11} vezes a primeira linha da segunda e subtraindo a_{31}/a_{11} vezes a primeira linha da terceira.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \rightarrow \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & \dfrac{a_{11}a_{22} - a_{21}a_{12}}{a_{11}} & \dfrac{a_{11}a_{23} - a_{21}a_{13}}{a_{11}} \\ 0 & \dfrac{a_{11}a_{32} - a_{31}a_{12}}{a_{11}} & \dfrac{a_{11}a_{33} - a_{31}a_{13}}{a_{11}} \end{pmatrix}$$

A matriz à direita é equivalente por linhas a I se e somente se

$$a_{11} \begin{vmatrix} \dfrac{a_{11}a_{22} - a_{21}a_{12}}{a_{11}} & \dfrac{a_{11}a_{23} - a_{21}a_{13}}{a_{11}} \\ \dfrac{a_{11}a_{32} - a_{31}a_{12}}{a_{11}} & \dfrac{a_{11}a_{33} - a_{31}a_{13}}{a_{11}} \end{vmatrix} \neq 0$$

Embora a álgebra seja um pouco trabalhosa, essa condição pode ser simplificada para

(2) $$\begin{aligned} a_{11}a_{22}a_{33} &- a_{11}a_{32}a_{23} - a_{12}a_{21}a_{33} + a_{12}a_{31}a_{23} \\ &+ a_{13}a_{21}a_{32} - a_{13}a_{31}a_{22} \neq 0 \end{aligned}$$

Logo, se definirmos

(3) $$\begin{aligned} \det(A) = {} & a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{12}a_{21}a_{33} \\ & + a_{12}a_{31}a_{23} + a_{13}a_{21}a_{32} - a_{13}a_{31}a_{22} \end{aligned}$$

então, para o caso $a_{11} \neq 0$, a matriz vai ser invertível se e somente se $\det(A) \neq 0$.

E se $a_{11} = 0$? Vamos considerar três possibilidades:

$$\begin{aligned}
&\text{(i)} \quad a_{11} = 0, a_{21} \neq 0 \\
&\text{(ii)} \quad a_{11} = a_{21} = 0, a_{31} \neq 0 \\
&\text{(iii)} \quad a_{11} = a_{21} = a_{31} = 0
\end{aligned}$$

No caso (i), não é difícil mostrar que A é equivalente por linhas a I se e somente se

$$-a_{12}a_{21}a_{33} + a_{12}a_{31}a_{23} + a_{13}a_{21}a_{32} - a_{13}a_{31}a_{22} \neq 0$$

Mas essa condição é a mesma que (2) com $a_{11} = 0$. Deixamos a cargo do leitor os detalhes do caso (i) (ver Exercício 7).

No caso (ii), temos que

$$A = \begin{pmatrix} 0 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

é equivalente por linhas a I se e somente se

$$a_{31}(a_{12}a_{23} - a_{22}a_{13}) \neq 0$$

Novamente, esse é um caso particular de (2) com $a_{11} = a_{21} = 0$.

É claro que no caso (iii) a matriz A não pode ser equivalente por linhas a I, logo é singular. Nesse caso, fazendo a_{11}, a_{21} e a_{31} iguais a 0 na fórmula (3), obtemos $\det(A) = 0$.

Em geral, portanto, a fórmula (2) nos dá uma condição necessária e suficiente para uma matriz 3×3 ser invertível (independentemente do valor de a_{11}).

Gostaríamos de definir o determinante de uma matriz $n \times n$. Para ver como fazer isso, observe que o determinante de uma matriz 2×2

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

pode ser definido em termos das matrizes 1×1

$$M_{11} = (a_{22}) \qquad \text{e} \qquad M_{12} = (a_{21})$$

A matriz M_{11} é formada retirando-se a primeira linha e a primeira coluna de A e M_{12} é formada retirando-se a primeira linha e a segunda coluna de A.

O determinante de A pode ser escrito na forma

(4) $$\det(A) = a_{11}a_{22} - a_{12}a_{21} = a_{11}\det(M_{11}) - a_{12}\det(M_{12})$$

Para uma matriz A 3×3, podemos colocar a equação (3) na forma

$$\det(A) = a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22})$$

Para $j = 1, 2, 3$, vamos denotar por M_{1j} a matriz 2×2 formada retirando-se a primeira linha e a j-ésima coluna de A. O determinante de A pode ser, então, colocado na forma

(5) $$\det(A) = a_{11}\det(M_{11}) - a_{12}\det(M_{12}) + a_{13}\det(M_{13})$$

onde

$$M_{11} = \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}, \quad M_{12} = \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix}, \quad M_{13} = \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

Para ver como generalizar (4) e (5) para o caso $n > 3$, vamos dar a seguinte definição.

Definição. Seja $A = (a_{ij})$ uma matriz $n \times n$. Seja M_{ij} a matriz $(n-1) \times (n-1)$ obtida retirando-se a linha e coluna de A que contém a_{ij}. O determinante de M_{ij} é chamado de **determinante menor** de a_{ij}. Definimos o **cofator** A_{ij} de a_{ij} por

$$A_{ij} = (-1)^{i+j}\det(M_{ij})$$

Em vista dessa definição, podemos colocar a equação (4) para uma matriz 2 × 2 na forma

(6) $$\det(A) = a_{11}A_{11} + a_{12}A_{12} \qquad (n = 2)$$

A Equação (6) é chamada de *expansão em cofatores* do det(A) em relação à primeira linha de A. Observe que poderíamos também ter escrito

(7) $$\det(A) = a_{21}(-a_{12}) + a_{22}a_{11} = a_{21}A_{21} + a_{22}A_{22}$$

A Equação (7) expressa det(A) em função dos elementos da segunda linha de A e de seus cofatores. Na verdade, não existe nenhuma razão especial para se expandir em relação a uma linha; o determinante poderia ser encontrado de maneira análoga expandindo-se em relação a uma coluna:

$$\begin{aligned}\det(A) &= a_{11}a_{22} + a_{21}(-a_{12}) \\ &= a_{11}A_{11} + a_{21}A_{21} \qquad \text{(primeira coluna)} \\ \det(A) &= a_{12}(-a_{21}) + a_{22}a_{11} \\ &= a_{12}A_{12} + a_{22}A_{22} \qquad \text{(segunda coluna)}\end{aligned}$$

Para uma matriz A 3 × 3, temos

(8) $$\det(A) = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13}$$

Portanto, o determinante de uma matriz 3 × 3 pode ser definido em termos dos elementos da primeira linha da matriz e seus fatores correspondentes.

EXEMPLO 1. Se

$$A = \begin{pmatrix} 2 & 5 & 4 \\ 3 & 1 & 2 \\ 5 & 4 & 6 \end{pmatrix}$$

então

$$\begin{aligned}\det(A) &= a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} \\ &= (-1)^2 a_{11}\det(M_{11}) + (-1)^3 a_{12}\det(M_{12}) \\ &\quad + (-1)^4 a_{13}\det(M_{13}) \\ &= 2\begin{vmatrix} 1 & 2 \\ 4 & 6 \end{vmatrix} - 5\begin{vmatrix} 3 & 2 \\ 5 & 6 \end{vmatrix} + 4\begin{vmatrix} 3 & 1 \\ 5 & 4 \end{vmatrix} \\ &= 2(6-8) - 5(18-10) + 4(12-5) \\ &= -16\end{aligned}$$

□

Como no caso de matrizes 2 × 2, o determinante de uma matriz 3 × 3 pode ser representado por uma expansão em cofatores em relação a qualquer linha ou coluna. Por exemplo, a equação (3) pode ser colocada na forma

$$\begin{aligned}\det(A) &= a_{12}a_{31}a_{23} - a_{13}a_{31}a_{22} - a_{11}a_{32}a_{23} + a_{13}a_{21}a_{32} + a_{11}a_{22}a_{33} \\ &\quad - a_{12}a_{21}a_{33} \\ &= a_{31}(a_{12}a_{23} - a_{31}a_{22}) - a_{32}(a_{11}a_{23} - a_{13}a_{21}) \\ &\quad + a_{33}(a_{11}a_{22} - a_{12}a_{21}) \\ &= a_{31}A_{31} + a_{32}A_{32} + a_{33}A_{33}\end{aligned}$$

Essa é a expansão em cofatores em relação à terceira linha de A.

EXEMPLO 2. Seja A a matriz do Exemplo 1. A expansão em cofatores do $\det(A)$ em relação à segunda coluna é dada por

$$\begin{aligned}\det(A) &= -5\begin{vmatrix}3 & 2\\ 5 & 6\end{vmatrix} + 1\begin{vmatrix}2 & 4\\ 5 & 6\end{vmatrix} - 4\begin{vmatrix}2 & 4\\ 3 & 2\end{vmatrix}\\ &= -5(18-10) + 1(12-20) - 4(4-12)\\ &= -16\end{aligned}$$

□

O determinante de uma matriz 4×4 pode ser definido como uma expansão em cofatores ao longo de qualquer linha ou coluna. Para calcular o determinante de uma matriz 4×4, teríamos que calcular quatro determinantes 3×3.

Definição. O **determinante** de uma matriz A $n \times n$, denotado por $\det(A)$, é um escalar associado à matriz A, definido indutivamente como se segue:

$$\det(A) = \begin{cases} a_{11} & \text{se } n = 1\\ a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n} & \text{se } n > 1\end{cases}$$

onde

$$A_{1j} = (-1)^{1+j}\det(M_{1j}) \qquad j = 1, \ldots, n$$

são os cofatores associados aos elementos na primeira linha de A.

Como vimos, não é necessário nos limitarmos a expandir em cofatores em relação à primeira linha. Enunciamos o teorema a seguir sem demonstração.

Teorema 2.1.1. *Se A é uma matriz $n \times n$ com $n \geq 2$, então $\det(A)$ pode ser expresso como uma expansão em cofatores em relação a qualquer linha ou coluna de A.*

$$\begin{aligned}\det(A) &= a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}\\ &= a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj}\end{aligned}$$

para $i = 1, \ldots, n$ e $j = 1, \ldots, n$.

A expansão em cofatores para um determinante 4×4 vai envolver quatro determinantes 3×3. Podemos simplificar nosso trabalho, muitas vezes, expandindo em relação à linha ou coluna que contém o maior número de zeros. Por exemplo, para calcular

$$\begin{vmatrix}0 & 2 & 3 & 0\\ 0 & 4 & 5 & 0\\ 0 & 1 & 0 & 3\\ 2 & 0 & 1 & 3\end{vmatrix}$$

expandiríamos em relação à primeira coluna. Os três primeiros termos são nulos e obtemos

$$-2\begin{vmatrix}2 & 3 & 0\\ 4 & 5 & 0\\ 1 & 0 & 3\end{vmatrix} = -2 \cdot 3 \cdot \begin{vmatrix}2 & 3\\ 4 & 5\end{vmatrix} = 12$$

A expansão em cofatores pode ser usada para se obterem alguns resultados importantes sobre determinantes. Esses resultados são dados nos teoremas a seguir.

Teorema 2.1.2. *Se A é uma matriz $n \times n$, então $\det(A^T) = \det(A)$.*

Demonstração. A demonstração é por indução em n. É claro que o resultado é válido para $n = 1$, já que uma matriz 1×1 é necessariamente simétrica. Suponha que o resultado é válido para todas as matrizes $k \times k$ e que A é uma matriz $(k + 1) \times (k + 1)$. Expandindo $\det(A)$ em relação à primeira linha, obtemos

$$\det(A) = a_{11}\det(M_{11}) - a_{12}\det(M_{12}) + - \cdots \pm a_{1,k+1}\det(M_{1,k+1})$$

Como as matrizes M_{ij} são todas $k \times k$, pela hipótese de indução temos que

$$\det(A) = a_{11}\det(M_{11}^T) - a_{12}\det(M_{12}^T) + - \cdots \pm a_{1,k+1}\det(M_{1,k+1}^T) \tag{9}$$

A expressão do lado direito do sinal de igualdade em (9) é simplesmente a expansão em determinantes menores de $\det(A^T)$ em relação à primeira coluna de A^T. Portanto,

$$\det(A^T) = \det(A)$$

□

Teorema 2.1.3. *Se A é uma matriz triangular n × n, então o determinante de A é igual ao produto dos elementos na diagonal de A.*

Demonstração. Em vista do Teorema 2.1.2, basta provar o teorema para matrizes triangulares inferiores. O resultado segue facilmente por indução em n, usando a expansão em cofatores. Os detalhes são deixados a cargo do leitor (ver Exercício 8). □

Teorema 2.1.4. *Seja A uma matriz n × n.*

(i) *Se A tem uma linha ou coluna contendo apenas zeros, então* $\det(A) = 0$.
(ii) *Se A tem duas linhas ou duas colunas idênticas, então* $\det(A) = 0$.

Esses dois resultados podem ser provados facilmente usando-se expansão em cofatores. As demonstrações ficam a cargo do leitor (ver Exercícios 9 e 10).

Na próxima seção, vamos examinar o efeito das operações elementares sobre o determinante. Isso vai nos permitir usar o Teorema 2.1.3 para obter um método mais eficiente de calcular o valor de um determinante.

EXERCÍCIOS

1. Seja

$$A = \begin{pmatrix} 3 & 2 & 4 \\ 1 & -2 & 3 \\ 2 & 3 & 2 \end{pmatrix}$$

(a) Encontre os valores de $\det(M_{21})$, $\det(M_{22})$ e $\det(M_{23})$.
(b) Encontre os valores de A_{21}, A_{22} e A_{23}.
(c) Use as respostas em (a) e (b) para calcular $\det(A)$.

2. Use determinantes para verificar, para cada uma das matrizes a seguir, se a matriz é ou não invertível.

$$\text{(a)} \begin{pmatrix} 3 & 5 \\ 2 & 4 \end{pmatrix} \quad \text{(b)} \begin{pmatrix} 3 & 6 \\ 2 & 4 \end{pmatrix} \quad \text{(c)} \begin{pmatrix} 3 & -6 \\ 2 & 4 \end{pmatrix}$$

3. Calcule cada um dos determinantes a seguir.

$$\text{(a)} \begin{vmatrix} 3 & 5 \\ -2 & -3 \end{vmatrix} \quad \text{(b)} \begin{vmatrix} 5 & -2 \\ -8 & 4 \end{vmatrix} \quad \text{(c)} \begin{vmatrix} 3 & 1 & 2 \\ 2 & 4 & 5 \\ 2 & 4 & 5 \end{vmatrix} \quad \text{(d)} \begin{vmatrix} 4 & 3 & 0 \\ 3 & 1 & 2 \\ 5 & -1 & -4 \end{vmatrix}$$

$$\text{(e)} \begin{vmatrix} 1 & 3 & 2 \\ 4 & 1 & -2 \\ 2 & 1 & 3 \end{vmatrix} \quad \text{(f)} \begin{vmatrix} 2 & -1 & 2 \\ 1 & 3 & 2 \\ 5 & 1 & 6 \end{vmatrix} \quad \text{(g)} \begin{vmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 6 & 2 & 0 \\ 1 & 1 & -2 & 3 \end{vmatrix} \quad \text{(h)} \begin{vmatrix} 2 & 1 & 2 & 1 \\ 3 & 0 & 1 & 1 \\ -1 & 2 & -2 & 1 \\ -3 & 2 & 3 & 1 \end{vmatrix}$$

4. Diga o valor de cada determinante a seguir diretamente, analisando cada matriz.

(a) $\begin{vmatrix} 3 & 5 \\ 2 & 4 \end{vmatrix}$ (b) $\begin{vmatrix} 2 & 0 & 0 \\ 4 & 1 & 0 \\ 7 & 3 & -2 \end{vmatrix}$

(c) $\begin{vmatrix} 3 & 0 & 0 \\ 2 & 1 & 1 \\ 1 & 2 & 2 \end{vmatrix}$ (d) $\begin{vmatrix} 4 & 0 & 2 & 1 \\ 5 & 0 & 4 & 2 \\ 2 & 0 & 3 & 4 \\ 1 & 0 & 2 & 3 \end{vmatrix}$

5. Calcule o determinante a seguir, escrevendo sua resposta como um polinômio em x.

$$\begin{vmatrix} a-x & b & c \\ 1 & -x & 0 \\ 0 & 1 & -x \end{vmatrix}$$

6. Encontre todos os valores de λ para os quais o determinante a seguir é igual a 0.

$$\begin{vmatrix} 2-\lambda & 4 \\ 3 & 3-\lambda \end{vmatrix}$$

7. Seja A uma matriz 3×3 com $a_{11} = 0$ e $a_{21} \neq 0$. Mostre que A é equivalente por linhas a I se e somente se

$$-a_{12}a_{21}a_{33} + a_{12}a_{31}a_{23} + a_{13}a_{21}a_{32} - a_{13}a_{31}a_{22} \neq 0$$

8. Escreva os detalhes da demonstração do Teorema 2.1.3.

9. Prove que se uma linha ou coluna de uma matriz A $n \times n$ tem todos os elementos iguais a zero, então $\det(A) = 0$.

10. Use indução matemática para provar que se A é uma matriz $(n+1) \times (n+1)$ com duas linhas idênticas, então $\det(A) = 0$.

11. Sejam A e B matrizes 2×2.
(a) $\det(A + B) = \det(A) + \det(B)$?
(b) $\det(AB) = \det(A)\det(B)$?
(c) $\det(AB) = \det(BA)$?
Justifique suas respostas.

12. Sejam A e B duas matrizes 2×2 e sejam

$$C = \begin{pmatrix} a_{11} & a_{12} \\ b_{21} & b_{22} \end{pmatrix}, \quad D = \begin{pmatrix} b_{11} & b_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad E = \begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix}$$

(a) Mostre que $\det(A + B) = \det(A) + \det(B) + \det(C) + \det(D)$.
(b) Mostre que, se $B = EA$, então $\det(A + B) = \det(A) + \det(B)$.

13. Seja A uma matriz simétrica tridiagonal (isto é, A é simétrica e $a_{ij} = 0$ sempre que $|i - j| > 1$). Seja B a matriz obtida retirando-se as duas primeiras linhas e colunas de A. Mostre que

$$\det(A) = a_{11}\det(M_{11}) - a_{12}^2\det(B)$$

2 PROPRIEDADES DE DETERMINANTES

Vamos considerar, nesta seção, os efeitos das operações elementares sobre o determinante de uma matriz. Uma vez estabelecidos esses efeitos, vamos provar que uma matriz é invertível se e somente se seu determinante é nulo e vamos desenvolver um método para calcular determinantes através de operações elementares. Além disso, vamos obter um resultado importante sobre o determinante de um produto de matrizes. Vamos começar com o seguinte lema:

Lema 2.2.1. *Seja A uma matriz $n \times n$. Se A_{jk} denota o cofator de a_{jk} para $k = 1, \ldots, n$, então*

$$(1) \qquad a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} = \begin{cases} \det(A) & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

Demonstração. Se $i = j$, (1) é simplesmente a expansão em cofatores de $\det(A)$ em relação à i-ésima linha de A. Para provar (1) no caso em que $i \neq j$, seja A^* a matriz obtida substituindo-se a j-ésima linha de A pela sua i-ésima linha.

$$A^* = \begin{array}{c} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & & & \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \end{array} \begin{array}{l} \\ \\ j\text{-ésima linha} \\ \\ \\ \\ \\ \end{array}$$

Como A^* tem duas linhas idênticas, seu determinante tem que ser zero. Expandindo $\det(A^*)$ em cofatores em relação à sua j-ésima linha, obtemos

$$\begin{aligned} 0 = \det(A^*) &= a_{i1}A^*_{j1} + a_{i2}A^*_{j2} + \cdots + a_{in}A^*_{jn} \\ &= a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} \end{aligned}$$ □

Vamos considerar agora os efeitos das operações elementares sobre o valor do determinante. Vamos começar com a operação elementar II.

OPERAÇÃO ELEMENTAR II

Uma linha de A é multiplicada por uma constante diferente de zero.

Vamos denotar por E a matriz elementar do tipo II obtida multiplicando-se a i-ésima linha de I por uma constante não-nula α. Expandindo $\det(EA)$ em cofatores em relação à i-ésima linha, obtemos

$$\begin{aligned} \det(EA) &= \alpha a_{i1}A_{i1} + \alpha a_{i2}A_{i2} + \cdots + \alpha a_{in}A_{in} \\ &= \alpha(a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}) \\ &= \alpha \det(A) \end{aligned}$$

Em particular,

$$\det(E) = \det(EI) = \alpha \det(I) = \alpha$$

e, portanto,

$$\det(EA) = \alpha \det(A) = \det(E)\det(A)$$

OPERAÇÃO ELEMENTAR III

Soma-se um múltiplo de uma linha a uma outra.

Seja E a matriz elementar de tipo III obtida somando-se c vezes a i-ésima linha de I à sua j-ésima linha. Como E é triangular e seus elementos diagonais são todos iguais a 1, temos $\det(E) = 1$. Vamos mostrar que

$$\det(EA) = \det(A) = \det(E)\det(A)$$

Expandindo $\det(EA)$ em cofatores em relação à j-ésima linha e usando o Lema 2.2.1, obtemos

$$\begin{aligned}\det(EA) &= (a_{j1} + ca_{i1})A_{j1} + (a_{j2} + ca_{i2})A_{j2} \\ &\quad + \cdots + (a_{jn} + ca_{in})A_{jn} \\ &= (a_{j1}A_{j1} + \cdots + a_{jn}A_{jn}) \\ &\quad + c(a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}) \\ &= \det(A)\end{aligned}$$

Logo,

$$\det(EA) = \det(A) = \det(E)\det(A)$$

OPERAÇÃO ELEMENTAR I

Troca-se a ordem de duas linhas de A.

Para ver o efeito da operação elementar I, observe que ela pode ser feita usando-se as operações elementares II e III. Vamos ilustrar isso para o caso de matrizes 3×3.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Subtraindo a terceira da segunda linha, obtemos

$$A^{(1)} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} - a_{31} & a_{22} - a_{32} & a_{23} - a_{33} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

A seguir, somamos a segunda linha de $A^{(1)}$ à sua terceira linha:

$$A^{(2)} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} - a_{31} & a_{22} - a_{32} & a_{23} - a_{33} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

Subtraindo a terceira da segunda linha, obtemos

$$A^{(3)} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ -a_{31} & -a_{32} & -a_{33} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

Como todas essas matrizes foram obtidas usando-se apenas as operações elementares III, temos

$$\det(A) = \det(A^{(1)}) = \det(A^{(2)}) = \det(A^{(3)})$$

Finalmente, multiplicando a segunda linha de $A^{(3)}$ por -1, obtemos

$$A^{(4)} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

Como usamos uma operação elementar II, temos

$$\det(A^{(4)}) = -1\det(A^{(3)}) = -\det(A)$$

$A^{(4)}$ é, simplesmente, a matriz obtida trocando-se a segunda e terceira linhas de A.

Esse mesmo argumento pode ser usado para matrizes $n \times n$ para mostrar que, sempre que trocamos duas linhas, o sinal do determinante muda. Então, se A é $n \times n$ e E_{ij} é a matriz elementar $n \times n$ obtida trocando-se as i-ésima e j-ésima linhas de I, então

$$\det(E_{ij}A) = -\det(A)$$

Em particular,

$$\det(E_{ij}) = \det(E_{ij}I) = -\det(I) = -1$$

Logo, qualquer que seja a matriz elementar E de tipo I,

$$\det(EA) = -\det(A) = \det(E)\det(A)$$

Resumindo, se E é uma matriz elementar, então

$$\det(EA) = \det(E)\det(A)$$

onde

(2) $$\det(E) = \begin{cases} -1 & \text{se } E \text{ é de tipo I} \\ \alpha \neq 0 & \text{se } E \text{ é de tipo II} \\ 1 & \text{se } E \text{ é de tipo III} \end{cases}$$

Resultados análogos são válidos para operações elementares sobre as colunas. De fato, se E é uma matriz elementar, então

$$\begin{aligned} \det(AE) &= \det\big((AE)^T\big) = \det\big(E^T A^T\big) \\ &= \det\big(E^T\big)\det\big(A^T\big) = \det(E)\det(A) \end{aligned}$$

Logo, o efeito que as operações elementares sobre as linhas ou colunas tem no valor do determinante pode ser resumido assim:

I. Trocar a ordem de duas linhas ou colunas de uma matriz troca o sinal do determinante.
II. Multiplicar uma única linha ou coluna de uma matriz por um escalar faz com que o determinante fique multiplicado por esse escalar.
III. Somar um múltiplo de uma linha a uma outra não altera o valor do determinante.

Observação. Como corolário de III, se uma linha ou coluna de uma matriz é um múltiplo de outra, então o determinante da matriz tem que ser igual a zero.

Segue de (2) que toda matriz elementar tem determinante diferente de zero. Essa observação pode ser usada para provar o teorema a seguir.

Teorema 2.2.2. *Uma matriz A $n \times n$ é singular se e somente se*

$$\det(A) = 0$$

Demonstração. A matriz A pode ser reduzida à sua forma escada reduzida por linhas através de um número finito de operações elementares; logo,

$$U = E_k E_{k-1} \cdots E_1 A$$

onde U está em forma escada reduzida por linhas e as matrizes E_i são elementares.

$$\begin{aligned} \det(U) &= \det(E_k E_{k-1} \cdots E_1 A) \\ &= \det(E_k)\det(E_{k-1}) \cdots \det(E_1)\det(A) \end{aligned}$$

Como os determinantes das matrizes E_i são todos diferentes de zero, temos que $\det(A) = 0$ se e somente se $\det(U) = 0$. Se A é singular, então U tem uma linha contendo apenas elementos nulos e, portanto, $\det(U) = 0$. Se A é invertível, então U é triangular e todos os elementos da diagonal são iguais a 1, logo $\det(U) = 1$. □

Podemos obter, da demonstração do Teorema 2.2.2, um método para calcular $\det(A)$. Coloque A em sua forma escada reduzida por linhas.

$$U = E_k E_{k-1} \cdots E_1 A$$

Se a última linha de U tiver todos os elementos nulos, então A é singular e $\det(A) = 0$. Caso contrário, A é invertível e

$$\det(A) = [\det(E_k)\det(E_{k-1})\cdots\det(E_1)]^{-1}$$

De fato, se A for invertível, é mais fácil colocar A em forma triangular. Isso pode ser feito usando-se apenas as operações I e III. Portanto,

$$T = E_m E_{m-1}\cdots E_1 A$$

logo,

$$\det(A) = \pm\det(T) = \pm t_{11}t_{22}\cdots t_{nn}$$

O sinal será positivo se a operação elementar I for usada um número par de vezes e negativo em caso contrário.

EXEMPLO 1. Calcule

$$\begin{vmatrix} 2 & 1 & 3 \\ 4 & 2 & 1 \\ 6 & -3 & 4 \end{vmatrix}$$

SOLUÇÃO

$$\begin{vmatrix} 2 & 1 & 3 \\ 4 & 2 & 1 \\ 6 & -3 & 4 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 3 \\ 0 & 0 & -5 \\ 0 & -6 & -5 \end{vmatrix} = (-1)\begin{vmatrix} 2 & 1 & 3 \\ 0 & -6 & -5 \\ 0 & 0 & -5 \end{vmatrix}$$

$$= (-1)(2)(-6)(-5) = -60$$ □

Temos agora dois métodos para calcular o determinante de uma matriz A $n \times n$. Se $n > 3$ e A não tem elementos nulos, a redução à forma escada reduzida por linhas é o método mais eficiente, no sentido de que envolve um menor número de operações aritméticas. A Tabela 1 mostra o número de operações aritméticas efetuadas em cada método para $n = 2, 3, 4, 5, 10$. Não é difícil obter fórmulas gerais para o número de operações em cada um dos métodos (ver Exercícios 16 e 17).

TABELA 1

	Cofatores		Redução	
n	Somas	Multiplicações	Somas	Multiplicações e Divisões
2	1	2	1	3
3	5	9	5	10
4	23	40	14	23
5	119	205	30	45
10	3.628.799	6.235.300	285	339

Vimos que, para qualquer matriz elementar E,

$$\det(EA) = \det(E)\det(A) = \det(AE)$$

Isso é um caso particular do teorema a seguir.

Teorema 2.2.3. *Se A e B são matrizes $n \times n$, então*

$$\det(AB) = \det(A)\det(B)$$

Demonstração. Se B é singular, pelo Teorema 1.4.3, AB também é singular (ver Exercício 15 do Cap. 1, Seção 4), e, portanto,

$$\det(AB) = 0 = \det(A)\det(B)$$

Se B é invertível, B pode ser escrita como um produto de matrizes elementares. Já vimos que o resultado é válido para matrizes elementares. Logo,

$$\begin{aligned}\det(AB) &= \det(AE_kE_{k-1}\cdots E_1)\\ &= \det(A)\det(E_k)\det(E_{k-1})\cdots\det(E_1)\\ &= \det(A)\det(E_kE_{k-1}\cdots E_1)\\ &= \det(A)\det(B)\end{aligned}$$

□

Se A é singular, então o valor calculado de $\det(A)$ usando aritmética exata tem que ser 0. No entanto, dificilmente vamos chegar a esse resultado se os cálculos forem feitos por computador. Como os computadores usam um sistema numérico finito, erros de aproximação são em geral inevitáveis. Conseqüentemente, é mais provável que o valor calculado de $\det(A)$ esteja apenas próximo de 0. Devido a erros de aproximação, é praticamente impossível determinar, utilizando um computador, se uma matriz é ou não exatamente singular. Em aplicações envolvendo computadores, muitas vezes faz mais sentido perguntar se uma matriz é "aproximadamente" singular. Em geral, o valor de $\det(A)$ não é um bom indicador de quão próxima uma matriz está de ser ou não singular. No Cap. 7 vamos discutir como determinar se uma matriz é aproximadamente singular ou não.

EXERCÍCIOS

1. Calcule cada um dos determinantes a seguir diretamente, analisando a matriz.

(a) $\begin{vmatrix} 0 & 0 & 3 \\ 0 & 4 & 1 \\ 2 & 3 & 1 \end{vmatrix}$ (b) $\begin{vmatrix} 1 & 1 & 1 & 3 \\ 0 & 3 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ -1 & -1 & -1 & 2 \end{vmatrix}$ (c) $\begin{vmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{vmatrix}$

2. Seja

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

(a) Use o método de redução para calcular $\det(A)$.
(b) Use o valor de $\det(A)$ para calcular

$$\begin{vmatrix} 0 & 1 & 2 & 3 \\ -2 & -2 & 3 & 3 \\ 1 & 2 & -2 & -3 \\ 1 & 1 & 1 & 1 \end{vmatrix} + \begin{vmatrix} 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & 4 & 4 \\ 2 & 3 & -1 & -2 \end{vmatrix}$$

3. Para cada uma das matrizes a seguir, calcule o determinante e diga se a matriz é singular ou invertível.

(a) $\begin{pmatrix} 3 & 1 \\ 6 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 3 & 1 \\ 4 & 2 \end{pmatrix}$ (c) $\begin{pmatrix} 3 & 3 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & 3 \end{pmatrix}$

(d) $\begin{pmatrix} 2 & 1 & 1 \\ 4 & 3 & 5 \\ 2 & 1 & 2 \end{pmatrix}$ (e) $\begin{pmatrix} 2 & -1 & 3 \\ -1 & 2 & -2 \\ 1 & 4 & 0 \end{pmatrix}$ (f) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & -1 & 3 & 2 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 7 & 3 \end{pmatrix}$

4. Escolha todos os valores possíveis de c que tornam a matriz a seguir singular.

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 9 & c \\ 1 & c & 3 \end{pmatrix}$$

5. Sejam A uma matriz $n \times n$ e α um escalar. Mostre que

$$\det(\alpha A) = \alpha^n \det(A)$$

6. Seja A uma matriz invertível. Mostre que

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

7. Sejam A e B matrizes 3×3 com $\det(A) = 4$ e $\det(B) = 5$. Encontre o valor de:
(a) $\det(AB)$ (b) $\det(3A)$ (c) $\det(2AB)$ (d) $\det(A^{-1}B)$

8. Sejam E_1, E_2, E_3 matrizes elementares de tipos I, II, III, respectivamente, e seja A uma matriz 3×3 com $\det(A) = 6$. Suponha que E_2 foi obtida multiplicando-se a segunda linha de I por 3. Encontre o valor de cada um dos determinantes a seguir.
(a) $\det(E_1 A)$ (b) $\det(E_2 A)$ (c) $\det(E_3 A)$
(d) $\det(AE_1)$ (e) $\det(E_1^2)$ (f) $\det(E_1 E_2 E_3)$

9. Sejam A e B matrizes equivalentes por linhas e suponha que B pode ser obtida de A usando-se apenas as operações elementares I e III. Qual a relação entre os valores de $\det(A)$ e $\det(B)$? Se B puder ser obtida de A usando-se apenas operações elementares III, qual a relação entre os valores de $\det(A)$ e $\det(B)$? Justifique suas respostas.

10. Considere a matriz de Vandermonde 3×3

$$V = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{pmatrix}$$

(a) Mostre que $\det(V) = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2)$.
(b) Que condições os escalares x_1, x_2 e x_3 têm que satisfazer para que V seja invertível?

11. Suponha que a matriz A 3×3 fatora em um produto

$$\begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix}$$

Determine o valor de $\det(A)$.

12. Sejam A e B matrizes $n \times n$. Prove que o produto AB é invertível se e somente se A e B são ambas invertíveis.

13. Sejam A e B matrizes $n \times n$. Prove que, se $AB = I$, então $BA = I$. Qual o significado desse resultado para a definição de uma matriz invertível?

14. Seja A uma matriz invertível $n \times n$ com um cofator não-nulo A_{nn} e defina

$$c = \frac{\det(A)}{A_{nn}}$$

Mostre que, se subtrairmos c de a_{nn}, a matriz resultante será singular.

15. Sejam $\mathbf{x}$ e $\mathbf{y}$ elementos de R^3 e seja $\mathbf{z}$ um vetor em R^3 cujas coordenadas são definidas por

$$z_1 = \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix}, \qquad z_2 = -\begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix}, \qquad z_3 = \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}$$

Sejam

$$X = (\mathbf{x}, \mathbf{x}, \mathbf{y})^T \qquad \text{e} \qquad Y = (\mathbf{x}, \mathbf{y}, \mathbf{y})^T$$

Mostre que

$$\mathbf{x}^T\mathbf{z} = \det(X) = 0 \qquad \text{e} \qquad \mathbf{y}^T\mathbf{z} = \det(Y) = 0$$

16. Mostre que o cálculo do determinante de uma matriz $n \times n$ por expansão em cofatores envolve $(n! - 1)$ somas e $\sum_{k-1}^{n-1} n!/k!$ multiplicações.

17. Mostre que o cálculo do determinante de uma matriz $n \times n$ pelo método de redução envolve $[n(n-1)(2n-1)]/6$ somas e $[(n-1)(n^2+n+3)]/3$ multiplicações e divisões. [***Sugestão:*** No i-ésimo passo do processo de redução, são necessárias $n - i$ divisões para se calcularem os múltiplos da i-ésima linha que vão ser subtraídos das linhas restantes abaixo do pivô. É necessário depois calcular os novos valores para os $(n-i)^2$ elementos nas linhas e colunas de $i + 1$ até n.]

3 REGRA DE CRAMER

Nesta seção, vamos aprender um método para calcular a inversa de uma matriz invertível A usando determinantes. Vamos, também, aprender um método para resolver $A\mathbf{x} = \mathbf{b}$ usando determinantes. Ambos os métodos dependem do Lema 2.2.1 da Seção 2.

A ADJUNTA DE UMA MATRIZ

Seja A uma matriz $n \times n$. Vamos definir uma nova matriz, chamada *adjunta* de A, por

$$\operatorname{adj} A = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$$

Então, para formar a adjunta, colocamos no lugar de cada elemento seu cofator e depois transpomos a matriz resultante. Pelo Lema 2.2.1,

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} = \begin{cases} \det(A) & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

Portanto,

$$A(\operatorname{adj} A) = \det(A)I$$

Se A é invertível, $\det(A)$ é um escalar diferente de zero e podemos escrever

$$A\left(\frac{1}{\det(A)} \operatorname{adj} A\right) = I$$

Logo,

$$\boxed{A^{-1} = \frac{1}{\det(A)} \operatorname{adj} A}$$

EXEMPLO 1. Para uma matriz 2×2,

$$\operatorname{adj} A = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

Se A é invertível, então

$$A^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \qquad \square$$

EXEMPLO 2. Seja

$$A = \begin{pmatrix} 2 & 1 & 2 \\ 3 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$$

Calcule adj A e A^{-1}.

SOLUÇÃO

$$\operatorname{adj} A = \begin{pmatrix} \begin{vmatrix} 2 & 2 \\ 2 & 3 \end{vmatrix} & -\begin{vmatrix} 3 & 2 \\ 1 & 3 \end{vmatrix} & \begin{vmatrix} 3 & 2 \\ 1 & 2 \end{vmatrix} \\ -\begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} & \begin{vmatrix} 2 & 2 \\ 1 & 3 \end{vmatrix} & -\begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} \\ \begin{vmatrix} 1 & 2 \\ 2 & 2 \end{vmatrix} & -\begin{vmatrix} 2 & 2 \\ 3 & 2 \end{vmatrix} & \begin{vmatrix} 2 & 1 \\ 3 & 2 \end{vmatrix} \end{pmatrix}^T$$

$$= \begin{pmatrix} 2 & 1 & -2 \\ -7 & 4 & 2 \\ 4 & -3 & 1 \end{pmatrix}$$

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj} A = \frac{1}{5} \begin{pmatrix} 2 & 1 & -2 \\ -7 & 4 & 2 \\ 4 & -3 & 1 \end{pmatrix} \qquad \square$$

Usando a fórmula

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj} A$$

podemos obter uma fórmula para representar a solução do sistema $A\mathbf{x} = \mathbf{b}$ em termos de determinantes.

Teorema 2.3.1 (Regra de Cramer). *Seja A uma matriz invertível $n \times n$ e seja $\mathbf{b} \in R^n$. Seja A_i a matriz obtida substituindo-se a i-ésima coluna de A por $\mathbf{b}$. Se $\mathbf{x}$ for a única solução de $A\mathbf{x} = \mathbf{b}$, então*

$$x_i = \frac{\det(A_i)}{\det(A)} \qquad \textit{para} \quad i = 1, 2, \ldots, n$$

Demonstração. Como

$$\mathbf{x} = A^{-1}\mathbf{b} = \frac{1}{\det(A)}(\operatorname{adj} A)\mathbf{b}$$

temos que

$$x_i = \frac{b_1 A_{1i} + b_2 A_{2i} + \cdots + b_n A_{ni}}{\det(A)} = \frac{\det(A_i)}{\det(A)} \qquad \square$$

EXEMPLO 3. Use a regra de Cramer para resolver

$$\begin{aligned} x_1 + 2x_2 + \ x_3 &= 5 \\ 2x_1 + 2x_2 + \ x_3 &= 6 \\ x_1 + 2x_2 + 3x_3 &= 9 \end{aligned}$$

SOLUÇÃO

$$\det(A) = \begin{vmatrix} 1 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 2 & 3 \end{vmatrix} = -4 \qquad \det(A_1) = \begin{vmatrix} 5 & 2 & 1 \\ 6 & 2 & 1 \\ 9 & 2 & 3 \end{vmatrix} = -4$$

$$\det(A_2) = \begin{vmatrix} 1 & 5 & 1 \\ 2 & 6 & 1 \\ 1 & 9 & 3 \end{vmatrix} = -4 \qquad \det(A_3) = \begin{vmatrix} 1 & 2 & 5 \\ 2 & 2 & 6 \\ 1 & 2 & 9 \end{vmatrix} = -8$$

Portanto,

$$x_1 = \frac{-4}{-4} = 1, \qquad x_2 = \frac{-4}{-4} = 1, \qquad x_3 = \frac{-8}{-4} = 2$$

□

A regra de Cramer nos dá um método conveniente para escrever a solução de um sistema de equações lineares $n \times n$ em função de determinantes. Para calcular a solução, no entanto, temos que calcular $n + 1$ determinantes de ordem n. O cálculo de apenas dois desses determinantes envolve, em geral, mais operações do que resolver o sistema pelo método de Gauss.

APLICAÇÃO: MENSAGENS CODIFICADAS

Um modo simples de codificar mensagens é associar um valor inteiro a cada letra do alfabeto e mandar a mensagem como uma lista de números. Por exemplo, a mensagem

MANDE FOTO

poderia ser codificada por

7, 4, 10, 21, 8, 11, 2, 20, 2

Aqui, M é representada por 7, A por 4, e assim por diante. Infelizmente, esse tipo de código é, em geral, muito fácil de quebrar. No entanto, podemos disfarçar ainda mais a mensagem usando multiplicação de matrizes. Se A é uma matriz cujos elementos são todos inteiros e cujo determinante é ± 1, então, como $A^{-1} = \pm \operatorname{adj} A$, os elementos de A^{-1} vão ser todos inteiros. Podemos usar tal matriz para transformar a mensagem. A mensagem transformada será mais difícil de quebrar. Para ilustrar essa técnica, considere

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 3 \\ 2 & 3 & 2 \end{pmatrix}$$

A mensagem codificada é colocada nas colunas de uma matriz B com três linhas.

$$B = \begin{pmatrix} 7 & 21 & 2 \\ 4 & 8 & 20 \\ 10 & 11 & 2 \end{pmatrix}$$

O produto

$$AB = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 3 \\ 2 & 3 & 2 \end{pmatrix} \begin{pmatrix} 7 & 21 & 2 \\ 4 & 8 & 20 \\ 10 & 11 & 2 \end{pmatrix} = \begin{pmatrix} 25 & 48 & 44 \\ 64 & 115 & 110 \\ 46 & 88 & 68 \end{pmatrix}$$

fornece a mensagem codificada que deve ser enviada:

$$25, 64, 46, 48, 115, 88, 44, 110, 68$$

A pessoa que receber a mensagem pode decodificá-la multiplicando-a por A^{-1}

$$\begin{pmatrix} 1 & -1 & 1 \\ 2 & 0 & -1 \\ -4 & 1 & 1 \end{pmatrix} \begin{pmatrix} 25 & 48 & 44 \\ 64 & 115 & 110 \\ 46 & 88 & 68 \end{pmatrix} = \begin{pmatrix} 7 & 21 & 2 \\ 4 & 8 & 20 \\ 10 & 11 & 2 \end{pmatrix}$$

Para construir uma matriz de codificação A, começamos com a identidade I e aplicamos, sucessivamente, operações elementares III, tendo o cuidado de somar múltiplos inteiros de uma linha a outra. A operação I também pode ser usada. A matriz resultante A vai ter elementos inteiros, e como

$$\det(A) = \pm \det(I) = \pm 1$$

A^{-1} também vai ter apenas elementos inteiros.

REFERÊNCIAS

1. Hansen, Robert, *Two-Year College Mathematics Journal, 13*(1), 1982.

EXERCÍCIOS

1. Para cada uma das matrizes a seguir, calcule (i) det(A), (ii) adj A e (iii) A^{-1}.

(a) $A = \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix}$ (b) $A = \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix}$

(c) $A = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 1 & 1 \\ -2 & 2 & -1 \end{pmatrix}$ (d) $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

2. Use a regra de Cramer para resolver cada um dos sistemas a seguir.

(a) $\begin{aligned} x_1 + 2x_2 &= 3 \\ 3x_1 - x_2 &= 1 \end{aligned}$ (b) $\begin{aligned} 2x_1 + 3x_2 &= 2 \\ 3x_1 + 2x_2 &= 5 \end{aligned}$ (c) $\begin{aligned} 2x_1 + x_2 - 3x_3 &= 0 \\ 4x_1 + 5x_2 + x_3 &= 8 \\ -2x_1 - x_2 + 4x_3 &= 2 \end{aligned}$

(d) $\begin{aligned} x_1 + 3x_2 + x_3 &= 1 \\ 2x_1 + x_2 + x_3 &= 5 \\ -2x_1 + 2x_2 - x_3 &= -8 \end{aligned}$ (e) $\begin{aligned} x_1 + x_2 \qquad\qquad &= 0 \\ x_2 + x_3 - 2x_4 &= 1 \\ x_1 \qquad + 2x_3 + x_4 &= 0 \\ x_1 + x_2 \qquad + x_4 &= 0 \end{aligned}$

3. Seja

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

Determine o elemento (2,3) de A^{-1} calculando um quociente entre dois determinantes.

4. Seja A a matriz do Exercício 3. Calcule a terceira coluna de A^{-1} usando a regra de Cramer para resolver $A\mathbf{x} = \mathbf{e}_3$.

5. Seja

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$$

(a) Calcule o determinante de A. A é invertível?
(b) Calcule adj A e o produto A adj A.

6. Se A é singular, o que você pode dizer sobre o produto A adj A?

7. Denote por B_j a matriz obtida substituindo-se a j-ésima coluna da matriz identidade por um vetor $\mathbf{b} = (b_1, \ldots, b_n)^T$. Use a regra de Cramer para mostrar que

$$b_j = \det(B_j) \quad \text{para } j = 1, \ldots, n$$

8. Seja A uma matriz $n \times n$ invertível com $n > 1$. Mostre que

$$\det(\text{adj } A) = (\det(A))^{n-1}$$

9. Seja A uma matriz 4×4. Seja

$$\text{adj } A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 4 & 3 & 2 \\ 0 & -2 & -1 & 2 \end{pmatrix}$$

(a) Calcule o valor de det(adj A). Qual deve ser o valor de det(A)? [***Sugestão:*** Use o resultado do Exercício 8.]
(b) Encontre A.

10. Mostre que, se A é invertível, então adj A é invertível e

$$(\text{adj } A)^{-1} = \det(A^{-1})A = \text{adj } A^{-1}$$

11. Mostre que, se A é singular, então adj A também é singular.

12. Mostre que, se det(A) = 1, então

$$\text{adj}(\text{adj } A) = A$$

13. Suponha que Q é uma matriz com a propriedade de que $Q^{-1} = Q^T$. Mostre que

$$q_{ij} = \frac{Q_{ij}}{\det(Q)}$$

14. Na codificação de uma mensagem, um espaço em branco é representado por 0, um A por 1, um B por 2, um C por 3 e assim por diante. A mensagem foi transformada usando a matriz

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

e enviada como

$15, 4, -4, 3, -32, 33, -1, 12, -34, 34, 5, 10, 7, 11, -15, 21, 6, 3, 6, -6, 13, 3, -15, 18, -19, 19, 3, 15, -18, 19, -1, 1$

Qual é a mensagem?

EXERCÍCIOS COM O MATLAB PARA O CAPÍTULO 2

Os quatro primeiros exercícios envolvem matrizes inteiras e ilustram algumas das propriedades de determinantes abordadas neste capítulo. Os dois últimos exercícios ilustram algumas das diferenças que podem aparecer quando se trabalha com aritmética de ponto flutuante.

Teoricamente, o valor do determinante deveria nos dizer se a matriz é invertível ou não. Entretanto, se a matriz é singular e seu determinante é calculado usando-se aritmética de precisão finita, devido aos erros de aproximação, o valor calculado do determinante pode não ser igual a zero. Um valor calculado próximo de zero não significa necessariamente que a matriz seja singular nem mesmo que ela esteja perto de ser singular. Além disso, uma matriz pode ser singular ou quase singular e ter um determinante cujo valor calculado não está nem perto de zero (ver Exercício 6).

1. Gere matrizes aleatórias 5×5 com elementos inteiros digitando

$$A = \texttt{round}(10 * \texttt{rand}(5)) \text{ and } B = \texttt{round}(20 * (\texttt{rand}(5) - 0.5))$$

Use o MATLAB para calcular cada um dos pares de números a seguir. Em cada caso, verifique se o primeiro é igual ou não ao segundo.

(a) $\det(A)$	$\det(A^T)$
(b) $\det(A+B)$	$\det(A) + \det(B)$
(c) $\det(AB)$	$\det(A)\det(B)$
(d) $\det(A^T B^T)$	$\det(A^T)\det(B^T)$
(e) $\det(A^{-1})$	$1/\det(A)$
(f) $\det(AB^{-1})$	$\dfrac{\det(A)}{\det(B)}$

2. Quadrados mágicos $n \times n$ são invertíveis? Use o MATLAB para calcular `det (magic(n))` para $n = 3, 4, \ldots, 10$. O que parece estar acontecendo? Verifique os casos $n = 24$ e $n = 25$ para ver se o padrão ainda é válido.

3. Defina $A = \texttt{round}(10*\texttt{rand}(6))$. Para cada item a seguir, use o MATLAB para calcular a matriz indicada. Diga qual a relação entre essa segunda matriz e A e calcule o determinante de ambas. Qual a relação entre os determinantes?

(a) $B = A;\ \ B(2,:) = A(1,:);\ \ B(1,:) = A(2,:)$

(b) $C = A;\ \ C(3,:) = 4 * A(3,:)$

(c) $D = A;\ \ D(5,:) = A(5,:) + 2 * A(4,:)$

4. Podemos gerar uma matriz aleatória A 6×6 cujos elementos são todos iguais a um ou zero digitando

$$A = \texttt{round}(\texttt{rand}(6))$$

(a) Qual a percentagem de matrizes contendo apenas 0 e 1 que são singulares? Você pode estimar esse percentual usando o MATLAB, digitando

$$y = \texttt{zeros}(1, 100);$$

e depois gerando 100 matrizes de teste e fazendo $y(j) = 1$, se a j-ésima matriz for singular e 0 em caso contrário. A maneira mais fácil de fazer isso no MATLAB é usar um *for loop*. Gere o código da seguinte maneira:

```
for j = 1 : 100
    A = round(rand(6));
    y(j) = (det(A) == 0);
end
```

(*Observação:* Um ponto-e-vírgula no final de uma linha faz com que não apareça o cálculo correspondente. Recomendamos que você coloque um ponto-e-vírgula ao final de cada linha de cálculo dentro de um `for`.) Para determinar quantas matrizes foram geradas, use o comando `sum(y)`. Qual a percentagem de matrizes singulares geradas?

(b) Para qualquer inteiro positivo n, podemos gerar uma matriz aleatória A 6×6 cujos elementos são inteiros entre 0 e n digitando

$$A = \texttt{round}(n * \texttt{rand}(6))$$

Qual a percentagem de matrizes aleatórias inteiras geradas dessa maneira que são singulares se $n = 3$? E se $n = 6$? E se $n = 10$? Podemos estimar as respostas para essas perguntas usando o MATLAB. Em cada caso, gere 100 matrizes e determine quantas são singulares.

5. Se uma matriz é sensitiva a erros de aproximação, o valor calculado de seu determinante pode ser drasticamente diferente de seu valor exato. Para um exemplo disso, defina

$$U = \texttt{round}(100 * \texttt{rand}(10)); \qquad U = \texttt{triu}(U, 1) + 0.1 * \texttt{eye}(10)$$

Teoricamente,

$$\det(U) = \det(U^T) = 10^{-10}$$

e

$$\det(UU^T) = \det(U)\det(U^T) = 10^{-20}$$

Calcule $\det(U)$, $\det(U')$ e $\det(U * U')$ usando o MATLAB. Os valores calculados coincidem com os valores teóricos?

6. Use o MATLAB para construir uma matriz A fazendo

$$A = \texttt{vander}(1:6); \qquad A = A - \texttt{diag}(\texttt{sum}(A'))$$

(a) Por construção, a soma dos elementos de cada linha deveria ser zero. Para verificar isso, defina $\mathbf{x} = \texttt{ones}(6, 1)$ e use o MATLAB para calcular o produto $A\mathbf{x}$. A matriz A deveria ser singular. Por quê? Explique. Use as funções `det` e `inv` para calcular os valores de $\det(A)$ e A^{-1}. Qual dessas funções do MATLAB é o indicador mais confiável de singularidade?

(b) Use o MATLAB para calcular $\det(A^T)$. Os valores calculados de $\det(A)$ e $\det(A^T)$ são iguais? Uma outra maneira de verificar se uma matriz é singular é calcular sua forma escada reduzida por linhas. Use o MATLAB para calcular as formas escadas reduzidas por linhas de A e A^T.

(c) Faça $B = A * A'$. O valor exato de $\det(B)$ deveria ser 0. Por quê? Explique. Use o MATLAB para calcular $\det(B)$. O valor calculado do determinante é igual ou aproximadamente igual ao valor exato? Calcule a forma escada reduzida por linhas de B e verifique que a matriz é, de fato, singular.

CAPÍTULO 3

ESPAÇOS VETORIAIS

As operações de soma e multiplicação por um escalar são usadas em diversos contextos em matemática. Independentemente do contexto, no entanto, essas operações obedecem, em geral, ao mesmo conjunto de regras aritméticas. Logo, uma teoria geral de sistemas matemáticos envolvendo soma e multiplicação por escalar vai ter aplicação em diversas áreas da matemática. Sistemas matemáticos desse tipo são chamados de espaços vetoriais ou espaços lineares. Vamos definir espaços vetoriais neste capítulo e desenvolver parte da teoria geral de espaços vetoriais.

1 DEFINIÇÃO E EXEMPLOS

Nesta seção, vamos apresentar a definição formal de espaço vetorial. Antes disso, no entanto, é instrutivo considerar alguns exemplos. Vamos começar com os espaços euclidianos R^n.

ESPAÇOS VETORIAIS EUCLIDIANOS

Talvez os espaços vetoriais mais elementares sejam os espaços vetoriais euclidianos R^n, $n = 1, 2, \ldots$. Vamos considerar primeiro, por simplicidade, R^2. Vetores não-nulos em R^2 podem ser representados geometricamente por segmentos de reta orientados. Essa representação geométrica nos ajuda a visualizar como as operações de multiplicação por um escalar e de soma funcionam em R^2. Dado um vetor não-nulo $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, podemos associar a ele um segmento orientado no plano do ponto $(0, 0)$ ao ponto (x_1, x_2) (ver Fig. 3.1.1). Identificando os segmentos orientados com mesmo comprimento, mesma direção e mesmo sentido (Fig. 3.1.2), $\mathbf{x}$ pode ser representado por qualquer segmento orientado de (a, b) a

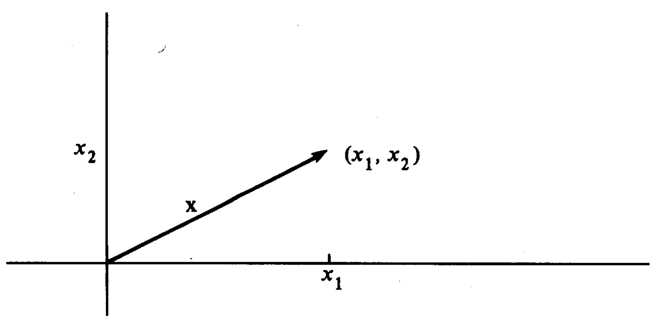


FIG. 3.1.1

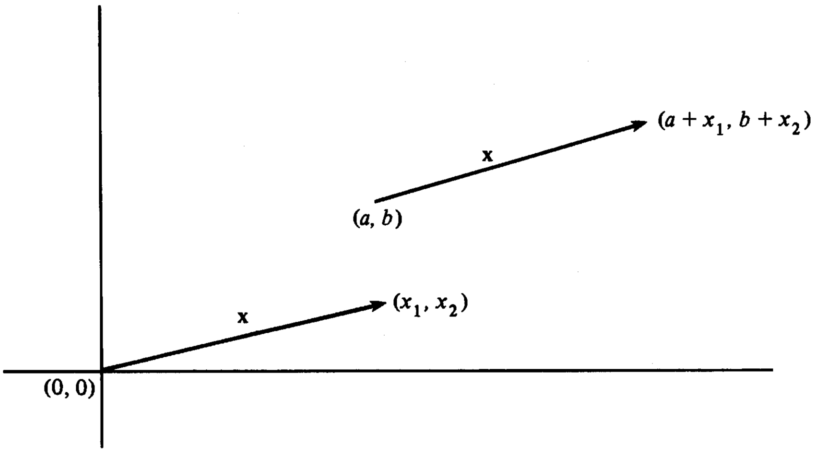


FIG. 3.1.2

$(a + x_1, b + x_2)$. Por exemplo, o vetor $\mathbf{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ em R^2 pode ser representado tanto pelo segmento orientado de (2, 2) a (4, 3) quanto pelo segmento orientado de $(-1, -1)$ a (1, 0), como ilustrado na Fig. 3.1.3.

Podemos pensar no comprimento de um vetor $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ como sendo o comprimento de qualquer segmento orientado que o representa. O comprimento do segmento orientado de (0, 0) a (x_1, x_2) é $\sqrt{x_1^2 + x_2^2}$ (ver Fig. 3.1.4). Para cada vetor $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ e para cada escalar α, o produto $\alpha\mathbf{x}$ é definido por

$$\alpha \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \end{pmatrix}$$

Por exemplo, como ilustrado na Fig. 3.1.5, se $\mathbf{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, então

$$3\mathbf{x} = \begin{pmatrix} 6 \\ 3 \end{pmatrix}, \quad -\mathbf{x} = \begin{pmatrix} -2 \\ -1 \end{pmatrix}, \quad -2\mathbf{x} = \begin{pmatrix} -4 \\ -2 \end{pmatrix}$$

O vetor $3\mathbf{x}$ tem a mesma direção e o mesmo sentido que $\mathbf{x}$, mas seu comprimento é três vezes o de $\mathbf{x}$. O vetor $-\mathbf{x}$ tem o mesmo comprimento e a mesma direção que $\mathbf{x}$, mas sentido oposto. O vetor $-2\mathbf{x}$ tem o

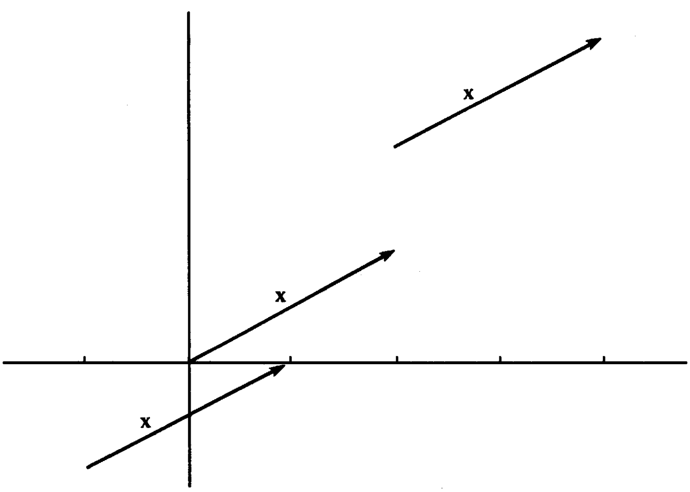


FIG. 3.1.3

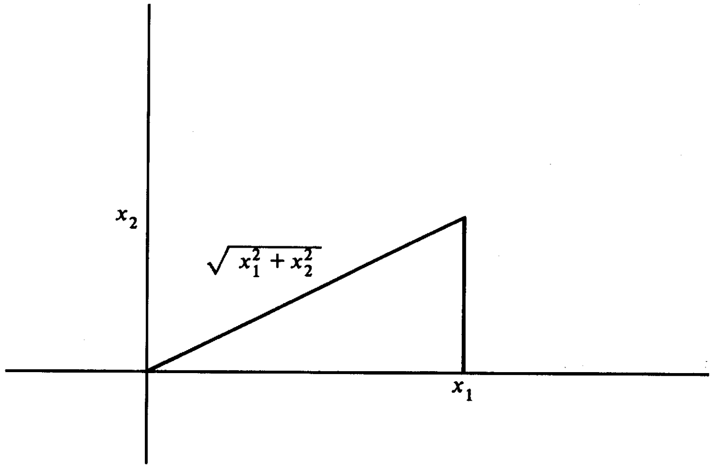


FIG. 3.1.4

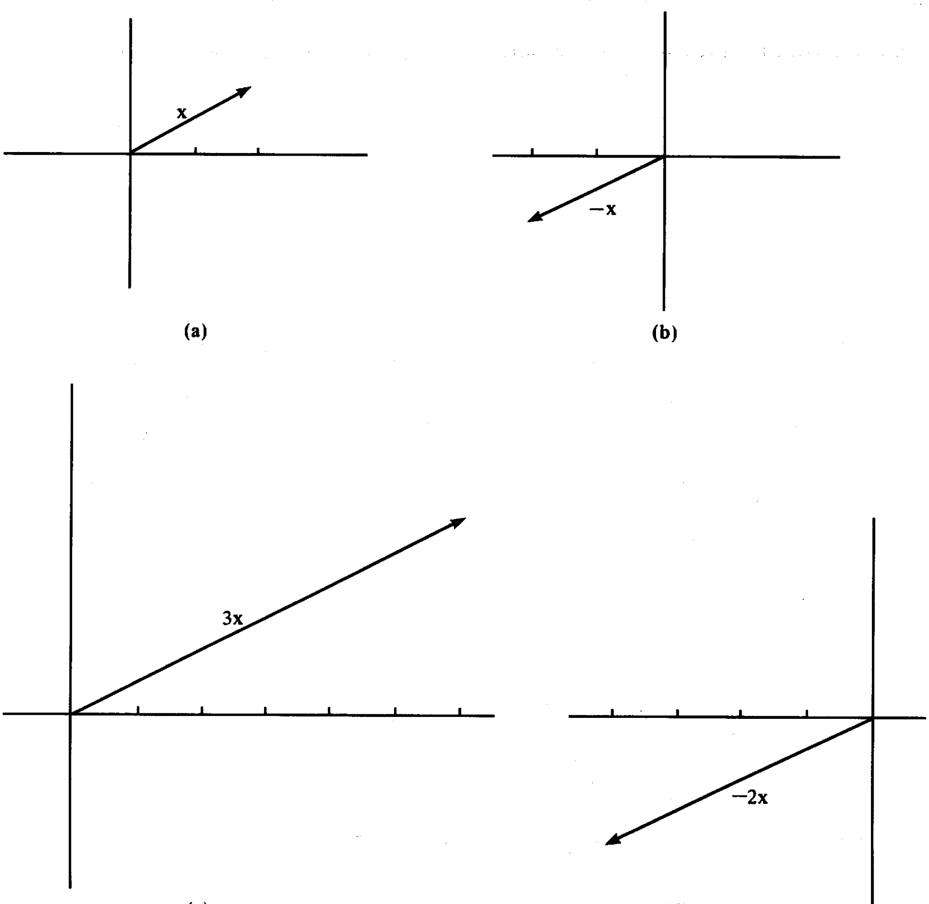


FIG. 3.1.5

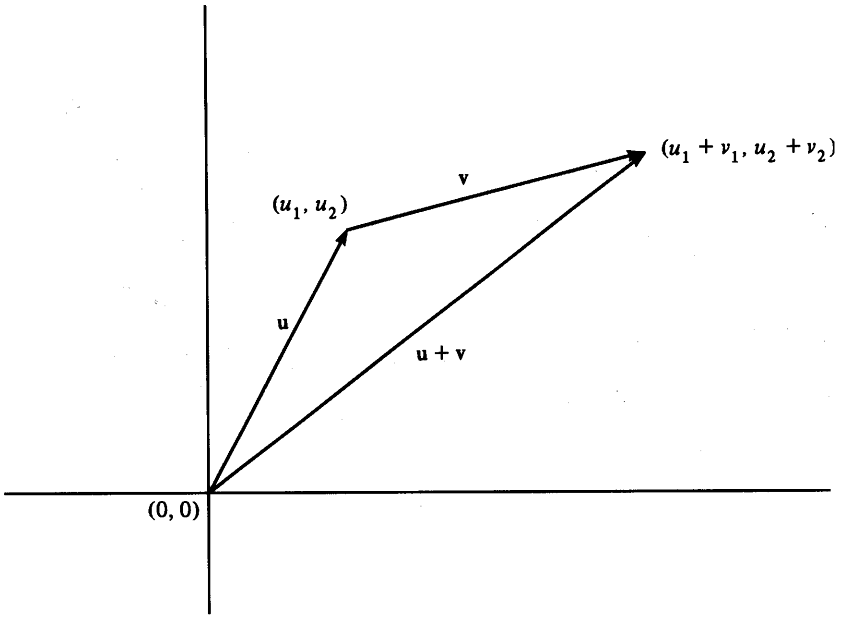


FIG. 3.1.6

dobro do comprimento de **x** e a mesma direção e o mesmo sentido que $-\mathbf{x}$. A soma de dois vetores $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ e $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ é definida por

$$\mathbf{u} + \mathbf{v} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix}$$

Observe que, se **v** é colocado no ponto final de **u**, então $\mathbf{u} + \mathbf{v}$ é representado pelo segmento de reta orientado com ponto inicial coincidindo com o ponto inicial de **u** e ponto final coincidindo com o ponto final de **v** (Fig. 3.1.6). Se **u** e **v** são colocados na origem e formamos um paralelogramo como na Fig. 1.3.7, as diagonais do paralelogramo representam a soma $\mathbf{u} + \mathbf{v}$ e a diferença $\mathbf{v} - \mathbf{u}$. De maneira análoga, vetores em R^3 podem ser representados por segmentos de reta orientados no espaço tridimensional (ver Fig. 3.1.8).

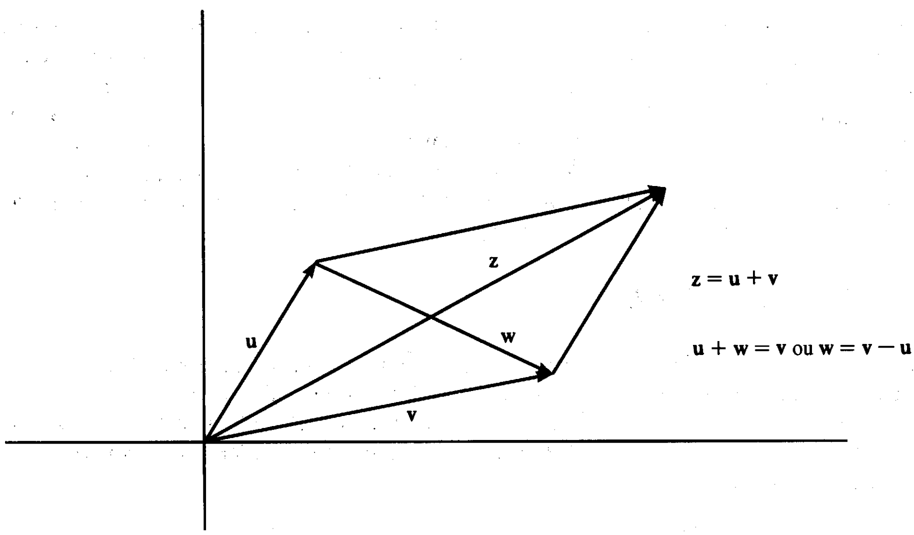
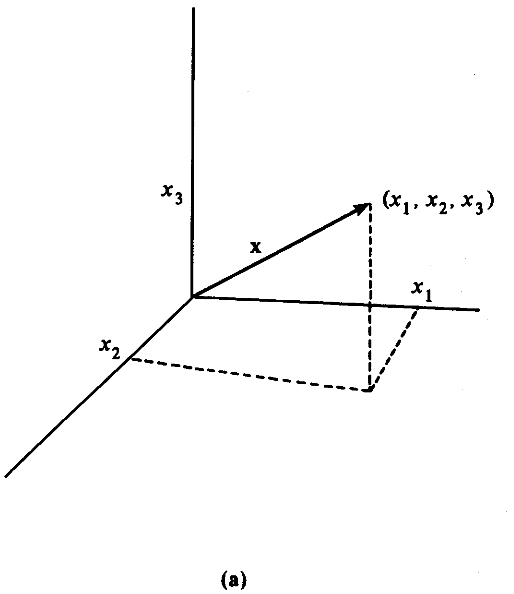


FIG. 3.1.7



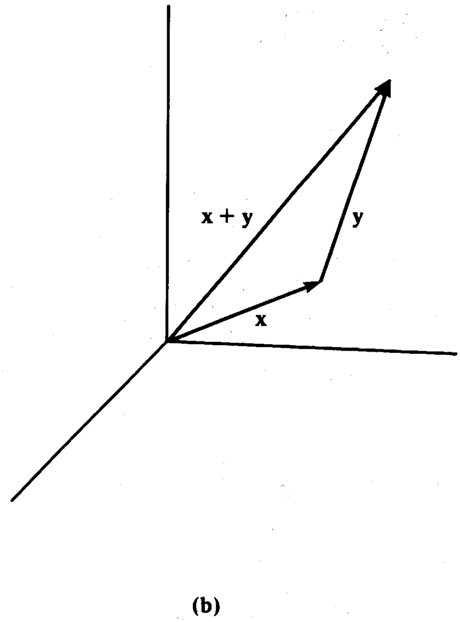


FIG. 3.1.8

Em geral, a multiplicação por um escalar e a soma em R^n são definidas por

$$\alpha\mathbf{x} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{pmatrix} \quad \text{e} \quad \mathbf{x}+\mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

quaisquer que sejam $\mathbf{x}, \mathbf{y} \in R^n$ e α um escalar.

O ESPAÇO VETORIAL $R^{m\times n}$

Podemos, também, olhar R^n como o conjunto de todas as matrizes $n \times 1$ com elementos reais. A soma e multiplicação por um escalar é, simplesmente, a soma e multiplicação por um escalar de matrizes. Mais geralmente, vamos denotar por $R^{m\times n}$ o conjunto de todas as matrizes $m \times n$ com elementos reais. Se $A = (a_{ij})$ e $B = (b_{ij})$, a soma $A + B$ é definida como a matriz $C = (c_{ij})$, onde $c_{ij} = a_{ij} + b_{ij}$. Dado um escalar α, podemos definir αA como a matriz $m \times n$ cujo elemento (i, j) é αa_{ij}. Desse modo, através da definição dessas operações no conjunto $R^{m\times n}$, criamos um sistema matemático. As operações de multiplicação por um escalar e soma em $R^{m\times n}$ obedecem a certas regras aritméticas. Essas regras formam os axiomas usados para definir o conceito de espaço vetorial.

AXIOMAS DO ESPAÇO VETORIAL

Definição. Seja V um conjunto onde estão definidas as operações de soma e multiplicação por um escalar. Isso significa que, a cada par de elementos $\mathbf{x}$ e $\mathbf{y}$ em V, podemos associar um único elemento $\mathbf{x} + \mathbf{y}$ em V, e a cada par $\mathbf{x}$ em V e α um escalar podemos associar um único elemento $\alpha\mathbf{x}$ em V. O conjunto V, junto com essas operações de soma e multiplicação por um escalar, forma um **espaço vetorial** se os seguintes axiomas são satisfeitos:

A1. $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ quaisquer que sejam $\mathbf{x}$ e $\mathbf{y}$ em V;

A2. $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$ quaisquer que sejam $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{z}$ em V;
A3. Existe um elemento $\mathbf{0}$ em V tal que $\mathbf{x} + \mathbf{0} = \mathbf{x}$ para todo $\mathbf{x}$ em V;
A4. Para cada $\mathbf{x}$ em V, existe um elemento $-\mathbf{x}$ em V tal que $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$;
A5. $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$ para todo número real α e para todos os $\mathbf{x}$ e $\mathbf{y}$ em V;
A6. $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$ para todos os números reais α e β e todos os $\mathbf{x}$ em V;
A7. $(\alpha\beta)\mathbf{x} = \alpha(\beta\mathbf{x})$ para todos os números reais α e β e para todo $\mathbf{x}$ em V;
A8. $1 \cdot \mathbf{x} = \mathbf{x}$ para todo $\mathbf{x}$ em V.

Os elementos de V são chamados de *vetores* e são, geralmente, denotados por letras do final do alfabeto: $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{z}$. Os números reais são chamados de *escalares*. O símbolo $\mathbf{0}$ foi usado para distinguir o vetor nulo do escalar 0. Em alguns contextos, os números complexos são usados como escalares. Neste livro, no entanto, escalares serão, em geral, números reais. Muitas vezes usa-se o termo *espaço vetorial real* para indicar que o conjunto dos escalares é o conjunto de números reais.

Um componente importante da definição é que o espaço é fechado em relação às duas operações. Essas propriedades podem ser resumidas da seguinte maneira:

C1. Se $\mathbf{x} \in V$ e se α é um escalar, então $\alpha\mathbf{x} \in V$;
C2. Se $\mathbf{x}, \mathbf{y} \in V$, então $\mathbf{x} + \mathbf{y} \in V$.

Para ver a importância dessas propriedades, considere o seguinte exemplo: seja

$$W = \{(a, 1) \mid a \text{ real}\}$$

com as operações de soma e multiplicação por um escalar definidas da maneira usual. Os elementos (3, 1) e (5, 1) estão em W, mas a soma

$$(3, 1) + (5, 1) = (8, 2)$$

não pertence a W. A operação + não é realmente uma operação em W, já que a propriedade C2 não é válida. Analogamente, a multiplicação por um escalar não está definida em W, pois a propriedade C1 não é válida. O conjunto W junto com essas operações *não* é um espaço vetorial.

Por outro lado, se temos um conjunto U no qual as operações de soma e multiplicação por um escalar estão definidas e satisfazem as propriedades C1 e C2, precisamos verificar se os oito axiomas são satisfeitos para determinar se U é ou não um espaço vetorial. Vamos deixar a cargo do leitor verificar que R^n e $R^{m\times n}$, com as operações usuais de soma e multiplicação por um escalar, são de fato espaços vetoriais reais. Existem diversos exemplos importantes de espaços vetoriais.

O ESPAÇO VETORIAL $C[a, b]$

Vamos denotar por $C[a, b]$ o conjunto de todas as funções definidas e contínuas em $[a, b]$ com valores reais. Nesse caso, nosso conjunto universal é o conjunto de todas as funções. Então, nossos vetores são funções em $C[a, b]$. A soma $f + g$ de duas funções em $C[a, b]$ é definida por

$$(f + g)(x) = f(x) + g(x)$$

para todo x em $[a, b]$. A nova função $f + g$ é um elemento de $C[a, b]$, já que a soma de duas funções contínuas é contínua. Se f é uma função em $C[a, b]$ e α é um número real, definimos αf por

$$(\alpha f)(x) = \alpha f(x)$$

para todo x em $[a, b]$. É claro que αf pertence a $C[a, b]$, já que uma constante vezes uma função contínua é contínua. Definimos, então, as operações de soma e multiplicação por um escalar em $C[a, b]$. Para mostrar o primeiro axioma, $f + g = g + f$, precisamos mostrar que

$$(f + g)(x) = (g + f)(x) \qquad \text{para todo } x \text{ em } [a, b]$$

Isso é válido porque

$$(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x)$$

para todo x em $[a, b]$. Deixamos a cargo do leitor a verificação dos axiomas de espaço vetorial restantes.

O ESPAÇO VETORIAL P_n

Vamos denotar por P_n o conjunto de todos os polinômios de grau menor do que n*. Defina $p + q$ e αp por

$$(p + q)(x) = p(x) + q(x)$$

e

$$(\alpha p)(x) = \alpha p(x)$$

para todo número real x. É fácil verificar que os axiomas de A1 a A8 são satisfeitos. Logo, P_n é um espaço vetorial em relação às operações usuais de soma e multiplicação por um escalar de funções.

PROPRIEDADES ADICIONAIS DE ESPAÇOS VETORIAIS

Vamos encerrar esta seção com um teorema que enuncia as três propriedades mais fundamentais de espaços vetoriais. Outras propriedades importantes são dadas nos Exercícios 7, 8 e 9.

Teorema 3.1.1. *Se V é um espaço vetorial e $\mathbf{x}$ é um elemento de V, então:*

(i) $0\mathbf{x} = \mathbf{0}$;
(ii) $\mathbf{x} + \mathbf{y} = \mathbf{0}$ *implica que* $\mathbf{y} = -\mathbf{x}$ (*isto é, a inversa aditiva de* $\mathbf{x}$ *é única*);
(iii) $(-1)\mathbf{x} = -\mathbf{x}$.

Demonstração. Dos axiomas A6 e A8, tem-se

$$\mathbf{x} = 1\mathbf{x} = (1 + 0)\mathbf{x} = 1\mathbf{x} + 0\mathbf{x} = \mathbf{x} + 0\mathbf{x}$$

Logo,

$$-\mathbf{x} + \mathbf{x} = -\mathbf{x} + (\mathbf{x} + 0\mathbf{x}) = (-\mathbf{x} + \mathbf{x}) + 0\mathbf{x} \qquad \text{(A2)}$$

$$\mathbf{0} = \mathbf{0} + 0\mathbf{x} = 0\mathbf{x} \qquad \text{(A1, A3 e A4)}$$

Para provar (ii), suponha que $\mathbf{x} + \mathbf{y} = \mathbf{0}$. Então

$$-\mathbf{x} = -\mathbf{x} + \mathbf{0} = -\mathbf{x} + (\mathbf{x} + \mathbf{y})$$

Portanto,

$$-\mathbf{x} = (-\mathbf{x} + \mathbf{x}) + \mathbf{y} = \mathbf{0} + \mathbf{y} = \mathbf{y} \qquad \text{(A1, A2, A3 e A4)}$$

Finalmente, para provar (iii), observe que

$$\mathbf{0} = 0\mathbf{x} = (1 + (-1))\mathbf{x} = 1\mathbf{x} + (-1)\mathbf{x} \qquad \text{[(i) e A6]}$$

Logo,

$$\mathbf{x} + (-1)\mathbf{x} = \mathbf{0} \qquad \text{(A8)}$$

e, de (ii), tem-se

$$(-1)\mathbf{x} = -\mathbf{x}$$

□

EXERCÍCIOS

1. Considere os vetores $\mathbf{x}_1 = (8, 6)^T$ e $\mathbf{x}_2 = (4, -1)^T$ em R^2.
   (a) Encontre o comprimento de cada vetor.
   (b) Seja $\mathbf{x}_3 = \mathbf{x}_1 + \mathbf{x}_2$. Determine o comprimento de $\mathbf{x}_3$. Qual a relação entre seu comprimento e a soma dos comprimentos de $\mathbf{x}_1$ e $\mathbf{x}_2$?

* Incluindo o polinômio identicamente nulo. (N.T.)

(c) Desenhe um gráfico ilustrando como $\mathbf{x}_3$ pode ser construído geometricamente usando $\mathbf{x}_1$ e $\mathbf{x}_2$. Use esse gráfico para dar uma interpretação geométrica da sua resposta em (b).

2. Repita o Exercício 1 para os vetores $\mathbf{x}_1 = (2, 1)^T$ e $\mathbf{x}_2 = (6, 3)^T$.

3. Seja C o conjunto dos números complexos. Defina a soma em C por

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

e defina a multiplicação por um escalar por

$$\alpha(a + bi) = \alpha a + \alpha bi$$

para todos os números reais α. Mostre que C é um espaço vetorial em relação a essas operações.

4. Mostre que $R^{m\times n}$, com as operações usuais de soma e multiplicação por um escalar, satisfaz os oito axiomas de espaços vetoriais.

5. Mostre que $C[a, b]$, com as operações usuais de soma e multiplicação por um escalar, satisfaz os oito axiomas de espaços vetoriais.

6. Seja P o conjunto de todos os polinômios. Mostre que P, com as operações usuais de soma e multiplicação por um escalar para funções, forma um espaço vetorial.

7. Mostre que o elemento $\mathbf{0}$ de um espaço vetorial é único.

8. Sejam $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{z}$ vetores em um espaço vetorial V. Mostre que, se

$$\mathbf{x} + \mathbf{y} = \mathbf{x} + \mathbf{z}$$

então $\mathbf{y} = \mathbf{z}$.

9. Seja V um espaço vetorial e seja $\mathbf{x} \in V$. Mostre que:
(a) $\beta\mathbf{0} = \mathbf{0}$ para todos os escalares β;
(b) se $\alpha\mathbf{x} = \mathbf{0}$, então $\alpha = 0$ ou $\mathbf{x} = \mathbf{0}$.

10. Seja S o conjunto de todos os pares ordenados de números reais. Defina a multiplicação por um escalar e a soma em S por

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

$$(x_1, x_2) \oplus (y_1, y_2) = (x_1 + y_1, 0)$$

Usamos o símbolo $\oplus$ para denotar a soma nesse sistema para evitar confusão com a soma usual $\mathbf{x} + \mathbf{y}$ de vetores linhas. Mostre que S, junto com a multiplicação usual por um escalar e a operação $\oplus$, não é um espaço vetorial. Quais dos oito axiomas não são válidos?

11. Seja V o conjunto de todos os pares ordenados de números reais com a soma definida por

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$$

e a multiplicação por um escalar definida por

$$\alpha \circ (x_1, x_2) = (\alpha x_1, x_2)$$

Como a multiplicação por um escalar é definida de maneira diferente da usual, usamos um símbolo diferente para evitar confusão com a multiplicação usual de um vetor linha por um escalar. V é um espaço vetorial em relação a essas operações? Justifique sua resposta.

12. Denote por R^+ o conjunto dos números reais positivos. Defina a operação de multiplicação por um escalar por

$$\alpha \circ x = x^\alpha$$

para cada $x \in R^+$ e para cada número real α. Defina a operação de soma por

$$x \oplus y = x \cdot y \quad \text{para todos} \quad x, y \in R^+$$

Então, para esse sistema, o produto do escalar -3 por ½ é dado por

$$-3 \circ \frac{1}{2} = \left(\frac{1}{2}\right)^{-3} = 8$$

e a soma de 2 com 5 é dada por

$$2 \oplus 5 = 2 \cdot 5 = 10$$

R^+ é um espaço vetorial em relação a essas operações? Justifique sua resposta.

13. Seja R o conjunto de todos os números reais. Defina a multiplicação por um escalar por

$$\alpha x = \alpha \cdot x \qquad \text{(a multiplicação usual de números reais)}$$

e a soma, denotada por $\oplus$, por

$$x \oplus y = \max(x, y) \qquad \text{(o máximo entre dois números)}$$

R é um espaço vetorial em relação a essas operações? Justifique sua resposta.

14. Denote por Z o conjunto de todos os números inteiros com a soma definida da maneira usual e a multiplicação por um escalar definida por

$$\alpha \circ k = [\![\alpha]\!] \cdot k \quad \text{para todos} \quad k \in Z$$

onde $[[\alpha]]$ denota o maior inteiro menor ou igual a α. Por exemplo,

$$2{,}25 \circ 4 = [\![2{,}25]\!] \cdot 4 = 2 \cdot 4 = 8$$

Mostre que Z não é um espaço vetorial em relação a essas operações. Quais dos axiomas não são válidos?

15. Denote por S o conjunto de todas as seqüências infinitas de números reais com a multiplicação por um escalar e a soma definidas por

$$\alpha\{a_n\} = \{\alpha a_n\}$$

$$\{a_n\} + \{b_n\} = \{a_n + b_n\}$$

Mostre que S é um espaço vetorial.

16. Podemos definir uma bijeção entre os elementos de P_n e de R^n por

$$p(x) = a_1 + a_2 x + \cdots + a_n x^{n-1} \leftrightarrow (a_1, \ldots, a_n)^T = \mathbf{a}$$

Mostre que, se $p \leftrightarrow \mathbf{a}$ e $q \leftrightarrow \mathbf{b}$, então

(a) $\alpha p \leftrightarrow \alpha\mathbf{a}$ qualquer que seja o escalar α;

(b) $p + q \leftrightarrow \mathbf{a} + \mathbf{b}$.

[Em geral, dois espaços vetoriais são ditos *isomorfos* se existe uma bijeção entre eles que preserva a multiplicação por um escalar e a soma como em (a) e (b).]

2 SUBESPAÇOS

Dado um espaço vetorial V, é muitas vezes possível formar um outro espaço vetorial usando um subconjunto S de V e as operações de V. Como V é um espaço vetorial, as operações de soma e multiplicação por um escalar sempre produzem um outro vetor em V. Para um novo sistema, usando um subconjunto S de V, ser um espaço vetorial, o conjunto S tem que ser fechado em relação às operações de soma e multiplicação por um escalar. Em outras palavras, a soma de dois elementos em S tem que ser sempre um elemento de S e a multiplicação de um elemento de S por um escalar tem que pertencer sempre à S.

EXEMPLO 1. Seja $S = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \middle| \; x_2 = 2x_1 \right\}$. S é um subconjunto de R^2. Se $\begin{pmatrix} c \\ 2c \end{pmatrix}$ é um elemento qualquer de S e α é um escalar arbitrário, então

$$\alpha \begin{pmatrix} c \\ 2c \end{pmatrix} = \begin{pmatrix} \alpha c \\ 2\alpha c \end{pmatrix}$$

é um elemento de S. Se $\begin{pmatrix} a \\ 2a \end{pmatrix}$ e $\begin{pmatrix} c \\ 2c \end{pmatrix}$ são dois elementos arbitrários de S, então sua soma

$$\begin{pmatrix} a + b \\ 2a + 2b \end{pmatrix} = \begin{pmatrix} a + b \\ 2(a + b) \end{pmatrix}$$

também pertence a S. É fácil ver que o sistema matemático consistindo no conjunto S (em vez de R^2), junto com as operações herdadas de R^2, é um espaço vetorial. □

Definição. Se S é um subconjunto não-vazio de um espaço vetorial V e se S satisfaz as seguintes condições:

(i) $\alpha\mathbf{x} \in S$ para todo escalar α, sempre que $\mathbf{x} \in S$
(ii) $\mathbf{x} + \mathbf{y} \in S$ sempre que $\mathbf{x} \in S$ e $\mathbf{y} \in S$

então dizemos que S é um **subespaço** de V.

A condição (i) diz que S é fechado sob a multiplicação por um escalar. Em outras palavras, toda vez que um elemento de S é multiplicado por um escalar, o resultado é um elemento de S. A condição (ii) diz que S é fechado em relação à soma. Em outras palavras, a soma de dois elementos de S é sempre um elemento de S. Então, se efetuarmos as operações usando as operações de V e os elementos de S, sempre obtemos elementos de S. Um subespaço de V, então, é um subconjunto S que é fechado em relação às operações de V.

Seja S um subespaço de um espaço vetorial V. Usando as operações de soma e multiplicação por um escalar como definidas em V, podemos formar um novo sistema matemático com S como conjunto universal. É fácil ver que os oito axiomas permanecem válidos para esse novo sistema. Os axiomas A3 e A4 seguem do Teorema 3.1.1 e da condição (i) na definição de subespaço. Os outros seis axiomas são válidos para todos os elementos de V, logo, em particular, são válidos para os elementos de S. Portanto, todo subespaço de um espaço vetorial é ele mesmo um espaço vetorial.

Observação. É fácil ver que os subconjuntos $\{\mathbf{0}\}$ e V de um espaço vetorial V são subespaços de V.* Todos os outros subespaços de V são chamados de *subespaços próprios*.

EXEMPLO 2. Seja $S = \{(x_1, x_2, x_3)^T \mid x_1 = x_2\}$. Então S é um subespaço de R^3, pois

(i) se $\mathbf{x} = (a, a, b)^T \in S$, então

$$\alpha\mathbf{x} = (\alpha a, \alpha a, \alpha b)^T \in S$$

(ii) se $(a, a, b)^T$ e $(c, c, d)^T$ são elementos arbitrários de S, então

$$(a, a, b)^T + (c, c, d)^T = (a + c, a + c, b + d)^T \in S$$

□

EXEMPLO 3. Seja $S = \left\{ \begin{pmatrix} x \\ 1 \end{pmatrix} \middle|\; x \text{ um número real} \right\}$ S não é um subespaço de R^2. Nesse caso as duas condições não são válidas. S não é fechado em relação à multiplicação por um escalar, pois $\alpha \begin{pmatrix} x \\ 1 \end{pmatrix} \notin S$ não pertence a S a menos que $\alpha = 1$. S não é fechado em relação à soma, pois

$$\begin{pmatrix} x \\ 1 \end{pmatrix} + \begin{pmatrix} y \\ 1 \end{pmatrix} = \begin{pmatrix} x + y \\ 2 \end{pmatrix} \notin S$$

□

EXEMPLO 4. Seja $S = \{A \in R^{2\times 2} \mid a_{12} = -a_{21}\}$. O conjunto S forma um subespaço de $R^{2\times 2}$, pois

(i) Se $A \in S$, então A tem que ser da forma

$$A = \begin{pmatrix} a & b \\ -b & c \end{pmatrix}$$

e, portanto,

$$\alpha A = \begin{pmatrix} \alpha a & \alpha b \\ -\alpha b & \alpha c \end{pmatrix}$$

* Esses subespaços são chamados de *subespaços triviais*. (N.T.)

Como o elemento (2, 1) de αA é menos o elemento (1, 2), $\alpha A \in S$.

(ii) Se $A, B \in S$, então as matrizes têm que ser da forma

$$A = \begin{pmatrix} a & b \\ -b & c \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} d & e \\ -e & f \end{pmatrix}$$

Então,

$$A + B = \begin{pmatrix} a+d & b+e \\ -(b+e) & c+f \end{pmatrix}$$

Portanto, $A + B \in S$. □

EXEMPLO 5. Seja S o conjunto de todos os polinômios de grau menor do que n com a propriedade de que $p(0) = 0$. O conjunto S não é vazio, já que contém o polinômio nulo. Vamos mostrar que S é um subespaço de P_n.

(i) Se $p(x) \in S$ e α é um escalar, então

$$\alpha p(0) = \alpha \cdot 0 = 0$$

e, portanto, $\alpha p \in S$.

(ii) Se $p(x)$ e $q(x)$ estão em S, então

$$(p+q)(0) = p(0) + q(0) = 0 + 0 = 0$$

e, portanto, $p + q \in S$. □

EXEMPLO 6. Seja $C^n[a, b]$ o conjunto de todas as funções f que têm a n-ésima derivada contínua em $[a, b]$. Deixamos a cargo do leitor verificar que $C^n[a, b]$ é um subespaço de $C[a, b]$. □

EXEMPLO 7. A função $f(x) = |x|$ está em $C[-1, 1]$, mas não é diferenciável em $x = 0$, logo não está em $C^1[-1, 1]$. Isso mostra que $C^1[-1, 1]$ é um subespaço próprio de $C[-1, 1]$. A função $g(x) = x|x|$ está em $C^1[-1, 1]$, já que é diferenciável em todos os pontos de $[-1, 1]$, e $g'(x) = 2|x|$ é contínua em $[-1, 1]$. No entanto, $g \notin C^2[-1, 1]$, já que $g''(x)$ não está definida em $x = 0$. Portanto, o espaço vetorial $C^2[-1, 1]$ é um subespaço próprio de $C[-1, 1]$ e de $C^1[-1, 1]$. □

EXEMPLO 8. Seja S o conjunto de todas as funções f em $C^2[a, b]$ tais que

$$f''(x) + f(x) = 0$$

para todo x em $[a, b]$. O conjunto S não é vazio, já que a função identicamente nula pertence a S. Se f está em S e α é um escalar, então, para todo x em $[a, b]$,

$$(\alpha f)''(x) + (\alpha f)(x) = \alpha f''(x) + \alpha f(x) = \alpha(f''(x) + f(x)) = \alpha \cdot 0 = 0$$

Logo $\alpha f \in S$. Se f e g pertencem a S, então

$$\begin{aligned}(f+g)''(x) + (f+g)(x) &= f''(x) + g''(x) + f(x) + g(x) \\ &= [f''(x) + f(x)] + [g''(x) + g(x)] = 0 + 0 = 0\end{aligned}$$

Portanto, o conjunto de todas as soluções em $[a, b]$ da equação diferencial $y'' + y = 0$ forma um subespaço de $C^2[a, b]$. Observe que ambas as funções $f(x) = \text{sen}\, x$ e $g(x) = \cos x$ estão em S, logo qualquer função da forma $c_1 \,\text{sen}\, x + c_2 \cos x$ também tem que estar em S. É fácil verificar que as funções dessa forma são soluções de $y'' + y = 0$. □

O NÚCLEO DE UMA MATRIZ

Seja A uma matriz $m \times n$. Vamos denotar por $N(A)$ o conjunto de todas as soluções do sistema homogêneo $A\mathbf{x} = \mathbf{0}$. Então

$$N(A) = \{\mathbf{x} \in R^n \mid A\mathbf{x} = \mathbf{0}\}$$

Vamos mostrar que $N(A)$ é um subespaço de R^n. Se $\mathbf{x} \in N(A)$ e α é um escalar, então

$$A(\alpha\mathbf{x}) = \alpha A\mathbf{x} = \alpha\mathbf{0} = \mathbf{0}$$

e, portanto, $\alpha\mathbf{x} \in N(A)$. Se $\mathbf{x}$ e $\mathbf{y}$ pertencem a $N(A)$, então

$$A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

Logo, $\mathbf{x} + \mathbf{y} \in N(A)$. Temos, então, que $N(A)$ é um subespaço de R^n, ou seja, o conjunto de todas as soluções do sistema homogêneo $A\mathbf{x} = \mathbf{0}$ forma um subespaço de R^n. Esse subespaço $N(A)$ é chamado de *núcleo* de A.

EXEMPLO 9. Determine $N(A)$ se

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{pmatrix}$$

SOLUÇÃO. Usando o método de Gauss-Jordan para resolver $A\mathbf{x} = \mathbf{0}$, obtemos

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -2 & 1 & 0 \end{array}\right)$$

$$\to \left(\begin{array}{cccc|c} 1 & 0 & -1 & 1 & 0 \\ 0 & -1 & -2 & 1 & 0 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & 2 & -1 & 0 \end{array}\right)$$

A forma escada reduzida por linhas tem duas variáveis livres, x_3 e x_4.

$$x_1 = x_3 - x_4$$

$$x_2 = -2x_3 + x_4$$

Logo, fazendo $x_3 = \alpha$ e $x_4 = \beta$, temos que

$$\mathbf{x} = \begin{pmatrix} \alpha - \beta \\ -2\alpha + \beta \\ \alpha \\ \beta \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

é uma solução de $A\mathbf{x} = \mathbf{0}$. O espaço vetorial $N(A)$ consiste em todos os vetores da forma

$$\alpha \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

onde α e β são escalares. □

ESPAÇO GERADO E CONJUNTOS GERADORES

Definição. Sejam $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ vetores em um espaço vetorial V. Uma soma da forma $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_n$, onde $\alpha_1, \ldots, \alpha_n$ são escalares, é chamada uma **combinação linear** de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. O conjunto de todas as combinações lineares de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ é o **espaço gerado** por $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Vamos denotar o espaço gerado por $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ por $[\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}]$.

Vimos, no Exemplo 9, que o núcleo de A era o espaço gerado pelos vetores $(1, -2, 1, 0)^T$ e $(-1, 1, 0, 1)^T$.

EXEMPLO 10. Em R^3, o espaço gerado por $\mathbf{e}_1$ e $\mathbf{e}_2$ é o conjunto de todos os vetores da forma

$$\alpha\mathbf{e}_1 + \beta\mathbf{e}_2 = \begin{pmatrix} \alpha \\ \beta \\ 0 \end{pmatrix}$$

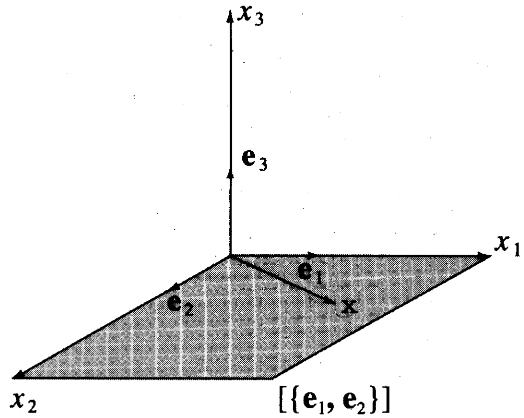


FIG. 3.2.1

O leitor pode verificar que $[\{\mathbf{e}_1, \mathbf{e}_2\}]$ é um subespaço de R^3. O espaço gerado pelos vetores $\mathbf{e}_1$, $\mathbf{e}_2$ e $\mathbf{e}_3$ é o conjunto de todos os vetores da forma

$$\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \alpha_3\mathbf{e}_3 = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$$

Logo, $[\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}] = R^3$. □

Teorema 3.2.1. *Se* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ *são elementos de um espaço vetorial* V, *então* $[\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}]$ *é um subespaço de V.*

Demonstração. Seja β um escalar e seja $\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_n$ um elemento arbitrário de $\{[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}]$. Como

$$\beta\mathbf{v} = (\beta\alpha_1)\mathbf{v}_1 + (\beta\alpha_2)\mathbf{v}_2 + \cdots + (\beta\alpha_n)\mathbf{v}_n$$

temos que $\beta\mathbf{v} \in [\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}]$. Temos que mostrar agora que qualquer soma de elementos em $[\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}]$ está em $[\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}]$. Sejam $\mathbf{v} = \alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n$ e $\mathbf{w} = \beta_1\mathbf{v}_1 + \ldots + \beta_n\mathbf{v}_n$.

$$\mathbf{v} + \mathbf{w} = (\alpha_1 + \beta_1)\mathbf{v}_1 + \cdots + (\alpha_n + \beta_n)\mathbf{v}_n \in [\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}]$$

Portanto, $[\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}]$ é um subespaço de V. □

Um vetor $\mathbf{x}$ em R^3 está em $[\{\mathbf{e}_1, \mathbf{e}_2\}]$ se e somente se ele pertence ao plano x_1x_2 no espaço tridimensional. Podemos, portanto, pensar no plano x_1x_2 como a representação geométrica do subespaço $[\{\mathbf{e}_1, \mathbf{e}_2\}]$ (ver Fig. 3.2.1). Analogamente, dados dois vetores $\mathbf{x}$ e $\mathbf{y}$, se $(0, 0, 0)$, (x_1, x_2, x_3) e (y_1, y_2, y_3) não são colineares, esses pontos determinam um plano. Se $\mathbf{z} = c_1\mathbf{x} + c_2\mathbf{y}$, então $\mathbf{z}$ é uma soma de vetores paralelos a $\mathbf{x}$ e $\mathbf{y}$ e pertence ao plano determinado pelos dois vetores (ver Fig. 3.2.2). Em geral, se dois vetores podem ser usados para determinar um plano no espaço tridimensional, o plano é a representação geométrica de $[\{\mathbf{x}, \mathbf{y}\}]$.

Sejam $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ vetores em um espaço vetorial V. Vamos nos referir a $[\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}]$ como o subespaço de V *gerado* por $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Pode acontecer que $[\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}] = V$ e, nesse caso, dizemos que os vetores $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *geram* V ou que $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ é um *conjunto gerador* para V. Temos, então, a seguinte definição.

Definição. O conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ é um **conjunto gerador** para V se e somente se todo vetor em V pode ser escrito como uma combinação linear de $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

EXEMPLO 11. Quais dos conjuntos a seguir são conjuntos geradores para R^3?

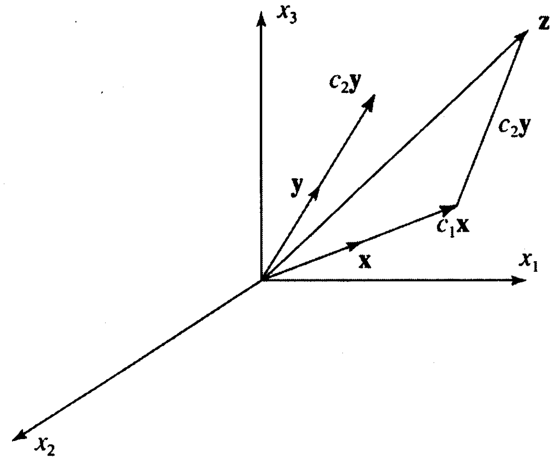


FIG. 3.2.2

(a) $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, (1, 2, 3)^T\}$
(b) $\{(1, 1, 1)^T, (1, 1, 0)^T, (1, 0, 0)^T\}$
(c) $\{(1, 0, 1)^T, (0, 1, 0)^T\}$
(d) $\{(1, 2, 4)^T, (2, 1, 3)^T, (4, -1, 1)^T\}$

SOLUÇÃO. Para determinar se um conjunto gera R^3, precisamos verificar se um vetor arbitrário $(a, b, c)^T$ em R^3 pode ser escrito como uma combinação linear dos vetores no conjunto. Para (a), é fácil ver que $(a, b, c)^T$ pode ser escrito como

$$(a, b, c)^T = a\mathbf{e}_1 + b\mathbf{e}_2 + c\mathbf{e}_3 + 0(1, 2, 3)^T$$

Para (b), precisamos verificar se é ou não possível encontrar constantes α_1, α_2, α_3 tais que

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Isso nos leva ao sistema de equações

$$\begin{aligned} \alpha_1 + \alpha_2 + \alpha_3 &= a \\ \alpha_1 + \alpha_2 &= b \\ \alpha_1 &= c \end{aligned}$$

Como a matriz de coeficientes é invertível, o sistema tem uma única solução. De fato, temos

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} c \\ b - c \\ a - b \end{pmatrix}$$

Logo,

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = c \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + (b - c) \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + (a - b) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

de modo que os três vetores geram o R^3.

Para (c), observe que qualquer combinação linear de $(1, 0, 1)^T$ e $(0, 1, 0)^T$ é um vetor da forma $(\alpha, \beta, \alpha)^T$. Logo, qualquer vetor $(a, b, c)^T$ em R^3 com $a \neq c$ não pertence ao espaço gerado por esses dois vetores.

O item (d) pode ser resolvido de maneira semelhante a (b). Se

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} + \alpha_2 \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} + \alpha_3 \begin{pmatrix} 4 \\ -1 \\ 1 \end{pmatrix}$$

então

$$\begin{aligned} \alpha_1 + 2\alpha_2 + 4\alpha_3 &= a \\ 2\alpha_1 + \alpha_2 - \alpha_3 &= b \\ 4\alpha_1 + 3\alpha_2 + \alpha_3 &= c \end{aligned}$$

Nesse caso, no entanto, a matriz de coeficientes é singular. O método de Gauss nos leva a um sistema da forma

$$\begin{aligned} \alpha_1 + 2\alpha_2 + 4\alpha_3 &= a \\ \alpha_2 + 3\alpha_3 &= \frac{2a - b}{3} \\ 0 &= 2a - 3c + 5b \end{aligned}$$

Se

$$2a - 3c + 5b \neq 0$$

então o sistema é incompatível. Portanto, para a maioria das escolhas para a, b, c, é impossível expressar $(a, b, c)^T$ como uma combinação linear de $(1, 2, 4)^T$, $(2, 1, 3)^T$, $(4, -1, 1)^T$. Os vetores não geram R^3. □

EXEMPLO 12. Os vetores $1 - x^2$, $x + 2$ e x^2 geram P_3. Então, se $ax^2 + bx + c$ é qualquer polinômio em P_3, é possível encontrar escalares α_1, α_2 e α_3 tais que

$$ax^2 + bx + c = \alpha_1(1 - x^2) + \alpha_2(x + 2) + \alpha_3 x^2$$

De fato,

$$\alpha_1(1 - x^2) + \alpha_2(x + 2) + \alpha_3 x^2 = (\alpha_3 - \alpha_1)x^2 + \alpha_2 x + (\alpha_1 + 2\alpha_2)$$

Fazendo

$$\begin{aligned} \alpha_3 - \alpha_1 &= a \\ \alpha_2 &= b \\ \alpha_1 + 2\alpha_2 &= c \end{aligned}$$

e resolvendo, obtemos $\alpha_1 = c - 2b$, $\alpha_2 = b$ e $\alpha_3 = a + c - 2b$. □

Vimos, no Exemplo 11(a), que os vetores $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, $(1, 2, 3)^T$ geram R^3. É claro que R^3 poderia ser gerado apenas pelos vetores $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. O vetor $(1, 2, 3)^T$ não é realmente necessário. Na próxima seção, vamos considerar o problema de encontrar conjuntos geradores mínimos para um espaço vetorial V (isto é, conjuntos geradores que contêm o menor número possível de vetores).

EXERCÍCIOS

1. Determine se cada conjunto a seguir é ou não um subespaço de R^2.
   (a) $\{(x_1, x_2)^T \mid x_1 + x_2 = 0\}$
   (b) $\{(x_1, x_2)^T \mid x_1 x_2 = 0\}$
   (c) $\{(x_1, x_2)^T \mid x_1 = 3x_2\}$
   (d) $\{(x_1, x_2)^T \mid x_1 = 3x_2 + 1\}$

2. Determine se cada conjunto a seguir é ou não um subespaço de R^3.
   (a) $\{(x_1, x_2, x_3)^T \mid x_1 + x_3 = 1\}$
   (b) $\{(x_1, x_2, x_3)^T \mid x_1 = x_2 = x_3\}$

(c) $\{(x_1, x_2, x_3)^T \mid x_3 = x_1 + x_2\}$

(d) $\{(x_1, x_2, x_3)^T \mid x_3 = x_1^2 + x_2^2\}$

3. Determine se cada conjunto a seguir é ou não um subespaço de $R^{2\times 2}$.

(a) O conjunto de todas as matrizes diagonais 2×2.

(b) O conjunto de todas as matrizes triangulares inferiores 2×2.

(c) O conjunto de todas as matrizes A 2×2 tais que $a_{12} = 1$.

(d) O conjunto de todas as matrizes B 2×2 tais que $b_{11} = 0$.

(e) O conjunto de todas as matrizes simétricas 2×2.

(f) O conjunto de todas as matrizes singulares 2×2.

4. Determine o núcleo de cada uma das matrizes a seguir.

(a) $\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 2 & -3 & -1 \\ -2 & -4 & 6 & 3 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 3 & -4 \\ 2 & -1 & -1 \\ -1 & -3 & 4 \end{pmatrix}$ (d) $\begin{pmatrix} 1 & 1 & -1 & 2 \\ 2 & 2 & -3 & 1 \\ -1 & -1 & 0 & -5 \end{pmatrix}$

5. Determine se cada conjunto a seguir é ou não um subespaço de P_4. (Cuidado!)

(a) O conjunto dos polinômios em P_4 de grau par.

(b) O conjunto dos polinômios de grau 3.

(c) O conjunto dos polinômios $p(x)$ em P_4 tais que $p(0) = 0$.

(d) O conjunto dos polinômios em P_4 que têm pelo menos uma raiz real.

6. Determine se cada conjunto a seguir é ou não um subespaço de $C[-1, 1]$.

(a) O conjunto das funções f em $C[-1, 1]$ tais $f(-1) = f(1)$.

(b) O conjunto das funções ímpares em $C[-1, 1]$.

(c) O conjunto das funções não-decrescentes em $[-1, 1]$.

(d) O conjunto das funções f em $C[-1, 1]$ tais $f(-1) = 0$ *e* $f(1) = 0$.

(e) O conjunto das funções f em $C[-1, 1]$ tais $f(-1) = 0$ *ou* $f(1) = 0$.

7. Mostre que $C^n[a, b]$ é um subespaço de $C[a, b]$.

8. Seja A um vetor particular em $R^{2\times 2}$. Determine se cada conjunto a seguir é ou não um subespaço de $R^{2\times 2}$.

(a) $S_1 = \{B \in R^{2\times 2} \mid AB = BA\}$

(b) $S_2 = \{B \in R^{2\times 2} \mid AB \neq BA\}$

(c) $S_3 = \{B \in R^{2\times 2} \mid BA = O\}$

9. Determine se cada conjunto a seguir é ou não um conjunto gerador para R^2.

(a) $\left\{\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix}\right\}$ (b) $\left\{\begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 6 \end{pmatrix}\right\}$ (c) $\left\{\begin{pmatrix} -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \end{pmatrix}\right\}$

(d) $\left\{\begin{pmatrix} -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 2 \\ -4 \end{pmatrix}\right\}$ (e) $\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}\right\}$

10. Quais dos conjuntos a seguir são conjuntos geradores para R^3? Justifique suas respostas.

(a) $\{(1, 0, 0)^T, (0, 1, 1)^T, (1, 0, 1)^T\}$ (b) $\{(1, 0, 0)^T, (0, 1, 1)^T, (1, 0, 1)^T, (1, 2, 3)^T\}$

(c) $\{(2, 1, -2)^T, (3, 2, -2)^T, (2, 2, 0)^T\}$ (d) $\{(2, 1, -2)^T, (-2, -1, 2)^T, (4, 2, -4)^T\}$

(e) $\{(1, 1, 3)^T, (0, 2, 1)^T\}$

11. Sejam

$$\mathbf{x}_1 = \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 2 \\ 6 \\ 6 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} -9 \\ -2 \\ 5 \end{pmatrix}$$

(a) $\mathbf{x} \in [\{\mathbf{x}_1, \mathbf{x}_2\}]$?
(b) $\mathbf{y} \in [\{\mathbf{x}_1, \mathbf{x}_2\}]$?
Justifique suas respostas.

12. Quais dos conjuntos a seguir são conjuntos geradores para P_3? Justifique suas respostas.
(a) $\{1, x^2, x^2 - 2\}$ (b) $\{2, x^2, x, 2x + 3\}$
(c) $\{x + 2, x + 1, x^2 - 1\}$ (d) $\{x + 2, x^2 - 1\}$

13. Em $R^{2\times 2}$, sejam

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Mostre que E_{11}, E_{12}, E_{21}, E_{22} geram $R^{2\times 2}$.

14. Seja S o espaço vetorial das seqüências infinitas definido no Exercício 15 da Seção 1. Seja S_0 o conjunto das seqüências $\{a_n\}$ tais que $a_n \to 0$ quando $n \to \infty$. Mostre que S_0 é um subespaço de S.

15. Prove que, se S é um subespaço de R^1, então $S = \{\mathbf{0}\}$ ou $S = R^1$.

16. Seja A uma matriz $n \times n$. Prove que as seguintes afirmações são equivalentes:
(a) $N(A) = \{\mathbf{0}\}$;
(b) A é invertível;
(c) para cada $\mathbf{b} \in R^n$, o sistema $A\mathbf{x} = \mathbf{b}$ tem uma única solução.

17. Sejam U e V subespaços de um espaço vetorial W. Prove que $U \cap V$ também é um subespaço de W.

18. Seja S o subespaço de R^2 gerado por $\mathbf{e}_1$ e seja T o subespaço de R^2 gerado por $\mathbf{e}_2$. $S \cup T$ é um subespaço de R^2? Explique.

19. Sejam U e V subespaços de um espaço vetorial W. Defina

$$U + V = \{\mathbf{z} \mid \mathbf{z} = \mathbf{u} + \mathbf{v} \text{ onde } \mathbf{u} \in U \text{ e } \mathbf{v} \in V\}$$

Mostre que $U + V$ é um subespaço de W.

3 INDEPENDÊNCIA LINEAR

Nesta seção, vamos olhar mais de perto a estrutura de um espaço vetorial. Para começar, vamos nos restringir a espaços vetoriais que podem ser gerados por um número finito de elementos. Cada vetor no espaço pode ser construído a partir dos elementos nesse conjunto gerador, usando-se apenas as operações de soma e multiplicação por um escalar. É desejável encontrar um conjunto gerador "mínimo". Por mínimo, queremos dizer um conjunto gerador sem elementos desnecessários (isto é, todos os elementos no conjunto são necessários para se gerar o espaço vetorial). Para ver como encontrar um conjunto gerador mínimo, é preciso considerar como os vetores no conjunto "dependem" um do outro. Vamos, então, introduzir os conceitos de *dependência linear* e *independência linear*. Esses conceitos simples vão nos dar a chave para entender a estrutura de espaços vetoriais.

Vamos considerar os seguintes vetores em R^3:

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} -1 \\ 3 \\ 8 \end{pmatrix}$$

Seja S o subespaço de R^3 gerado por $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$. Observe que S pode ser representado, de fato, pelos vetores $\mathbf{x}_1$ e $\mathbf{x}_2$, já que $\mathbf{x}_3$ pertence ao espaço gerado por $\mathbf{x}_1$ e $\mathbf{x}_2$.

$$\mathbf{x}_3 = 3\mathbf{x}_1 + 2\mathbf{x}_2 \tag{1}$$

Qualquer combinação linear de $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ pode ser reduzida a uma combinação linear de $\mathbf{x}_1$ e $\mathbf{x}_2$:

$$\begin{aligned} \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \alpha_3\mathbf{x}_3 &= \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \alpha_3(3\mathbf{x}_1 + 2\mathbf{x}_2) \\ &= (\alpha_1 + 3\alpha_3)\mathbf{x}_1 + (\alpha_2 + 2\alpha_3)\mathbf{x}_2 \end{aligned}$$

Logo,

$$S = [\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}] = [\{\mathbf{x}_1, \mathbf{x}_2\}].$$

A Equação (1) pode ser colocada na forma

(2) $$3\mathbf{x}_1 + 2\mathbf{x}_2 - 1\mathbf{x}_3 = \mathbf{0}$$

Como os três coeficientes em (2) são diferentes de zero, podemos resolver para qualquer um dos vetores em função dos outros dois.

$$\mathbf{x}_1 = -\tfrac{2}{3}\mathbf{x}_2 + \tfrac{1}{3}\mathbf{x}_3, \quad \mathbf{x}_2 = -\tfrac{3}{2}\mathbf{x}_1 + \tfrac{1}{2}\mathbf{x}_3, \quad \mathbf{x}_3 = 3\mathbf{x}_1 + 2\mathbf{x}_2$$

Temos, então, que

$$[\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}] = [\{\mathbf{x}_2, \mathbf{x}_3\}] = [\{\mathbf{x}_1, \mathbf{x}_3\}] = [\{\mathbf{x}_1, \mathbf{x}_2\}].$$

Por causa da relação de dependência (2), o subespaço S pode ser gerado por qualquer dois dos vetores dados.

Por outro lado, não existe nenhuma relação de dependência entre $\mathbf{x}_1$ e $\mathbf{x}_2$. De fato, se c_1 e c_2 forem escalares tais que um deles é diferente de 0, e se

(3) $$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 = \mathbf{0}$$

então poderíamos resolver para um vetor em função do outro.

$$\mathbf{x}_1 = -\frac{c_2}{c_1}\mathbf{x}_2 \quad (c_1 \neq 0) \qquad \text{ou} \qquad \mathbf{x}_2 = -\frac{c_1}{c_2}\mathbf{x}_1 \quad (c_2 \neq 0)$$

No entanto, nenhum dos dois vetores em pauta é múltiplo do outro. Logo, $[\{\mathbf{x}_1\}]$ e $[\{\mathbf{x}_2\}]$ são subespaços próprios de $[\{\mathbf{x}_1, \mathbf{x}_2\}]$ e (3) só é válida se $c_1 = c_2 = 0$.

Podemos generalizar esse exemplo fazendo as seguintes observações.

(i) Se $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ gera um espaço vetorial V e um desses vetores pode ser escrito como uma combinação linear dos outros $n - 1$ vetores, então esses outros $n - 1$ vetores geram V.

(ii) Dados n vetores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, é possível escrever um dos vetores como uma combinação linear dos outros $n - 1$ vetores se e somente se existem escalares $c_1, c_2, \ldots, c_n$, nem todos nulos, tais que

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

Demonstração de (i). Suponha que $\mathbf{v}_n$ pode ser escrito como uma combinação linear de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}$.

$$\mathbf{v}_n = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \cdots + \beta_{n-1}\mathbf{v}_{n-1}$$

Seja $\mathbf{v}$ um elemento qualquer de V. Como $\mathbf{v}_1, \ldots, \mathbf{v}_n$ geram V, podemos escrever

$$\begin{aligned}\mathbf{v} &= \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_{n-1}\mathbf{v}_{n-1} + \alpha_n\mathbf{v}_n \\ &= \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_{n-1}\mathbf{v}_{n-1} + \alpha_n(\beta_1\mathbf{v}_1 + \cdots + \beta_{n-1}\mathbf{v}_{n-1}) \\ &= (\alpha_1 + \alpha_n\beta_1)\mathbf{v}_1 + (\alpha_2 + \alpha_n\beta_2)\mathbf{v}_2 + \cdots + (\alpha_{n-1} + \alpha_n\beta_{n-1})\mathbf{v}_{n-1}\end{aligned}$$

Logo, qualquer vetor $\mathbf{v}$ em V pode ser escrito como combinação linear de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}$ e, portanto, esses vetores geram V. □

Demonstração de (ii). Suponha que um dos vetores, por exemplo $\mathbf{v}_n$, pode ser escrito como uma combinação linear dos outros.

$$\mathbf{v}_n = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_{n-1}\mathbf{v}_{n-1}$$

Escolhendo $c_i = \alpha_i$ para $i = 1, \ldots, n - 1$ e $c_n = -1$, temos que

$$\sum_{i=1}^{n} c_i\mathbf{v}_i = \sum_{i=1}^{n-1} \alpha_i\mathbf{v}_i - \sum_{i=1}^{n-1} \alpha_i\mathbf{v}_i = \mathbf{0}$$

Por outro lado, se

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

e pelo menos um dos c_i, por exemplo c_n, é diferente de zero, então

$$\mathbf{v}_n = \frac{-c_1}{c_n}\mathbf{v}_1 + \frac{-c_2}{c_n}\mathbf{v}_2 + \cdots + \frac{-c_{n-1}}{c_n}\mathbf{v}_{n-1}$$

□

Definição. Os vetores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ em um espaço vetorial V são ditos **linearmente independentes** se

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

implica que todos os escalares $c_1, \ldots, c_n$ têm que ser iguais a 0.

Como conseqüência de (i) e (ii), vê-se que, se $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ é um conjunto gerador mínimo, então $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são linearmente independentes. Em contrapartida, se $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são linearmente independentes e geram V, então $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ é um conjunto gerador mínimo para V (ver Exercício 16). Um conjunto gerador mínimo é chamado de *base*. O conceito de base será estudado com mais detalhes na próxima seção.

EXEMPLO 1. Os vetores $\begin{pmatrix}1\\1\end{pmatrix}$ e $\begin{pmatrix}1\\2\end{pmatrix}$ são linearmente independentes, pois, se

$$c_1\begin{pmatrix}1\\1\end{pmatrix} + c_2\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

então

$$\begin{aligned} c_1 + c_2 &= 0 \\ c_1 + 2c_2 &= 0 \end{aligned}$$

e a única solução desse sistema é $c_1 = 0$, $c_2 = 0$. □

Definição. Os vetores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ em um espaço vetorial V são ditos **linearmente dependentes** se existem escalares $c_1, c_2, \ldots, c_n$, nem todos nulos, tais que

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

EXEMPLO 2. Seja $\mathbf{x} = (1, 2, 3)^T$. Os vetores $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{x}$ são linearmente dependentes, já que

$$\mathbf{e}_1 + 2\mathbf{e}_2 + 3\mathbf{e}_3 - \mathbf{x} = \mathbf{0}$$

(Nesse caso, $c_1 = 1, c_2 = 2, c_3 = 3, c_4 = -1$.) □

Dado um conjunto de vetores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ em um espaço vetorial V, é trivial encontrar escalares c_1, c_2, c_n tais que

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

Basta definir

$$c_1 = c_2 = \cdots = c_n = 0$$

Se existem escolhas não-triviais de escalares para os quais a combinação linear $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n$ é igual ao vetor nulo, então $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são linearmente dependentes. Se a única maneira da combinação linear $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n$ ser igual ao vetor nulo é quando todos os escalares $c_1, c_2, \ldots, c_n$ são iguais a 0, então $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são linearmente independentes.

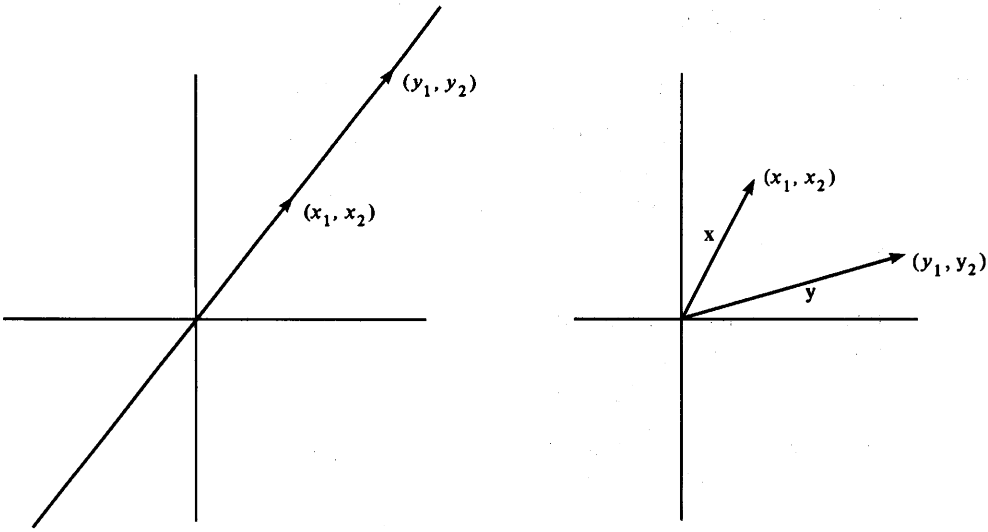
INTERPRETAÇÃO GEOMÉTRICA

Se $\mathbf{x}$ e $\mathbf{y}$ são linearmente dependentes em R^2, então

$$c_1\mathbf{x} + c_2\mathbf{y} = \mathbf{0}$$

onde c_1 e c_2 não são ambos iguais a 0. Se, por exemplo, $c_1 \neq 0$, temos

$$\mathbf{x} = -\frac{c_2}{c_1}\mathbf{y}$$



(a) **x** e **y** são linearmente dependentes.

(b) **x** e **y** são linearmente independentes.

FIG. 3.3.1

Se dois vetores em R^2 são linearmente dependentes, um deles pode ser escrito como um múltiplo escalar do outro. Logo, se os dois vetores forem colocados na origem, eles vão estar contidos na mesma reta (ver Fig. 3.3.1).

Se

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \text{e} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

são linearmente independentes em R^3, então os pontos (x_1, x_2, x_3) e (y_1, y_2, y_3) não pertencem a uma mesma reta contendo a origem no espaço tridimensional. Como $(0, 0, 0)$, (x_1, x_2, x_3) e (y_1, y_2, y_3) não são colineares, eles determinam um plano. Se (z_1, z_2, z_3) pertence a esse plano, o vetor $\mathbf{z} = (z_1, z_2, z_3)^T$ pode ser escrito como uma combinação linear de **x** e **y** e, portanto, **x**, **y** e **z** são linearmente dependentes. Se (z_1, z_2, z_3) não pertence a esse plano, os três vetores vão ser linearmente independentes (ver Fig. 3.3.2).

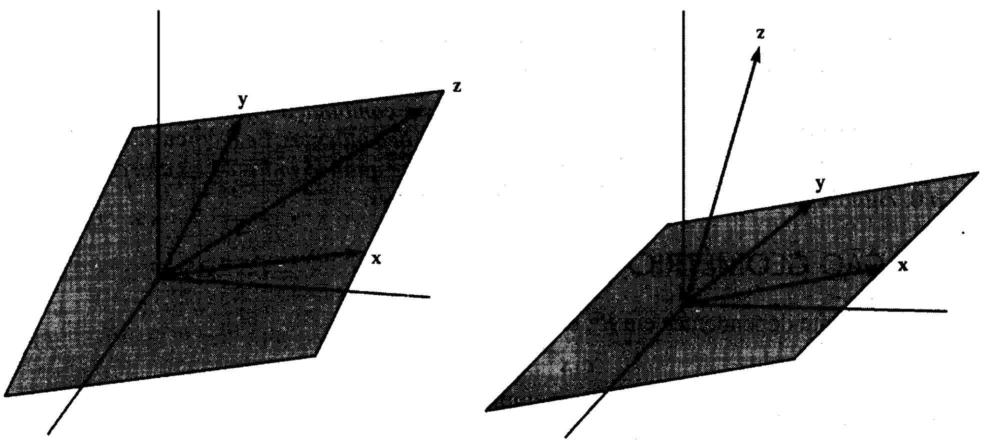


FIG. 3.3.2

TEOREMAS E EXEMPLOS

EXEMPLO 3. Quais dos conjuntos de vetores a seguir são linearmente independentes em R^3?

(a) $(1, 1, 1)^T, (1, 1, 0)^T, (1, 0, 0)^T$
(b) $(1, 0, 1)^T, (0, 1, 0)^T$
(c) $(1, 2, 4)^T, (2, 1, 3)^T, (4, -1, 1)^T$

SOLUÇÃO

(a) Esses três vetores são linearmente independentes. Para verificar isso, precisamos mostrar que a única maneira de se obter

(4) $$c_1(1, 1, 1)^T + c_2(1, 1, 0)^T + c_3(1, 0, 0)^T = (0, 0, 0)^T$$

é quando os escalares c_1, c_2, c_3 forem todos nulos. A Equação (4) pode ser escrita como um sistema linear com incógnitas c_1, c_2, c_3.

$$\begin{aligned} c_1 + c_2 + c_3 &= 0 \\ c_1 + c_2 \phantom{{}+c_3} &= 0 \\ c_1 \phantom{{}+c_2+c_3} &= 0 \end{aligned}$$

A única solução desse sistema é $c_1 = 0, c_2 = 0, c_3 = 0$.

(b) Se

$$c_1(1, 0, 1)^T + c_2(0, 1, 0)^T = (0, 0, 0)^T$$

então

$$(c_1, c_2, c_1)^T = (0, 0, 0)^T$$

logo $c_1 = c_2 = 0$. Portanto, os dois vetores são linearmente independentes.

(c) Se

$$c_1(1, 2, 4)^T + c_2(2, 1, 3)^T + c_3(4, -1, 1)^T = (0, 0, 0)^T$$

então

$$\begin{aligned} c_1 + 2c_2 + 4c_3 &= 0 \\ 2c_1 + c_2 - c_3 &= 0 \\ 4c_1 + 3c_2 + c_3 &= 0 \end{aligned}$$

A matriz de coeficientes desse sistema é singular, logo o sistema tem soluções não-triviais e, portanto, os vetores são linearmente dependentes. □

Observe que, nos itens (a) e (c) do Exemplo 3, foi necessário resolver um sistema 3×3 para determinar se os três vetores eram ou não linearmente independentes. No item (a), onde a matriz de coeficientes era invertível, os vetores eram linearmente independentes, enquanto no item (c), onde a matriz de coeficientes era singular, os vetores eram linearmente dependentes. Isso ilustra um caso particular do seguinte teorema:

Teorema 3.3.1. *Sejam $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ n vetores em R^n, com $\mathbf{x}_i = (x_{1i}, x_{2i}, \ldots, x_{ni})^T$ para $i = 1, \ldots, n$. Se $X = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$, então os vetores $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ são linearmente dependentes se e somente se X é singular.*

Demonstração. A equação

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n = \mathbf{0}$$

é equivalente ao sistema de equações

$$\begin{aligned} c_1x_{11} + c_2x_{12} + \cdots + c_nx_{1n} &= 0 \\ c_1x_{21} + c_2x_{22} + \cdots + c_nx_{2n} &= 0 \\ &\vdots \\ c_1x_{n1} + c_2x_{n2} + \cdots + c_nx_{nn} &= 0 \end{aligned}$$

Definindo $\mathbf{c} = (c_1, c_2, \ldots, c_n)^T$, o sistema pode ser escrito em forma matricial como

$$X\mathbf{c} = \mathbf{0}$$

Essa equação tem uma solução não-trivial se e somente se X é singular. Portanto, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ são linearmente dependentes se e somente se X é singular. □

Podemos usar o Teorema 3.3.1 para testar se n vetores são linearmente independentes em R^n. Basta formar a matriz X, cujas colunas são os vetores a serem testados. Para determinar se X é ou não singular, basta calcular o valor do determinante de X. Se $\det(X) = 0$, os vetores são linearmente dependentes. Se $\det(X) \neq 0$, os vetores são linearmente independentes.

EXEMPLO 4. Determine se os vetores $(4, 2, 3)^T$, $(2, 3, 1)^T$ e $(2, -5, 3)^T$ são ou não linearmente dependentes.

SOLUÇÃO. Como

$$\begin{vmatrix} 4 & 2 & 2 \\ 2 & 3 & -5 \\ 3 & 1 & 3 \end{vmatrix} = 0$$

os vetores são linearmente dependentes. □

Vamos considerar, agora, uma propriedade muito importante de vetores linearmente independentes. Combinações lineares de vetores linearmente independentes são únicas. O próximo teorema enuncia esse resultado de modo mais preciso.

Teorema 3.3.2. *Sejam $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ vetores em um espaço vetorial V. Um vetor $\mathbf{v} \in [\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}]$ pode ser escrito de maneira única como um combinação linear de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ se e somente se $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são linearmente independentes.*

Demonstração. Se $\mathbf{v} \in [\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}]$, então $\mathbf{v}$ pode ser escrito como uma combinação linear

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n \tag{5}$$

Suponha que $\mathbf{v}$ também possa ser expresso como uma combinação linear

$$\mathbf{v} = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \cdots + \beta_n\mathbf{v}_n \tag{6}$$

Vamos mostrar que, se $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são linearmente independentes, então $\beta_i = \alpha_i$, $i = 1, \ldots, n$, e que, se $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são linearmente dependentes, então é possível escolher os β_i diferentes dos α_i.

Se $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são linearmente independentes, subtraindo (6) de (5), obtemos

$$(\alpha_1 - \beta_1)\mathbf{v}_1 + (\alpha_2 - \beta_2)\mathbf{v}_2 + \cdots + (\alpha_n - \beta_n)\mathbf{v}_n = \mathbf{0} \tag{7}$$

Pela independência linear de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, os coeficientes de (7) têm que ser todos iguais a 0. Logo,

$$\alpha_1 = \beta_1,\ \alpha_2 = \beta_2,\ \ldots,\ \alpha_n = \beta_n$$

Portanto, a representação (5) é única quando $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são linearmente independentes.

Por outro lado, se $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são linearmente dependentes, existem $c_1, c_2, \ldots, c_n$, nem todos nulos, tais que

$$\mathbf{0} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n \tag{8}$$

Escolhendo, então,

$$\beta_1 = \alpha_1 + c_1,\ \beta_2 = \alpha_2 + c_2,\ \ldots,\ \beta_n = \alpha_n + c_n$$

e depois somando (5) e (8), obtemos

$$\begin{aligned}\mathbf{v} &= (\alpha_1 + c_1)\mathbf{v}_1 + (\alpha_2 + c_2)\mathbf{v}_2 + \cdots + (\alpha_n + c_n)\mathbf{v}_n \\ &= \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \cdots + \beta_n\mathbf{v}_n\end{aligned}$$

Como os c_i não são todos nulos, $\beta_i \neq \alpha_i$ para pelo menos um valor de i. Portanto, se $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são linearmente dependentes, a representação de um vetor como uma combinação linear de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ não é única. □

ESPAÇOS VETORIAIS DE FUNÇÕES

Para determinar se um conjunto de vetores é ou não linearmente independente em R^n, precisamos resolver um sistema homogêneo de equações lineares. Uma situação semelhante ocorre com o espaço vetorial P_n.

O Espaço Vetorial P_n. Para testar se os polinômios $p_1, p_2, \ldots, p_k$ são ou não linearmente independentes em P_n, fazemos

(9)
$$c_1p_1 + c_2p_2 + \cdots + c_kp_k = z$$

onde z representa o polinômio nulo,

$$z(x) = 0x^{n-1} + 0x^{n-2} + \cdots + 0x + 0$$

Se o polinômio do lado esquerdo do sinal de igualdade em (9) é colocado na forma $a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n$, então, como dois polinômios são iguais se e somente se todos os seus coeficientes são iguais, os coeficientes a_i têm que ser todos nulos. Como cada um dos a_i é uma combinação linear dos c_j, isso nos leva a um sistema linear homogêneo com incógnitas $c_1, c_2, \ldots, c_k$. Se o sistema tem apenas a solução trivial, os polinômios são linearmente independentes; caso contrário, eles são linearmente dependentes.

EXEMPLO 5. Para testar se os vetores

$$p_1(x) = x^2 - 2x + 3 \quad p_2(x) = 2x^2 + x + 8 \quad p_3(x) = x^2 + 8x + 7$$

são linearmente independentes, suponha que

$$c_1p_1(x) + c_2p_2(x) + c_3p_3(x) = 0x^2 + 0x + 0$$

Agrupando os termos de mesma potência, obtemos

$$(c_1+2c_2+c_3)x^2+(-2c_1+c_2+8c_3)x+(3c_1+8c_2+7c_3) = 0x^2+0x+0$$

Igualando os coeficientes, chegamos ao sistema

$$\begin{aligned} c_1 + 2c_2 + c_3 &= 0 \\ -2c_1 + c_2 + 8c_3 &= 0 \\ 3c_1 + 8c_2 + 7c_3 &= 0 \end{aligned}$$

A matriz de coeficientes desse sistema é singular e então existem soluções não-triviais. Portanto, p_1, p_2, p_3 são linearmente dependentes. □

O Espaço Vetorial $C^{(n-1)}[a, b]$. No Exemplo 4, usamos um determinante para testar se três vetores eram ou não linearmente independentes em R^3. Determinantes também podem ser usados para se decidir se um conjunto de n vetores é linearmente independente em $C^{(n-1)}[a, b]$. De fato, sejam $f_1, f_2, \ldots, f_n$

elementos de $C^{(n-1)}[a, b]$. Se esses vetores forem linearmente dependentes, existem escalares $c_1, c_2, \ldots, c_n$, nem todos nulos, tais que

$$c_1 f_1(x) + c_2 f_2(x) + \cdots + c_n f_n(x) = 0 \tag{10}$$

para todo x em $[a, b]$. Derivando em relação a x os dois lados de (10), obtemos

$$c_1 f_1'(x) + c_2 f_2'(x) + \cdots + c_n f_n'(x) = 0$$

Se continuarmos a derivar ambos os lados, terminamos com o sistema

$$\begin{array}{ccccccc} c_1 f_1(x) & + & c_2 f_2(x) & + \cdots + & c_n f_n(x) & = & 0 \\ c_1 f_1'(x) & + & c_2 f_2'(x) & + \cdots + & c_n f_n'(x) & = & 0 \\ \vdots & & & & & & \\ c_1 f_1^{(n-1)}(x) & + & c_2 f_2^{(n-1)}(x) & + \cdots + & c_n f_n^{(n-1)}(x) & = & 0 \end{array}$$

Para cada x em $[a, b]$ fixo, a equação matricial

$$\begin{pmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & & & \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{11}$$

tem a mesma solução não-trivial $(c_1, c_2, \ldots, c_n)^T$. Logo, se $f_1, f_2, \ldots, f_n$ são linearmente dependentes em $C^{(n-1)}[a, b]$, para cada x fixo em $[a, b]$, a matriz de coeficientes do sistema (11) é singular. Se a matriz é singular, seu determinante é zero.

Definição. Sejam $f_1, f_2, \ldots, f_n$ funções em $C^{(n-1)}[a, b]$ e defina a função $W[f_1, f_2, \ldots, f_n](x)$ em $[a, b]$ por

$$W[f_1, f_2, \ldots, f_n](x) = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & & & \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}$$

A função $W[f_1, f_2, \ldots, f_n]$ é chamada o **wronskiano** de $f_1, f_2, \ldots, f_n$.

Teorema 3.3.3. *Sejam $f_1, f_2, \ldots, f_n$ elementos de $C^{(n-1)}[a, b]$. Se existe um ponto x_0 em $[a, b]$ tal que $W[f_1, f_2, \ldots, f_n](x_0) \neq 0$, então $f_1, f_2, \ldots, f_n$ são linearmente independentes.*

Demonstração. Se $f_1, f_2, \ldots, f_n$ fossem linearmente dependentes, pela discussão anterior, a matriz de coeficientes de (11) seria singular para todo x em $[a, b]$ e, portanto, $W[f_1, f_2, \ldots, f_n](x)$ seria identicamente nula em $[a, b]$. □

Se $f_1, f_2, \ldots, f_n$ são linearmente independentes em $C^{(n-1)}[a, b]$, eles também são linearmente independentes em $C[a, b]$.

EXEMPLO 6. Mostre que e^x e e^{-x} são linearmente independentes em $C(-\infty, \infty)$.

SOLUÇÃO

$$W[e^x, e^{-x}] = \begin{vmatrix} e^x & e^{-x} \\ e^x & -e^{-x} \end{vmatrix} = -2$$

Como $W[e^x, e^{-x}]$ não é identicamente nulo, e^x e e^{-x} são linearmente independentes. □

EXEMPLO 7. Considere as funções x^2 e $x|x|$ em $C[-1, 1]$. Ambas pertencem ao subespaço $C^1[-1, 1]$ (ver Exemplo 7 da Seção 2), logo, podemos calcular o wronskiano

$$W[x^2, x|x|] = \begin{vmatrix} x^2 & x|x| \\ 2x & 2|x| \end{vmatrix} \equiv 0$$

Como o wronskiano é identicamente nulo, ele não nos dá informação sobre se as funções são linearmente independentes ou não. Para responder essa pergunta, suponha que

$$c_1x^2 + c_2x|x| = 0$$

para todo x em $[-1, 1]$. Em particular, para $x = 1$ e para $x = -1$, temos

$$c_1 + c_2 = 0$$

$$c_1 - c_2 = 0$$

e a única solução desse sistema é $c_1 = c_2 = 0$. Portanto, as funções x^2 e $x|x|$ são linearmente independentes em $C[-1, 1]$, apesar de $W[x^2, x|x|] \equiv 0$.

Esse exemplo mostra que a recíproca do Teorema 3.3.3 não é válida. □

EXEMPLO 8. Mostre que os vetores $1, x, x^2, x^3$ são linearmente independentes em P_4.

SOLUÇÃO

$$W[1, x, x^2, x^3] = \begin{vmatrix} 1 & x & x^2 & x^3 \\ 0 & 1 & 2x & 3x^2 \\ 0 & 0 & 2 & 6x \\ 0 & 0 & 0 & 6 \end{vmatrix} = 12$$

Como $W[1, x, x^2, x^3] \not\equiv 0$, os vetores são linearmente independentes. □

EXERCÍCIOS

1. Determine se os vetores dados são ou não linearmente independentes em R^2.

(a) $\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ (b) $\begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 6 \end{pmatrix}$

(c) $\begin{pmatrix} -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \end{pmatrix}$ (d) $\begin{pmatrix} -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 2 \\ -4 \end{pmatrix}$

(e) $\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

2. Determine se os vetores dados são ou não linearmente independentes em R^3.

(a) $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ (b) $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

(c) $\begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ -2 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}$ (d) $\begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 4 \\ 2 \\ -4 \end{pmatrix}$

(e) $\begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$

3. Descreva geometricamente o espaço gerado por cada um dos conjuntos de vetores no Exercício 2.

4. Determine se os vetores dados são ou não linearmente independentes em $R^{2\times 2}$.

(a) $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}$

5. Determine se os vetores dados são ou não linearmente independentes em P_3.

(a) $1, x^2, x^2 - 2$ (b) $2, x^2, x, 2x + 3$
(c) $x + 2, x + 1, x^2 - 1$ (d) $x + 2, x^2 - 1$

6. Mostre que os vetores dados são linearmente independentes em $C\,[0, 1]$.

(a) $\cos \pi x, \operatorname{sen} \pi x$ (b) $x^{3/2}, x^{5/2}$
(c) $1, e^x + e^{-x}, e^x - e^{-x}$ (d) e^x, e^{-x}, e^{2x}

7. Determine se os vetores $\cos x$, 1, $\operatorname{sen}^2(x/2)$ são linearmente independentes em $C\,[-\pi, \pi]$.

8. Considere os vetores $\cos(x + \alpha)$ e $\operatorname{sen} x$ em $C\,[-\pi, \pi]$. Para que valores de α os dois vetores vão ser linearmente dependentes? Interprete graficamente sua resposta.

9. Dadas as funções $2x$ e $|x|$, mostre que:
(a) esses dois vetores são linearmente independentes em $C\,[-1, 1]$;
(b) esses dois vetores são linearmente dependentes em $C\,[0, 1]$.

10. Prove que qualquer conjunto finito de vetores contendo o vetor nulo tem que ser linearmente dependente.

11. Sejam $\mathbf{v}_1$ e $\mathbf{v}_2$ dois vetores em um espaço vetorial V. Mostre que $\mathbf{v}_1$ e $\mathbf{v}_2$ são linearmente dependentes se e somente se um dos vetores é um múltiplo do outro.

12. Prove que qualquer subconjunto não-vazio de um conjunto linearmente independente de vetores $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ também é linearmente independente.

13. Seja A uma matriz $m \times n$. Mostre que, se os vetores colunas de A são linearmente independentes, então $N(A) = \{\mathbf{0}\}$.

[***Sugestão:*** para todo $\mathbf{x} \in R^n$, $A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n\mathbf{a}_n$.]

14. Sejam $\mathbf{x}_1, \ldots, \mathbf{x}_k$ vetores linearmente independentes em R^n e seja A uma matriz invertível $n \times n$. Defina $\mathbf{y}_i = A\mathbf{x}_i$ para $i = 1, \ldots, k$. Mostre que $\mathbf{y}_1, \ldots, \mathbf{y}_k$ são linearmente independentes.

15. Seja $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ um conjunto gerador para o espaço vetorial V e seja $\mathbf{v}$ um outro vetor qualquer em V. Mostre que $\mathbf{v}, \mathbf{v}_1, \ldots, \mathbf{v}_n$ são linearmente dependentes.

16. Sejam $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ vetores linearmente independentes em um espaço vetorial V. Mostre que $\mathbf{v}_2, \ldots, \mathbf{v}_n$ não podem gerar V.

4 BASE E DIMENSÃO

Mostramos, na Seção 3, que um conjunto gerador para um espaço vetorial é mínimo se seus elementos são linearmente independentes. Os elementos de um conjunto gerador mínimo formam as peças básicas para a construção de todo o espaço vetorial e, por causa disso, dizemos que eles formam uma "base" para o espaço vetorial.

Definição. Os vetores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ formam uma **base** para um espaço vetorial V se e somente se

(i) $\mathbf{v}_1, \ldots, \mathbf{v}_n$ são linearmente independentes;
(ii) $\mathbf{v}_1, \ldots, \mathbf{v}_n$ geram V.

EXEMPLO 1. A "base canônica" para o R^3 é $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. No entanto, poderíamos usar outra base qualquer, como, por exemplo, $\{(1, 1, 1)^T, (0, 1, 1)^T, (2, 0, 1)^T\}$ ou $\{(1, 1, 1)^T, (1, 1, 0)^T, (1, 0, 1)^T\}$. Veremos, em breve, que qualquer base para R^3 tem exatamente três elementos. □

EXEMPLO 2. Considere o conjunto $\{E_{11}, E_{12}, E_{21}, E_{22}\}$ em $R^{2\times 2}$, onde

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

$$E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Se

$$c_1E_{11} + c_2E_{12} + c_3E_{21} + c_4E_{22} = O$$

então

$$\begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

de modo que $c_1 = c_2 = c_3 = c_4 = 0$. Logo, E_{11}, E_{12}, E_{21}, E_{22} são linearmente independentes. Se A pertence a $R^{2\times2}$, então

$$A = a_{11}E_{11} + a_{12}E_{12} + a_{21}E_{21} + a_{22}E_{22}$$

Logo, E_{11}, E_{12}, E_{21}, E_{22} geram $R^{2\times2}$ e, portanto, formam uma base para $R^{2\times2}$. □

Em muitas aplicações é necessário encontrar um subespaço particular de um espaço vetorial V. Isso pode ser feito encontrando-se os elementos de uma base para o subespaço. Por exemplo, para encontrar todas as soluções do sistema

$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ 2x_1 + x_2 + x_4 &= 0 \end{aligned}$$

precisamos encontrar o núcleo da matriz

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{pmatrix}$$

Vimos, no Exemplo 9 da Seção 2, que $N(A)$ é o subespaço de R^4 gerado pelos vetores

$$\begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

Como esses dois vetores são linearmente independentes, eles formam uma base para $N(A)$.

Teorema 3.4.1. *Se $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ é um conjunto gerador para um espaço vetorial V, então qualquer coleção de m vetores em V, onde $m > n$, é linearmente dependente.*

Demonstração. Sejam $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ m vetores em V, onde $m > n$. Então, como $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ geram V, tem-se

$$\mathbf{u}_i = a_{i1}\mathbf{v}_1 + a_{i2}\mathbf{v}_2 + \cdots + a_{in}\mathbf{v}_n \quad \text{para} \quad i = 1, 2, \ldots, m$$

Uma combinação linear $b_1\mathbf{u}_1 + b_2\mathbf{u}_2 + \ldots + b_m\mathbf{u}_m$ pode ser, então, escrita na forma

$$b_1 \sum_{j=1}^{n} a_{1j}\mathbf{v}_j + b_2 \sum_{j=1}^{n} a_{2j}\mathbf{v}_j + \cdots + b_m \sum_{j=1}^{n} a_{mj}\mathbf{v}_j$$

Arrumando os termos, vemos que

$$b_1\mathbf{u}_1 + b_2\mathbf{u}_2 + \cdots + b_m\mathbf{u}_m = \sum_{i=1}^{m} \left[b_i \left(\sum_{j=1}^{n} a_{ij}\mathbf{v}_j \right) \right] = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ij}b_i \right) \mathbf{v}_j$$

Considere, agora, o sistema de equações

$$\sum_{i=1}^{m} a_{ij} b_i = 0 \qquad j = 1, 2, \ldots, n$$

Esse é um sistema homogêneo com mais equações do que incógnitas. Portanto, pelo Teorema 1.2.1, o sistema tem uma solução não-trivial $(\hat{b}_1, \hat{b}_2, \ldots \hat{b}_m)$. Mas, então,

$$\hat{b}_1 \mathbf{u}_1 + \hat{b}_2 \mathbf{u}_2 + \cdots + \hat{b}_m \mathbf{u}_m = \sum_{j=1}^{n} 0 \mathbf{v}_j = \mathbf{0}$$

logo $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ são linearmente dependentes. □

Corolário 3.4.2. *Se* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ e $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ *são duas bases para um espaço vetorial V, então* $n = m$.

Demonstração. Sejam $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ e $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ duas bases para V. Como $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ geram V e $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ são linearmente independentes, pelo Teorema 3.4.1 temos que $m \leq n$. Pelo mesmo argumento, como $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ geram V e $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são linearmente independentes, $n \leq m$. □

Em vista do Corolário 3.4.2, podemos nos referir ao número de elementos em qualquer base para um espaço vetorial dado. Isso nos leva à seguinte definição:

Definição. Seja V um espaço vetorial. Se V tem uma base com n vetores, dizemos que V tem **dimensão** n. Dizemos que o subespaço $\{\mathbf{0}\}$ de V tem dimensão zero. V tem **dimensão finita** se existe um número finito de vetores que geram V; caso contrário, dizemos que V tem **dimensão infinita.**

Se $\mathbf{x}$ é um vetor não-nulo em R^3, então $\mathbf{x}$ gera um espaço unidimensional, $[\{\mathbf{x}\}] = \{\alpha\mathbf{x} \mid \alpha \text{ é um escalar}\}$. Um vetor $(a, b, c)^T$ pertence a $[\{\mathbf{x}\}]$ se e somente se o ponto (a, b, c) pertence à reta determinada por $(0, 0, 0)$ e (x_1, x_2, x_3). Portanto, um subespaço unidimensional de R^3 pode ser representado geometricamente por uma reta contendo a origem.

Se $\mathbf{x}$ e $\mathbf{y}$ são linearmente independentes em R^3, então $[\{\mathbf{x}, \mathbf{y}\}] = \{\alpha\mathbf{x} + \beta\mathbf{y} \mid \alpha \text{ e } \beta \text{ são escalares}\}$ é um subespaço bidimensional de R^3. Um vetor $(a, b, c)^T$ pertence a $[\{\mathbf{x}, \mathbf{y}\}]$ se e somente se o ponto (a, b, c) pertence ao plano determinado por $(0, 0, 0)$, (x_1, x_2, x_3) e (y_1, y_2, y_3). Portanto, podemos pensar em um subespaço bidimensional de R^3 como sendo um plano contendo a origem. Se $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{z}$ são linearmente independentes em R^3, eles formam uma base para R^3 e $[\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}] = R^3$. Então, qualquer outro ponto $(a, b, c)^T$ tem que pertencer a $[\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}]$ (ver a Fig. 3.4.1).

EXEMPLO 3. Seja P o espaço vetorial de todos os polinômios. Vamos mostrar que P tem dimensão infinita. Se P tivesse dimensão finita, por exemplo, dimensão n, qualquer conjunto de $n + 1$ vetores seria linearmente dependente. No entanto, $1, x, x^2, \ldots, x^n$ são linearmente independentes, já que $W[1, x, x^2, \ldots, x^n] > 0$. Portanto, P não pode ter dimensão n. Como n era arbitrário, P tem que ser de dimensão infinita. O mesmo argumento mostra que $C[a, b]$ tem dimensão infinita. □

Teorema 3.4.3. *Seja V um espaço vetorial de dimensão* $n > 0$. *Então:*

(I) *Qualquer conjunto de n vetores linearmente independentes gera V;*
(II) *Quaisquer n vetores que geram V são linearmente independentes.*

Demonstração. Para provar (I), vamos supor que $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são linearmente independentes e que $\mathbf{v}$ é outro vetor qualquer em V. Como V tem dimensão n, ele tem uma base com n vetores e esses vetores geram V. Pelo Teorema 3.4.1, temos, então, que $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}$ são linearmente dependentes. Logo, existem escalares $c_1, c_2, \ldots, c_n, c_{n+1}$, nem todos nulos, tais que

(1) $$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n + c_{n+1}\mathbf{v} = \mathbf{0}$$

O escalar c_{n+1} não pode ser nulo, senão teríamos $\mathbf{v}_1, \ldots, \mathbf{v}_n$ linearmente dependentes. Podemos, então, resolver (1) para $\mathbf{v}$.

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n$$

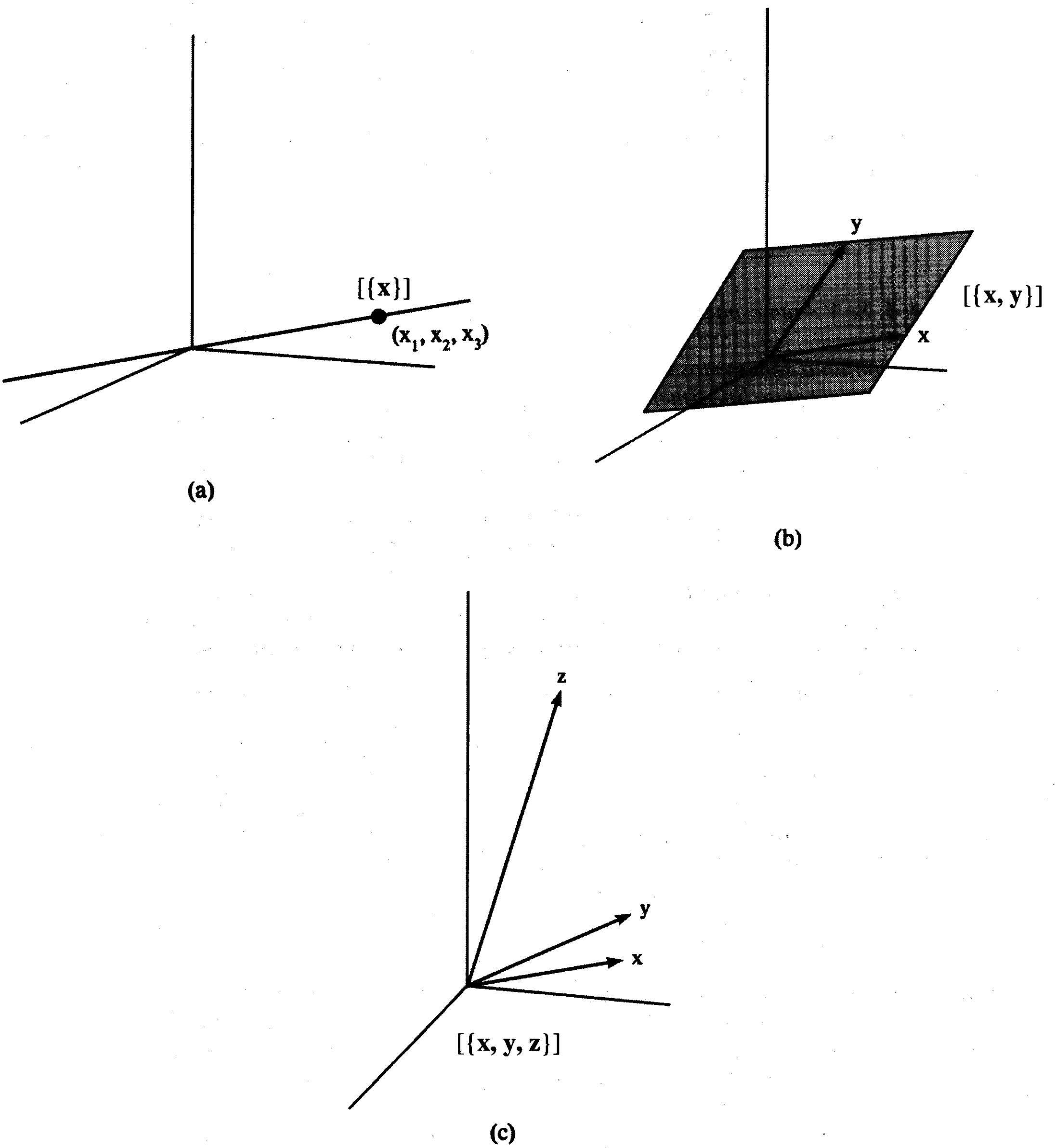


FIG. 3.4.1

onde $\alpha_i = -c_i/c_{n+1}$ para $i = 1, 2, \ldots, n$. Como $\mathbf{v}$ era um vetor arbitrário em V, temos que $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ geram V.

Para provar (II), suponha que $\mathbf{v}_1, \ldots, \mathbf{v}_n$ geram V. Se $\mathbf{v}_1, \ldots, \mathbf{v}_n$ fossem linearmente dependentes, então um dos $\mathbf{v}_i$, por exemplo, $\mathbf{v}_n$, poderia ser escrito como combinação linear dos outros. Concluiríamos, então, que $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ ainda geraria V. Se $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ também fossem linearmente dependentes, poderíamos eliminar outro vetor e ainda continuar com um conjunto gerador. Poderíamos continuar eliminando vetores dessa maneira até chegar a um conjunto gerador linearmente independente com $k < n$ elementos. Mas isso contradiz o fato de que $\dim V = n$. Portanto, $\mathbf{v}_1, \ldots, \mathbf{v}_n$ têm que ser linearmente independentes. □

EXEMPLO 4. Mostre que $\left\{\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}-2\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}\right\}$ é uma base para R^3.

SOLUÇÃO. Como dim $R^3 = 3$, basta mostrar que esses três vetores são linearmente independentes. Isso segue do fato de que

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

□

Teorema 3.4.4. *Se V é um espaço vetorial de dimensão n > 0, então:*

(i) *nenhum conjunto com menos de n vetores pode gerar V;*
(ii) *qualquer subconjunto linearmente independente com menos de n elementos pode ser estendido para formar uma base para V;*
(iii) *podem-se retirar elementos de qualquer conjunto gerador contendo mais de n vetores de modo a se obter uma base para V.*

Demonstração. A observação (i) segue pelo mesmo argumento utilizado no Teorema 3.4.3 para provar (II). Para provar (ii), suponha que $\mathbf{v}_1, \ldots, \mathbf{v}_k$ são vetores linearmente independente e que $k < n$. De (i), $[\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}]$ é um subespaço próprio de V, logo existe um vetor $\mathbf{v}_{k+1}$ que está em V, mas não pertence a $[\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}]$. Temos, então, que os vetores $\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}$ são linearmente independentes. Se $k + 1 < n$, podemos estender $\{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}\}$, da mesma maneira, a um conjunto linearmente independente com $k + 2$ vetores. Esse processo pode ser continuado até obtermos um conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ de vetores linearmente independentes.

Para provar (iii), suponha que $\mathbf{v}_1, \ldots, \mathbf{v}_m$ geram V e que $m > n$. Pelo Teorema 3.4.1, $\mathbf{v}_1, \ldots, \mathbf{v}_m$ são linearmente dependentes. Temos, então, que um dos vetores, por exemplo, $\mathbf{v}_m$, pode ser escrito como uma combinação linear dos outros. Logo, se retirarmos $\mathbf{v}_m$ do conjunto, os $m - 1$ vetores restantes ainda geram V. Se $m - 1 > n$, podemos continuar a retirar vetores do conjunto até chegarmos a um conjunto gerador contendo n elementos. □

BASES CANÔNICAS

No Exemplo 1 dissemos que o conjunto $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ era a *base canônica* para R^3. Chamamos essa base de canônica por ela ser a mais natural para se representar vetores em R^3. Mais geralmente, a base canônica para R^n é o conjunto $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$.

A maneira mais natural de representar matrizes em $R^{2\times2}$ é em termos da base $\{E_{11}, E_{12}, E_{21}, E_{22}\}$ dada no Exemplo 2. Essa é, então, a base canônica para $R^{2\times2}$.

A maneira padrão de representar um polinômio em P_n é em termos das funções $1, x, x^2, \ldots, x^{n-1}$ e, por isso, a base canônica para P_n é $(1, x, x^2, \ldots, x^{n-1}\}$.

Embora essas bases canônicas pareçam ser as mais simples e naturais para se usar, elas não são as bases mais apropriadas para muitos problemas aplicados. (Veja, por exemplo, o problema de mínimos quadráticos no Cap. 5 ou as aplicações de autovalores no Cap. 6.) De fato, a chave na resolução de muitos problemas aplicados é mudar de uma das bases canônicas para uma base que é, de alguma forma, mais natural para a aplicação em questão. Uma vez resolvido o problema na nova base, é fácil voltar e representar a solução em termos da base canônica. Na próxima seção vamos aprender a mudar de uma base para outra.

EXERCÍCIOS

1. Indique se os vetores dados no Exercício 1 da Seção 3 formam ou não uma base para R^2.
2. Indique se os vetores dados no Exercício 2 da Seção 3 formam ou não uma base para R^3.

3. Considere os vetores

$$\mathbf{x}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} 4 \\ 3 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} 7 \\ -3 \end{pmatrix}$$

(a) Mostre que $\mathbf{x}_1$ e $\mathbf{x}_2$ formam uma base para R^2.
(b) Por que $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ têm que ser linearmente dependentes?
(c) Qual a dimensão de $[\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}]$?

4. Considere os vetores

$$\mathbf{x}_1 = \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} -3 \\ 2 \\ -4 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} -6 \\ 4 \\ -8 \end{pmatrix}$$

Qual a dimensão de $[\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}]$?

5. Considere

$$\mathbf{x}_1 = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} 2 \\ 6 \\ 4 \end{pmatrix}$$

(a) Mostre que $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ são linearmente dependentes.
(b) Mostre que $\mathbf{x}_1$, $\mathbf{x}_2$ são linearmente independentes.
(c) Qual a dimensão de $[\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}]$?
(d) Descreva geometricamente $[\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}]$.

6. Alguns dos conjuntos no Exercício 2 da Seção 2 formavam subespaços de R^3. Em cada um desses casos, encontre uma base para o subespaço e determine sua dimensão.

7. Encontre uma base para o subespaço S de R^4 formado por todos os vetores da forma $(a + b, a - b + 2c, b, c)^T$, onde a, b e c são números reais. Qual a dimensão de S?

8. Considere os vetores $\mathbf{x}_1 = (1, 1, 1)^T$ e $\mathbf{x}_2 = (3, -1, 4)^T$.
(a) $\mathbf{x}_1$ e $\mathbf{x}_2$ geram R^3? Explique.
(b) Seja $\mathbf{x}_3$ um terceiro vetor em R^3 e defina $X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$. Que condição (ou condições) X tem que satisfazer para que $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ formem uma base para R^3?
(c) Encontre um terceiro vetor $\mathbf{x}_3$ que estenda o conjunto $\{\mathbf{x}_1, \mathbf{x}_2\}$ a uma base para R^3.

9. Os vetores

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} 2 \\ 5 \\ 4 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \quad \mathbf{x}_4 = \begin{pmatrix} 2 \\ 7 \\ 4 \end{pmatrix}, \quad \mathbf{x}_5 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

geram R^3. Retire algum (ou alguns) elementos de $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5\}$ de modo a obter uma base para R^3.

10. Seja S o subespaço de P_3 formado por todos os polinômios da forma $ax^2 + bx + 2a + 3b$. Encontre uma base para S.

11. Alguns dos conjuntos no Exercício 3 da Seção 2 formavam subespaços de $R^{2\times2}$. Em cada um desses casos, encontre uma base para o subespaço e determine sua dimensão.

12. Encontre a dimensão do espaço gerado por 1, $\cos 2x$, $\cos^2 x$ em $C[-\pi, \pi]$.

13. Encontre a dimensão do subespaço de P_3 gerado pelos vetores dados em cada um dos itens a seguir.

(a) $x, x - 1, x^2 + 1$ (b) $x, x - 1, x^2 + 1, x^2 - 1$
(c) $x^2, x^2 - x - 1, x + 1$ (d) $2x, x - 2$

14. Seja S o subespaço de P_3 formado por todos os polinômios $p(x)$ satisfazendo $p(0) = 0$, e seja T o subespaço de todos os polinômios $q(x)$ tais que $q(1) = 0$. Encontre bases para

(a) S (b) T (c) $S \cap T$

15. Seja U o subespaço de R^4 formado pelos vetores da forma $(u_1, u_2, 0, 0)^T$ e seja V o subespaço de todos os vetores da forma $(0, v_2, v_3, 0)^T$. Quais as dimensões de U, V, $U \cap V$, $U + V$? Encontre uma base para cada um desses subespaços.

16. É possível encontrar um par de subespaços bidimensionais U e V de R^3 tais que $U \cap V = \{\mathbf{0}\}$? Justifique sua resposta. Interprete geometricamente sua conclusão.

[***Sugestão:*** sejam $\{\mathbf{u}_1, \mathbf{u}_2\}$ e $\{\mathbf{v}_1, \mathbf{v}_2\}$ bases para U e V, respectivamente; mostre que $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{v}_1$, $\mathbf{v}_2$ são linearmente dependentes.]

5 MUDANÇA DE BASES

Muitos problemas aplicados podem ser simplificados mudando-se de um sistema de coordenadas para outro. Mudar sistemas de coordenadas em um espaço vetorial é, essencialmente, a mesma coisa que mudar de base. Por exemplo, ao descrever o movimento de uma partícula no plano em um instante particular, é muitas vezes conveniente usar uma base de R^2 formada por um vetor tangente unitário $\mathbf{t}$ e um vetor normal unitário $\mathbf{n}$, em vez da base canônica $\{\mathbf{e}_1, \mathbf{e}_2\}$.

Nesta seção, vamos discutir o problema de mudar de um sistema de coordenadas para outro. Vamos mostrar que isso pode ser feito multiplicando-se um vetor de coordenadas dado $\mathbf{x}$ por uma matriz invertível S. O produto $\mathbf{y} = S\mathbf{x}$ vai ser o vetor de coordenadas para o novo sistema.

MUDANÇA DE COORDENADAS EM R^2

A base canônica para R^2 é $\{\mathbf{e}_1, \mathbf{e}_2\}$. Qualquer vetor $\mathbf{x}$ em R^2 pode ser escrito como uma combinação linear

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$$

Os escalares x_1 e x_2 são as *coordenadas* de $\mathbf{x}$ em relação à base canônica. De fato, para qualquer base $\{\mathbf{y}, \mathbf{z}\}$ para R^2, pelo Teorema 3.3.2, um dado vetor $\mathbf{x}$ pode ser representado de maneira única como uma combinação linear

$$\mathbf{x} = \alpha\mathbf{y} + \beta\mathbf{z}$$

Os escalares α e β são as coordenadas de $\mathbf{x}$ em relação à base $\{\mathbf{y}, \mathbf{z}\}$. Vamos ordenar os elementos da base de modo que $\mathbf{y}$ seja o primeiro vetor da base e $\mathbf{z}$ seja o segundo, e vamos denotar a base ordenada por $[\mathbf{y}, \mathbf{z}]$*. Podemos, então, nos referir ao vetor $(\alpha, \beta)^T$ como sendo o *vetor de coordenadas* de $\mathbf{x}$ em relação à base $[\mathbf{y}, \mathbf{z}]$.

EXEMPLO 1. Sejam $\mathbf{y} = (2, 1)^T$ e $\mathbf{z} = (1, 4)^T$. Os vetores $\mathbf{y}$ e $\mathbf{z}$ são linearmente independentes e, portanto, formam uma base para R^2. O vetor $\mathbf{x} = (7, 7)^T$ pode ser escrito como uma combinação linear

$$\mathbf{x} = 3\mathbf{y} + \mathbf{z}$$

Logo, o vetor de coordenadas de $\mathbf{x}$ em relação a $[\mathbf{y}, \mathbf{z}]$ é $(3, 1)^T$. Geometricamente, esse vetor nos diz como sair da origem e chegar em $(7, 7)$, movendo-nos primeiro na direção de $\mathbf{y}$ e depois na direção de $\mathbf{z}$. O vetor de coordenadas de $\mathbf{x}$ em relação à base ordenada $[\mathbf{z}, \mathbf{y}]$ é $(1, 3)^T$. Geometricamente, esse vetor nos diz como sair da origem e chegar em $(7, 7)$ movendo-nos primeiro na direção de $\mathbf{z}$ e depois na direção de $\mathbf{y}$ (ver Fig. 3.5.1). □

Uma vez decididos a trabalhar com uma nova base, temos o problema de encontrar as coordenadas em relação a essa nova base. Suponha, por exemplo, que, em vez de usarmos a base canônica $\{\mathbf{e}_1, \mathbf{e}_2\}$ para o R^2, queira usar uma base diferente, por exemplo,

$$\mathbf{u}_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

* Não confundir com o espaço gerado por y e z, que é denotado por $[\{\mathbf{y}, \mathbf{z}\}]$. A notação com colchetes para bases ordenadas não é padrão. (N.T.)

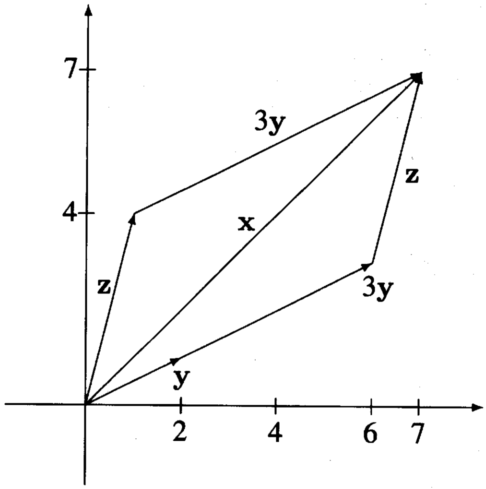


FIG. 3.5.1

De fato, podemos querer mudar nos dois sentidos entre os dois sistemas de coordenadas. Vamos considerar os dois problemas seguintes:

I. dado um vetor $\mathbf{x} = (x_1, x_2)^T$, encontre suas coordenadas em relação a $\mathbf{u}_1$ e $\mathbf{u}_2$;
II. dado um vetor $\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2$, encontre suas coordenadas em relação a $\mathbf{e}_1$ e $\mathbf{e}_2$.

Vamos resolver o problema II primeiro, que é o mais fácil. Para mudar da base $[\mathbf{u}_1, \mathbf{u}_2]$ para a base $[\mathbf{e}_1, \mathbf{e}_2]$, precisamos expressar os elementos da base antiga, $\mathbf{u}_1$ e $\mathbf{u}_2$, em termos dos elementos da nova base, $\mathbf{e}_1$ e $\mathbf{e}_2$.

$$\mathbf{u}_1 = 3\mathbf{e}_1 + 2\mathbf{e}_2$$
$$\mathbf{u}_2 = \mathbf{e}_1 + \mathbf{e}_2$$

Temos, então, que

$$\begin{aligned} \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 &= (3\alpha_1\mathbf{e}_1 + 2\alpha_1\mathbf{e}_2) + (\alpha_2\mathbf{e}_1 + \alpha_2\mathbf{e}_2) \\ &= (3\alpha_1 + \alpha_2)\mathbf{e}_1 + (2\alpha_1 + \alpha_2)\mathbf{e}_2 \end{aligned}$$

Logo, o vetor de coordenadas de $\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2$ em relação a $[\mathbf{e}_1, \mathbf{e}_2]$ é

$$\mathbf{x} = \begin{pmatrix} 3\alpha_1 + \alpha_2 \\ 2\alpha_1 + \alpha_2 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$$

Definindo

$$U = (\mathbf{u}_1, \mathbf{u}_2) = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$$

temos que, dado qualquer vetor de coordenadas α em relação a $[\mathbf{u}_1, \mathbf{u}_2]$, para encontrar o vetor de coordenadas correspondente $\mathbf{x}$ em relação a $[\mathbf{e}_1, \mathbf{e}_2]$, basta multiplicar U e $\boldsymbol{\alpha}$.

(1) $$\mathbf{x} = U\boldsymbol{\alpha}$$

A matriz U é chamada de *matriz mudança de base* de $[\mathbf{u}_1, \mathbf{u}_2]$ para $[\mathbf{e}_1, \mathbf{e}_2]$.

Para resolver o problema I, precisamos encontrar a matriz mudança de base de $[\mathbf{e}_1, \mathbf{e}_2]$ para $[\mathbf{u}_1, \mathbf{u}_2]$. A matriz U é invertível, já que suas colunas são vetores linearmente independentes. Temos, então, de (1),

$$\boldsymbol{\alpha} = U^{-1}\mathbf{x}$$

Logo, dado um vetor

$$\mathbf{x} = (x_1, x_2)^T = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$$

basta multiplicá-lo por U^{-1} para encontrar seu vetor de coordenadas em relação a $[\mathbf{u}_1, \mathbf{u}_2]$. U^{-1} é a matriz mudança de base de $[\mathbf{e}_1, \mathbf{e}_2]$ para $[\mathbf{u}_1, \mathbf{u}_2]$.

EXEMPLO 2. Sejam $\mathbf{u}_1 = (3, 2)^T$, $\mathbf{u}_2 = (1, 1)^T$ e $\mathbf{x} = (7, 4)^T$. Encontre as coordenadas de $\mathbf{x}$ em relação a $[\mathbf{u}_1, \mathbf{u}_2]$.

SOLUÇÃO. Pela discussão precedente, a matriz mudança de base de $[\mathbf{e}_1, \mathbf{e}_2]$ para $[\mathbf{u}_1, \mathbf{u}_2]$ é a inversa de

$$U = (\mathbf{u}_1, \mathbf{u}_2) = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$$

Logo,

$$\boldsymbol{\alpha} = U^{-1}\mathbf{x} = \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix}\begin{pmatrix} 7 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$$

é o vetor de coordenadas desejado e

$$\mathbf{x} = 3\mathbf{u}_1 - 2\mathbf{u}_2$$ □

EXEMPLO 3. Sejam $\mathbf{b}_1 = (1, -1)^T$ e $\mathbf{b}_2 = (-2, 3)^T$. Encontre a matriz mudança de base de $[\mathbf{e}_1, \mathbf{e}_2]$ para $[\mathbf{b}_1, \mathbf{b}_2]$ e determine as coordenadas de $\mathbf{x} = (1, 2)^T$ em relação a $[\mathbf{b}_1, \mathbf{b}_2]$.

SOLUÇÃO. A matriz mudança de base de $[\mathbf{b}_1, \mathbf{b}_2]$ para $[\mathbf{e}_1, \mathbf{e}_2]$ é

$$B = (\mathbf{b}_1, \mathbf{b}_2) = \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix}$$

logo a matriz mudança de base de $[\mathbf{e}_1, \mathbf{e}_2]$ para $[\mathbf{b}_1, \mathbf{b}_2]$ é

$$B^{-1} = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}$$

O vetor de coordenadas de $\mathbf{x}$ em relação a $[\mathbf{b}_1, \mathbf{b}_2]$ é

$$\boldsymbol{\alpha} = B^{-1}\mathbf{x} = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \end{pmatrix}$$

e, portanto,

$$\mathbf{x} = 7\mathbf{b}_1 + 3\mathbf{b}_2$$ □

Se

$$S = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix}$$

é a matriz mudança de base de uma base ordenada $[\mathbf{v}_1, \mathbf{v}_2]$ em R^2 para outra base ordenada $[\mathbf{u}_1, \mathbf{u}_2]$, então, como

$$\mathbf{v}_1 = 1\mathbf{v}_1 + 0\mathbf{v}_2$$

o vetor de coordenadas de $\mathbf{v}_1$ em relação a $[\mathbf{u}_1, \mathbf{u}_2]$ é dado por

$$\mathbf{s}_2 = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} s_{11} \\ s_{21} \end{pmatrix}$$

Analogamente,

$$\mathbf{v}_2 = 0\mathbf{v}_1 + 1\mathbf{v}_2$$

e seu vetor de coordenadas em relação a $[\mathbf{u}_1, \mathbf{u}_2]$ é dado por

$$\mathbf{s}_2 = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} s_{12} \\ s_{22} \end{pmatrix}$$

Logo,

$$\begin{aligned} \mathbf{v}_1 &= s_{11}\mathbf{u}_1 + s_{21}\mathbf{u}_2 \\ \mathbf{v}_2 &= s_{12}\mathbf{u}_1 + s_{22}\mathbf{u}_2 \end{aligned} \tag{2}$$

Em geral, se os elementos da base antiga $\mathbf{v}_1$ e $\mathbf{v}_2$ são escritos em termos da nova base $[\mathbf{u}_1, \mathbf{u}_2]$, o vetor de coordenadas $\mathbf{s}_1 = (s_{11}, s_{21})^T$ correspondente a $\mathbf{v}_1$ é a primeira coluna da matriz mudança de base S e o vetor de coordenadas $\mathbf{s}_2 = (s_{12}, s_{22})^T$ correspondente a $\mathbf{v}_2$ é a segunda coluna de S. Logo, S é a transposta da matriz de coeficientes em (2).

EXEMPLO 4. Encontre a matriz mudança de base de $[\mathbf{v}_1, \mathbf{v}_2]$ para $[\mathbf{u}_1, \mathbf{u}_2]$, onde

$$\mathbf{v}_1 = \begin{pmatrix} 5 \\ 2 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 7 \\ 3 \end{pmatrix} \quad \text{e} \quad \mathbf{u}_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

SOLUÇÃO. Precisamos escrever $\mathbf{v}_1$ e $\mathbf{v}_2$ em termos dos elementos da nova base, $\mathbf{u}_1$ e $\mathbf{u}_2$:

$$\mathbf{v}_1 = s_{11}\mathbf{u}_1 + s_{21}\mathbf{u}_2$$

$$\mathbf{v}_2 = s_{12}\mathbf{u}_1 + s_{22}\mathbf{u}_2$$

A primeira equação pode ser escrita como

$$\begin{pmatrix} 5 \\ 2 \end{pmatrix} = \begin{pmatrix} 3s_{11} + s_{21} \\ 2s_{11} + s_{21} \end{pmatrix}$$

A solução desse sistema é $(s_{11}, s_{21})^T = (3, -4)^T$. Analogamente, a segunda equação nos leva ao sistema

$$\begin{pmatrix} 7 \\ 3 \end{pmatrix} = \begin{pmatrix} 3s_{12} + s_{22} \\ 2s_{12} + s_{22} \end{pmatrix}$$

cuja solução é $(s_{12}, s_{22})^T = (4, -5)^T$. Portanto,

$$S = \begin{pmatrix} 3 & 4 \\ -4 & -5 \end{pmatrix}$$

é a matriz mudança de base de $[\mathbf{v}_1, \mathbf{v}_2]$ para $[\mathbf{u}_1, \mathbf{u}_2]$. □

Um método alternativo de mudar de uma base $[\mathbf{v}_1, \mathbf{v}_2]$ para outra base $[\mathbf{u}_1, \mathbf{u}_2]$ é mudar primeiro de $[\mathbf{v}_1, \mathbf{v}_2]$ para a base canônica e depois mudar para $[\mathbf{u}_1, \mathbf{u}_2]$. Dado um vetor $\mathbf{x}$ em R^2, se $\mathbf{c}$ é o vetor de coordenadas de $\mathbf{x}$ em relação a $[\mathbf{v}_1, \mathbf{v}_2]$ e $\mathbf{d}$ é o vetor de coordenadas de $\mathbf{x}$ em relação a $[\mathbf{u}_1, \mathbf{u}_2]$, então

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 = d_1\mathbf{u}_1 + d_2\mathbf{u}_2$$

Como V é a matriz mudança de base de $[\mathbf{v}_1, \mathbf{v}_2]$ para $[\mathbf{e}_1, \mathbf{e}_2]$ e U^{-1} é a matriz mudança de base de $[\mathbf{e}_1, \mathbf{e}_2]$ para $[\mathbf{u}_1, \mathbf{u}_2]$, tem-se que

$$V\mathbf{c} = \mathbf{x} \qquad \text{e} \qquad U^{-1}\mathbf{x} = \mathbf{d}$$

e, portanto,

$$U^{-1}V\mathbf{c} = U^{-1}\mathbf{x} = \mathbf{d}$$

Logo, $U^{-1}V$ é a matriz mudança de base de $[\mathbf{v}_1, \mathbf{v}_2]$ para $[\mathbf{u}_1, \mathbf{u}_2]$ (ver Fig. 3.5.2).

EXEMPLO 5. Sejam $[\mathbf{v}_1, \mathbf{v}_2]$ e $[\mathbf{u}_1, \mathbf{u}_2]$ as bases ordenadas do Exemplo 4. A matriz mudança de base de $[\mathbf{v}_1, \mathbf{v}_2]$ para $[\mathbf{u}_1, \mathbf{u}_2]$ é dada por

$$U^{-1}V = \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 5 & 7 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ -4 & -5 \end{pmatrix}$$

□

$$\begin{tikzcd}
{[\mathbf{v}_1, \mathbf{v}_2]} \arrow[r, "V"] \arrow[dr, "U^{-1}V"'] & {[\mathbf{e}_1, \mathbf{e}_2]} \arrow[d, "U^{-1}"] \\
& {[\mathbf{u}_1, \mathbf{u}_2]}
\end{tikzcd}$$

FIG. 3.5.2

MUDANÇA DE BASE EM UM ESPAÇO VETORIAL GERAL

Tudo que fizemos até agora pode ser generalizado facilmente para qualquer espaço vetorial de dimensão finita. Vamos começar definindo vetores de coordenadas em um espaço vetorial de dimensão n.

Definição. Seja V um espaço vetorial com base ordenada $E = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$. Se $\mathbf{v}$ é um elemento qualquer de V, então $\mathbf{v}$ pode ser escrito na forma

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

onde $c_1, c_2, \ldots, c_n$ são escalares. Podemos associar, então, a cada vetor $\mathbf{v}$ um único vetor $\mathbf{c} = (c_1, c_2, \ldots, c_n)^T$ em R^n. O vetor $\mathbf{c}$ assim definido é chamado de **vetor de coordenadas** de $\mathbf{v}$ em relação à base ordenada E e é denotado por $[\mathbf{v}]_E$. Os c_i são as **coordenadas** de $\mathbf{v}$ em relação a E.

Os exemplos considerados até agora trataram apenas de mudança de coordenadas em R^2. Técnicas análogas podem ser usadas em R^n. No caso de R^n, as matrizes de mudança de base serão $n \times n$.

EXEMPLO 6. Sejam

$$\begin{aligned} E &= [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3] = [(1,1,1)^T, (2,3,2)^T, (1,5,4)^T] \\ F &= [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3] = [(1,1,0)^T, (1,2,0)^T, (1,2,1)^T] \end{aligned}$$

Encontre a matriz mudança de base de E para F. Se

$$\mathbf{x} = 3\mathbf{v}_1 + 2\mathbf{v}_2 - \mathbf{v}_3 \qquad \text{e} \qquad \mathbf{y} = \mathbf{v}_1 - 3\mathbf{v}_2 + 2\mathbf{v}_3$$

encontre as coordenadas de $\mathbf{x}$ e $\mathbf{y}$ em relação à base ordenada F.

SOLUÇÃO. Como no Exemplo 5, a matriz mudança de base é dada por

$$U^{-1}V = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 5 \\ 1 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -3 \\ -1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Os vetores de coordenadas de $\mathbf{x}$ e de $\mathbf{y}$ em relação à base ordenada F são dados por

$$[\mathbf{x}]_F = \begin{pmatrix} 1 & 1 & -3 \\ -1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 8 \\ -5 \\ 3 \end{pmatrix}$$

e

$$[\mathbf{y}]_F = \begin{pmatrix} 1 & 1 & -3 \\ -1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix} = \begin{pmatrix} -8 \\ 2 \\ 3 \end{pmatrix}$$

O leitor pode verificar que

$$\begin{aligned} 8\mathbf{u}_1 - 5\mathbf{u}_2 + 3\mathbf{u}_3 &= 3\mathbf{v}_1 + 2\mathbf{v}_2 - \mathbf{v}_3 \\ -8\mathbf{u}_1 + 2\mathbf{u}_2 + 3\mathbf{u}_3 &= \mathbf{v}_1 - 3\mathbf{v}_2 + 2\mathbf{v}_3 \end{aligned}$$ □

Se V é um espaço vetorial qualquer de dimensão n, é possível mudar de uma base para outra através de uma matriz mudança de base $n \times n$. Vamos mostrar que uma tal matriz é necessariamente invertível. Para ver como fazer isso, suponha que $E = [\mathbf{w}_1, \ldots, \mathbf{w}_n]$ e $F = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ são duas bases ordenadas para V. Cada vetor $\mathbf{w}_j$ pode ser, então, expresso como uma combinação linear dos $\mathbf{v}_i$.

$$\begin{aligned} \mathbf{w}_1 &= s_{11}\mathbf{v}_1 + s_{21}\mathbf{v}_2 + \cdots + s_{n1}\mathbf{v}_n \\ \mathbf{w}_2 &= s_{12}\mathbf{v}_1 + s_{22}\mathbf{v}_2 + \cdots + s_{n2}\mathbf{v}_n \\ &\vdots \\ \mathbf{w}_n &= s_{1n}\mathbf{v}_1 + s_{2n}\mathbf{v}_2 + \cdots + s_{nn}\mathbf{v}_n \end{aligned} \tag{3}$$

Seja $\mathbf{v} \in V$. Se $\mathbf{x} = [\mathbf{v}]_E$, obtemos, de (3),

$$\begin{aligned} \mathbf{v} &= x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + \cdots + x_n\mathbf{w}_n \\ &= \left(\sum_{j=1}^n s_{1j}x_j\right)\mathbf{v}_1 + \left(\sum_{j=1}^n s_{2j}x_j\right)\mathbf{v}_2 + \cdots + \left(\sum_{j=1}^n s_{nj}x_j\right)\mathbf{v}_n \end{aligned}$$

Logo, se $\mathbf{y} = [\mathbf{v}]_F$, temos

$$y_i = \sum_{j=1}^n s_{ij}x_j \qquad i = 1, \ldots, n$$

e, portanto,

$$\mathbf{y} = S\mathbf{x}$$

A matriz S definida por (3) é a *matriz mudança de base*. Uma vez determinada S, é fácil mudar de sistema de coordenadas. Para encontrar as coordenadas de $\mathbf{v} = x_1\mathbf{w}_1 + \ldots + x_n\mathbf{w}_n$ em relação a $[\mathbf{v}_1, \ldots, \mathbf{v}_n]$, basta calcular $\mathbf{y} = S\mathbf{x}$.

A matriz mudança de base S correspondente à mudança da base $[\mathbf{w}_1, \ldots, \mathbf{w}_n]$ para a base $[\mathbf{v}_1, \ldots, \mathbf{v}_n]$ pode ser caracterizada pela condição

$$S\mathbf{x} = \mathbf{y} \text{ se e somente se } x_1\mathbf{w}_1 + \cdots + x_n\mathbf{w}_n = y_1\mathbf{v}_1 + \cdots + y_n\mathbf{v}_n \tag{4}$$

Fazendo $\mathbf{y} = 0$ em (4), vemos que $S\mathbf{x} = \mathbf{0}$ implica que

$$x_1\mathbf{w}_1 + \cdots + x_n\mathbf{w}_n = \mathbf{0}$$

Como os $\mathbf{w}_i$ são linearmente independentes, segue que $\mathbf{x} = \mathbf{0}$. Logo, a equação $S\mathbf{x} = \mathbf{0}$ tem apenas a solução trivial e, portanto, a matriz S é invertível. A matriz inversa é caracterizada pela condição

$$S^{-1}\mathbf{y} = \mathbf{x} \text{ se e somente se } y_1\mathbf{v}_1 + \cdots + y_n\mathbf{v}_n = x_1\mathbf{w}_1 + \cdots + x_n\mathbf{w}_n$$

Então S^{-1} é a matriz mudança de base que muda da base $[\mathbf{v}_1, \ldots, \mathbf{v}_n]$ para a base $[\mathbf{w}_1, \ldots, \mathbf{w}_n]$.

EXEMPLO 7. Suponha que queremos mudar, em P_3, da base ordenada $[1, x, x^2]$ para a base ordenada $[1, 2x, 4x^2 - 2]$. Como $[1, x, x^2]$ é a base canônica para P_3, é mais fácil encontrar a matriz mudança de base de $[1, 2x, 4x^2 - 2]$ para $[1, x, x^2]$. Como

$$\begin{aligned} 1 &= 1 \cdot 1 + 0x + 0x^2 \\ 2x &= 0 \cdot 1 + 2x + 0x^2 \\ 4x^2 - 2 &= -2 \cdot 1 + 0x + 4x^2 \end{aligned}$$

a matriz mudança de base é

$$S = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

A inversa de S vai ser a matriz que muda da base $[1, x, x^2]$ para a base $[1, 2x, 4x^2 - 2]$:

$$S^{-1} = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} \end{pmatrix}$$

Dado qualquer $p(x) = a + bx + cx^2$ em P_3, para encontrar as coordenadas de $p(x)$ em relação a $[1, 2x, 4x^2 - 2]$, basta multiplicar

$$\begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a + \frac{1}{2}c \\ \frac{1}{2}b \\ \frac{1}{4}c \end{pmatrix}$$

Logo,

$$p(x) = (a + \tfrac{1}{2}c) \cdot 1 + (\tfrac{1}{2}b) \cdot 2x + \tfrac{1}{4}c \cdot (4x^2 - 2) \qquad \square$$

Vimos que cada matriz mudança de base é invertível. De fato, podemos pensar em qualquer matriz invertível como uma matriz mudança de base. Se S é uma matriz invertível $n \times n$ e $[\mathbf{v}_1, \ldots, \mathbf{v}_n]$ é uma base ordenada para V, defina $[\mathbf{w}_1, \ldots, \mathbf{w}_n]$ por (3). Para ver que os $\mathbf{w}_j$ são linearmente independentes, suponha que

$$\sum_{j=1}^{n} x_j \mathbf{w}_j = \mathbf{0}$$

De (3), tem-se que

$$\sum_{i=1}^{n} \left(\sum_{j=1}^{n} s_{ij} x_j \right) \mathbf{v}_i = \mathbf{0}$$

Pela independência linear dos $\mathbf{v}_i$, tem-se que

$$\sum_{j=1}^{n} s_{ij} x_j = 0 \qquad i = 1, \ldots, n$$

ou, equivalentemente,

$$S\mathbf{x} = \mathbf{0}$$

Como S é invertível, $\mathbf{x}$ tem que ser igual a $\mathbf{0}$. Logo, $\mathbf{w}_1, \ldots, \mathbf{w}_n$ são linearmente independentes e, portanto, formam uma base para V. A matriz S é a matriz que efetua a mudança da base ordenada $[\mathbf{w}_1, \ldots, \mathbf{w}_n]$ para $[\mathbf{v}_1, \ldots, \mathbf{v}_n]$.

Em muitos problemas aplicados é importante usar o tipo certo de base para a aplicação em questão. Veremos, no Cap. 5, que a chave para a resolução de problemas de mínimos quadráticos é usar um tipo especial de base, uma base *ortonormal*. No Cap. 6, vamos considerar um número de aplicações envolvendo *autovalores* e *autovetores* associados a uma matriz A $n \times n$. A chave para resolver esse tipo de problema é mudar para uma base para R^n formada por autovetores de A.

EXERCÍCIOS

1. Para um dos itens a seguir, encontre a matriz que corresponde à mudança da base $[\mathbf{u}_1, \mathbf{u}_2]$ para a base $[\mathbf{e}_1, \mathbf{e}_2]$.

(a) $\mathbf{u}_1 = (1, 1)^T, \quad \mathbf{u}_2 = (-1, 1)^T$
(b) $\mathbf{u}_1 = (1, 2)^T, \quad \mathbf{u}_2 = (2, 5)^T$
(c) $\mathbf{u}_1 = (0, 1)^T, \quad \mathbf{u}_2 = (1, 0)^T$

2. Para cada uma das bases ordenadas $[\mathbf{u}_1, \mathbf{u}_2]$ no Exercício 1, encontre a matriz mudança de base de $[\mathbf{e}_1, \mathbf{e}_2]$ para $[\mathbf{u}_1, \mathbf{u}_2]$.

3. Sejam $\mathbf{v}_1 = (3, 2)^T$ e $\mathbf{v}_2 = (4, 3)^T$. Para cada uma das bases ordenadas $[\mathbf{u}_1, \mathbf{u}_2]$ no Exercício 1, encontre a matriz mudança de base de $[\mathbf{v}_1, \mathbf{v}_2]$ para $[\mathbf{u}_1, \mathbf{u}_2]$.

4. Seja $E = [(5, 3)^T, (3, 2)^T]$ e sejam $\mathbf{x} = (1, 1)^T$, $\mathbf{y} = (1, -1)^T$ e $\mathbf{z} = (10, 7)^T$. Encontre os vetores de coordenadas $[\mathbf{x}]_E$, $[\mathbf{y}]_E$ e $[\mathbf{z}]_E$.

5. Sejam $\mathbf{u}_1 = (1, 1, 1)^T$, $\mathbf{u}_2 = (1, 2, 2)^T$ e $\mathbf{u}_3 = (2, 3, 4)^T$.
(a) Encontre a matriz mudança de base de $[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$ para $[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3]$.
(b) Encontre as coordenadas de cada um dos vetores a seguir em relação a $[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3]$.

$$\text{(i) } (3, 2, 5)^T \qquad \text{(ii) } (1, 1, 2)^T \qquad \text{(iii) } (2, 3, 2)^T$$

6. Sejam $\mathbf{v}_1 = (4, 6, 7)^T$, $\mathbf{v}_2 = (0, 1, 1)^T$ e $\mathbf{v}_3 = (0, 1, 2)^T$ e sejam $\mathbf{u}_1$, $\mathbf{u}_2$ e $\mathbf{u}_3$ os vetores dados no Exercício 5.
(a) Encontre a matriz mudança de base de $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$ para $[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3]$.
(b) Se $\mathbf{x} = 2\mathbf{v}_1 + 3\mathbf{v}_2 - 4\mathbf{v}_3$, determine as coordenadas de $\mathbf{x}$ em relação a $[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3]$.

7. Considere

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \qquad S = \begin{pmatrix} 3 & 5 \\ 1 & -2 \end{pmatrix}$$

Encontre vetores $\mathbf{w}_1$ e $\mathbf{w}_2$ tais que S é a matriz mudança de base de $[\mathbf{w}_1, \mathbf{w}_2]$ para $[\mathbf{v}_1, \mathbf{v}_2]$.

8. Considere

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ 6 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 4 \end{pmatrix}, \qquad S = \begin{pmatrix} 4 & 1 \\ 2 & 1 \end{pmatrix}$$

Encontre vetores $\mathbf{u}_1$ e $\mathbf{u}_2$ tais que S é a matriz mudança de base de $[\mathbf{v}_1, \mathbf{v}_2]$ para $[\mathbf{u}_1, \mathbf{u}_2]$.

9. Sejam $[x, 1]$ e $[2x - 1, 2x + 1]$ duas bases ordenadas para P_2.
(a) Encontre a matriz mudança de base que representa a mudança de coordenadas de $[2x - 1, 2x + 1]$ para $[x, 1]$.
(b) Encontre a matriz mudança de base que representa a mudança de coordenadas de $[x, 1]$ para $[2x - 1, 2x + 1]$.

10. Encontre a matriz mudança de base que representa a mudança de coordenadas em P_3 da base ordenada $[1, x, x^2]$ para a base ordenada

$$[1, 1 + x, 1 + x + x^2]$$

11. Sejam $E = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$ e $F = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ duas bases ordenadas para R^n e defina

$$U = (\mathbf{u}_1, \ldots, \mathbf{u}_n), \qquad V = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$$

Mostre que a matriz mudança de base de E para F pode ser determinada calculando-se a forma escada reduzida por linhas de $(U|V)$.

6 ESPAÇOS LINHA E COLUNA

Se A é uma matriz $m \times n$, cada linha de A é uma n-upla de números reais e pode ser considerada, portanto, como um vetor em $R^{1\times n}$. Vamos nos referir aos m vetores correspondentes às linhas de A como os *vetores linhas* de A. Analogamente, cada coluna de A pode ser considerada como um vetor em R^m e podemos associar à matriz A n *vetores colunas*.

Definição. Se A é uma matriz $m \times n$, o subespaço de $R^{1\times n}$ gerado pelos vetores linhas de A é chamado de **espaço linha** de A. O subespaço de R^m gerado pelos vetores colunas de A é chamado de **espaço coluna** de A.

EXEMPLO 1. Seja

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

O espaço linha de A é o conjunto de todas as triplas da forma

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

O espaço coluna de A é o conjunto de todos os vetores da forma

$$\alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

Portanto, o espaço linha de A é um subespaço bidimensional de $R^{1\times 3}$ e o espaço coluna de A é R^2. □

Teorema 3.6.1. *Duas matrizes equivalentes por linhas têm o mesmo espaço linha.*

Demonstração. Se B é equivalente por linhas a A, então B pode ser formada por uma seqüência finita de operações sobre as linhas de A. Logo, as linhas de B são combinações lineares dos vetores linhas de A. Conseqüentemente, o espaço linha de B tem que ser um subespaço do espaço linha de A. Como A é equivalente por linhas a B, o espaço linha de A tem que ser um subespaço do espaço linha de B pelo mesmo argumento. □

Definição. O **posto** de uma matriz A é a dimensão de seu espaço linha.

Para determinar o posto de uma matriz, podemos colocá-la em forma escada. As linhas não-nulas da forma escada formam uma base para o espaço linha.

EXEMPLO 2. Seja

$$A = \begin{pmatrix} 1 & -2 & 3 \\ 2 & -5 & 1 \\ 1 & -4 & -7 \end{pmatrix}$$

Reduzindo A à sua forma escada, obtemos

$$U = \begin{pmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{pmatrix}$$

É claro que $(1, -2, 3)$ e $(0, 1, 5)$ formam uma base para o espaço linha de U. Como as matrizes A e U são equivalentes por linhas, elas têm o mesmo espaço linha e, portanto, o posto de A é 2. □

SISTEMAS LINEARES

Os conceitos de espaços linha e coluna são úteis no estudo de sistemas lineares. O sistema $A\mathbf{x} = \mathbf{b}$ pode ser escrito na forma

$$\text{(1)} \qquad x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Segue de (1) que o sistema $A\mathbf{x} = \mathbf{b}$ é compatível se e somente se $\mathbf{b}$ pode ser escrito como uma combinação linear dos vetores colunas de A. Temos, então, a seguinte caracterização de sistemas compatíveis:

$$\boxed{A\mathbf{x} = \mathbf{b} \text{ é compatível se e somente se } \mathbf{b} \text{ pertence ao espaço coluna de } A.}$$

Se $\mathbf{b}$ for substituído pelo vetor nulo, (1) fica

(2) $$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{0}$$

Então, de (2), o sistema $A\mathbf{x} = \mathbf{0}$ tem apenas a solução trivial $\mathbf{x} = \mathbf{0}$ se e somente se os vetores colunas de A são linearmente independentes.

Teorema 3.6.2. *Seja A uma matriz $m \times n$. O sistema linear $A\mathbf{x} = \mathbf{b}$ é compatível para todo $\mathbf{b} \in R^m$ se e somente se os vetores colunas de A geram R^m. O sistema $A\mathbf{x} = \mathbf{b}$ tem no máximo uma solução qualquer que seja o vetor $\mathbf{b} \in R^m$ se e somente se os vetores colunas de A são linearmente independentes.*

Demonstração. Vimos que o sistema $A\mathbf{x} = \mathbf{b}$ é compatível se e somente se $\mathbf{b}$ pertence ao espaço coluna de A. Temos, então, que $A\mathbf{x} = \mathbf{b}$ é compatível para todo $\mathbf{b} \in R^m$ se e somente se os vetores colunas de A geram R^m. Para provar a segunda afirmação, observe que, se $A\mathbf{x} = \mathbf{b}$ tem no máximo uma solução qualquer que seja o vetor $\mathbf{b}$, então, em particular, $A\mathbf{x} = \mathbf{0}$ tem apenas a solução trivial e, portanto, os vetores colunas de A são linearmente independentes. Por outro lado, se os vetores colunas de A são linearmente independentes, $A\mathbf{x} = \mathbf{0}$ tem apenas a solução trivial. Mas então, se $\mathbf{x}_1$ e $\mathbf{x}_2$ são duas soluções de $A\mathbf{x} = \mathbf{b}$, $\mathbf{x}_1 - \mathbf{x}_2$ é uma solução de $A\mathbf{x} = \mathbf{0}$,

$$A(\mathbf{x}_1 - \mathbf{x}_2) = A\mathbf{x}_1 - A\mathbf{x}_2 = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

Logo, $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$ e, portanto, $\mathbf{x}_1$ tem que ser igual a $\mathbf{x}_2$. □

Seja A uma matriz $m \times n$. Se os vetores colunas de A geram R^m, então n tem que ser maior ou igual a m, já que um conjunto com menos que m vetores não pode gerar R^m. Se as colunas de A são linearmente independentes, então n tem que ser menor ou igual a m, já que qualquer conjunto com mais de m vetores em R^m é linearmente dependente. Portanto, se os vetores colunas de A formarem uma base para R^m, então n tem que ser igual a m.

Corolário 3.6.3. *Uma matriz A $n \times n$ é invertível se e somente se os vetores colunas de A formam uma base para R^n.*

Em geral, a soma do posto e da dimensão do núcleo é igual ao número de colunas da matriz. A dimensão do núcleo de uma matriz é chamada a *nulidade* da matriz.

Teorema 3.6.4. *Se A é uma matriz $m \times n$, então a soma do posto de A com a nulidade de A é igual a n.*

Demonstração. Seja U a forma escada reduzida por linhas de A. O sistema $A\mathbf{x} = \mathbf{0}$ é equivalente ao sistema $U\mathbf{x} = \mathbf{0}$. Se A tem posto r, então U tem r linhas não-nulas e, portanto, o sistema $U\mathbf{x} = \mathbf{0}$ tem r variáveis líderes e $n - r$ variáveis livres. A dimensão de $N(A)$ é igual ao número de variáveis livres. □

EXEMPLO 3. Seja

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

Encontre uma base para o espaço linha de A e uma base para $N(A)$. Verifique que $\dim N(A) = n - r$.

SOLUÇÃO. A forma escada reduzida por linhas de A é dada por

$$U = \begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Logo, $\{(1, 2, 0, 3), (0, 0, 1, 2)\}$ é uma base para o espaço linha de A e A tem posto 2. Como os sistemas $A\mathbf{x} = \mathbf{0}$ e $U\mathbf{x} = \mathbf{0}$ são equivalentes, $\mathbf{x}$ pertence a $N(A)$ se e somente se

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

Podemos resolver para as variáveis líderes x_1 e x_3 em termos das variáveis livres x_2 e x_4.

$$\begin{aligned} x_1 &= -2x_2 - 3x_4 \\ x_3 &= -2x_4 \end{aligned}$$

Fazendo $x_2 = \alpha$ e $x_4 = \beta$, vemos que $N(A)$ é formado por todos os vetores da forma

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -2\alpha - 3\beta \\ \alpha \\ -2\beta \\ \beta \end{pmatrix} = \alpha \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -3 \\ 0 \\ -2 \\ 1 \end{pmatrix}$$

Os vetores $(-2, 1, 0, 0)^T$ e $(-3, 0, -2, 1)^T$ formam uma base para $N(A)$. Observe que

$$n - r = 4 - 2 = 2 = \dim N(A)$$ □

O ESPAÇO COLUNA

As matrizes A e U no Exemplo 3 têm espaços colunas diferentes; no entanto, seus vetores colunas satisfazem as mesmas relações de dependência. Para a matriz U, os vetores colunas $\mathbf{u}_1$ e $\mathbf{u}_3$ são linearmente independentes, enquanto

$$\begin{aligned} \mathbf{u}_2 &= 2\mathbf{u}_1 \\ \mathbf{u}_4 &= 3\mathbf{u}_1 + 2\mathbf{u}_3 \end{aligned}$$

As mesmas relações valem para as colunas de A. Os vetores $\mathbf{a}_1$ e $\mathbf{a}_3$ são linearmente independentes, enquanto

$$\begin{aligned} \mathbf{a}_2 &= 2\mathbf{a}_1 \\ \mathbf{a}_4 &= 3\mathbf{a}_1 + 2\mathbf{a}_3 \end{aligned}$$

Em geral, se A é uma matriz $m \times n$ e U é a forma escada reduzida por linhas de A, então, como $A\mathbf{x} = 0$ se e somente se $U\mathbf{x} = \mathbf{0}$, seus vetores colunas satisfazem as mesmas relações de dependência. Usaremos esse fato para provar que a dimensão do espaço coluna de A é igual à dimensão do espaço linha de A.

Teorema 3.6.5. *Se A é uma matriz $m \times n$, então a dimensão do espaço linha de A é igual à dimensão do espaço coluna de A.*

Demonstração. Se A é uma matriz $m \times n$ de posto r, a forma escada reduzida por linhas U de A tem r linhas cujos primeiros elementos não-nulos são iguais a 1; as colunas correspondentes a esses elementos são linearmente independentes. Elas não formam, no entanto, uma base para o espaço coluna de A, já que, em geral, A e U têm espaços colunas diferentes. Denote por U_L a matriz obtida de U retirando-se todas as colunas correspondentes às variáveis livres. Retire as mesmas colunas de A e denote por A_L a matriz assim obtida. As matrizes A_L e U_L são equivalentes por linhas. Então, se $\mathbf{x}$ é uma solução de $A_L\mathbf{x} = \mathbf{0}$, $\mathbf{x}$ também tem que ser solução de $U_L\mathbf{x} = \mathbf{0}$. Como as colunas de U_L são linearmente independentes, $\mathbf{x}$ tem que ser igual a 0. Das observações antecedendo o Teorema 3.6.2, tem-se que as colunas de A_L são linearmente independentes. Como A_L tem r colunas, a dimensão do espaço coluna de A é, pelo menos, r.

Acabamos de provar que, para qualquer matriz, a dimensão do espaço coluna é maior ou igual à dimensão do espaço linha. Aplicando esse resultado à matriz A^T, vemos que

$$\begin{aligned}\dim(\text{espaço linha de } A) &= \dim(\text{espaço coluna de } A^T)\\ &\geq \dim(\text{espaço linha de } A^T)\\ &= \dim(\text{espaço coluna de } A^T)\end{aligned}$$

Portanto, qualquer que seja a matriz A, a dimensão do espaço linha tem que ser igual à dimensão do espaço coluna. □

Podemos usar a forma escada reduzida por linhas U de A para encontrar uma base para o espaço coluna de A. Basta determinar as colunas de U que correspondem aos primeiros elementos não-nulos, iguais a 1, de cada linha. As colunas correspondentes de A são linearmente independentes e formam uma base para o espaço coluna de A.

Observação. A forma escada reduzida por linhas U nos diz quais colunas de A devem ser usadas para se obter uma base. Não podemos usar os vetores colunas de U já que, em geral, U e A têm espaços colunas diferentes.

EXEMPLO 4. Seja

$$A = \begin{pmatrix} 1 & -2 & 1 & 1 & 2 \\ -1 & 3 & 0 & 2 & -2 \\ 0 & 1 & 1 & 3 & 4 \\ 1 & 2 & 5 & 13 & 5 \end{pmatrix}$$

A forma escada reduzida por linhas de A é dada por

$$U = \begin{pmatrix} 1 & -2 & 1 & 1 & 2 \\ 0 & 1 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Os primeiros elementos não-nulos de cada linha ocorrem nas primeira, segunda e quinta colunas. Logo,

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} -2 \\ 3 \\ 1 \\ 2 \end{pmatrix}, \quad \mathbf{a}_5 = \begin{pmatrix} 2 \\ -2 \\ 4 \\ 5 \end{pmatrix}$$

formam uma base espaço coluna de A. □

EXEMPLO 5. Encontre a dimensão do subespaço de R^4 gerado por

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 0 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} 2 \\ 5 \\ -3 \\ 2 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} 2 \\ 4 \\ -2 \\ 0 \end{pmatrix}, \quad \mathbf{x}_4 = \begin{pmatrix} 3 \\ 8 \\ -5 \\ 4 \end{pmatrix}$$

SOLUÇÃO. O subespaço $[\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}]$ é igual ao espaço coluna da matriz

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

A forma escada reduzida por linhas de X é

$$\begin{pmatrix} 1 & 2 & 2 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

As duas primeiras colunas $\mathbf{x}_1$ e $\mathbf{x}_2$ de X formam uma base para o espaço coluna de X, logo dim $[\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}] = 2$. □

EXERCÍCIOS

1. Para cada uma das matrizes a seguir, encontre uma base para o espaço linha, uma base para o espaço coluna e uma base para o núcleo.

(a) $\begin{pmatrix} 1 & 3 & 2 \\ 2 & 1 & 4 \\ 4 & 7 & 8 \end{pmatrix}$ (b) $\begin{pmatrix} -3 & 1 & 3 & 4 \\ 1 & 2 & -1 & -2 \\ -3 & 8 & 4 & 2 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 3 & -2 & 1 \\ 2 & 1 & 3 & 2 \\ 3 & 4 & 5 & 6 \end{pmatrix}$

2. Em cada um dos itens a seguir, determine a dimensão do subespaço de R^3 gerado pelos vetores dados.

(a) $\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \\ 4 \end{pmatrix}, \begin{pmatrix} -3 \\ 3 \\ 6 \end{pmatrix}$ (b) $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$

(c) $\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 2 \\ -4 \end{pmatrix}, \begin{pmatrix} 3 \\ -2 \\ 5 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$

3. Seja

$$A = \begin{pmatrix} 1 & 2 & 2 & 3 & 1 & 4 \\ 2 & 4 & 5 & 5 & 4 & 9 \\ 3 & 6 & 7 & 8 & 5 & 9 \end{pmatrix}$$

(a) Calcule a forma escada reduzida por linhas U de A. Quais os vetores colunas de U que correspondem às variáveis livres? Escreva cada um desses vetores colunas como uma combinação linear dos vetores colunas correspondentes às variáveis líderes.

(b) Quais os vetores colunas de A que correspondem às variáveis líderes de U? Esses vetores colunas formam uma base para o espaço coluna de A. Escreva cada um dos vetores colunas de A como uma combinação linear dos vetores dessa base.

4. Para cada uma das escolhas de A e b a seguir, determine se $\mathbf{b}$ pertence ao espaço coluna de A e diga se o sistema $A\mathbf{x} = \mathbf{b}$ é ou não compatível.

(a) $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 4 \\ 8 \end{pmatrix}$

(b) $A = \begin{pmatrix} 3 & 6 \\ 1 & 2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

(c) $A = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$

(d) $A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 1 & 1 & 2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

(e) $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 2 \\ 5 \\ 2 \end{pmatrix}$

(f) $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 1 & 2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 5 \\ 10 \\ 5 \end{pmatrix}$

5. Para cada um dos sistemas compatíveis no Exercício 4, examine os vetores colunas da matriz de coeficientes para determinar se o sistema tem uma ou uma infinidade de soluções.

6. Quantas soluções os sistema $A\mathbf{x} = \mathbf{b}$ vai ter se $\mathbf{b}$ pertencer ao espaço coluna de A e se os vetores colunas de A forem linearmente independentes? Explique.]

7. Seja A uma matriz $m \times n$ com $m > n$. Seja $\mathbf{b} \in R^m$ e suponha que $N(A) = \{\mathbf{0}\}$.
(a) O que você pode concluir sobre os vetores colunas de A? Eles são linearmente independentes? Eles geram R^m? Explique.
(b) Quantas soluções o sistema $A\mathbf{x} = \mathbf{b}$ vai ter se $\mathbf{b}$ não pertencer ao espaço coluna de A? Quantas soluções o sistema vai ter se $\mathbf{b}$ pertencer ao espaço coluna de A? Explique.

8. Sejam A e B matrizes 6×5. Se $\dim N(A) = 2$, qual o posto de A? Se o posto de B for 4, qual vai ser a $\dim N(B)$?

9. Sejam A e B matrizes equivalentes por linhas.
(a) Mostre que a dimensão do espaço coluna de A é igual à dimensão do espaço coluna de B.
(b) Os espaços colunas de A e B são necessariamente iguais? Justifique sua resposta.

10. Prove que um sistema linear $A\mathbf{x} = \mathbf{b}$ é compatível se e somente se o posto de $(A|\mathbf{b})$ é igual ao posto de A.

11. Seja A uma matriz $m \times n$.
(a) Se B é uma matriz $m \times m$ invertível, mostre que BA e A têm o mesmo núcleo e, portanto, o mesmo posto.
(b) Se C é uma matriz $n \times n$ invertível, mostre que AC e A têm o mesmo posto.

12. Prove o Corolário 3.6.3.

13. Suponha que A e B são matrizes $n \times n$ com a propriedade de que $A\mathbf{x} = B\mathbf{x}$ para todo $\mathbf{x} \in R^n$. Mostre que:
(a) $N(A - B) = R^n$;
(b) $A - B$ têm que ter posto nulo e, portanto, $A = B$.

14. Sejam A e B matrizes $n \times n$. Mostre que $AB = O$ se e somente se o espaço coluna de B é um subespaço do núcleo de A.

15. Sejam $A \in R^{m\times n}$, $\mathbf{b} \in R^m$ e $\mathbf{x}_0$ uma solução particular do sistema $A\mathbf{x} = \mathbf{b}$. Prove as afirmações a seguir.
(a) Um vetor $\mathbf{y}$ em R^n é uma solução de $A\mathbf{x} = \mathbf{b}$ se e somente se $\mathbf{y} = \mathbf{x}_0 + \mathbf{z}$, onde $\mathbf{z} \in N(A)$.
(b) Se $N(A) = \{\mathbf{0}\}$, então a solução $\mathbf{x}_0$ é única.

16. Sejam $\mathbf{x}$ e $\mathbf{y}$ vetores não-nulos em R^m e R^n, respectivamente, e seja $A = \mathbf{x}\mathbf{y}^T$.
(a) Mostre que $\{\mathbf{x}\}$ é uma base para o espaço coluna de A e que $\{\mathbf{y}^T\}$ é uma base para o espaço linha de A.
(b) Qual a dimensão de $N(A)$?

17. Sejam $A \in R^{m\times n}$, $B \in R^{n\times r}$ e $C = AB$. Mostre que:
(a) O espaço coluna de C é um subespaço do espaço coluna de A;
(b) O espaço linha de C é um subespaço do espaço linha de B;
(c) $\text{Posto}(C) \leq \min\{\text{posto}(A), \text{posto}(B)\}$.

18. Sejam $A \in R^{m\times n}$, $B \in R^{n\times r}$ e $C = AB$. Mostre que:
(a) Se ambos A e B têm vetores colunas linearmente independentes, então os vetores colunas de C também são linearmente independentes.
(b) Se ambos A e B têm vetores linhas linearmente independentes, então os vetores linhas de C também são linearmente independentes.

[***Sugestão:*** aplique a parte (a) a C^T.]

19. Sejam $A \in R^{m\times n}$; $B \in R^{n\times r}$ e $C = AB$. Mostre que:
(a) Se os vetores colunas de B são linearmente dependentes, então os vetores colunas de C também são linearmente dependentes.
(b) Se os vetores linhas de A são linearmente dependentes, então os vetores linhas de C também são linearmente dependentes.

[***Sugestão:*** aplique a parte (a) a C^T.]

20. Dizemos que uma matriz A $m \times n$ tem uma *inversa à direita* se existe uma matriz C $n \times m$ tal que $AC = I_m$. Dizemos que A tem uma *inversa à esquerda* se existe uma matriz D $n \times m$ tal que $DA = I_n$.

(a) Mostre que, se A tem inversa à direita, então os vetores colunas de A geram R^m.

(b) É possível para uma matriz $m \times n$ ter uma inversa à direita se $n < m$? E se $n \geq m$? Explique.

21. Prove que, se A é uma matriz $m \times n$ tal que os vetores colunas de A geram R^m, então A tem uma inversa à direita.

[***Sugestão:*** denote por $\mathbf{e}_j$ a j-ésima coluna de I_m e resolva $A\mathbf{x} = \mathbf{e}_j$ para $j = 1, \ldots, m$.]

22. Mostre que uma matriz B tem inversa à esquerda se e somente se B^T tem inversa à direita.

23. Seja B uma matriz $n \times m$ cujas colunas são linearmente independentes. Mostre que B tem inversa à esquerda.

24. Prove que, se uma matriz B tem inversa à esquerda, então as colunas de B são linearmente independentes.

25. Se uma matriz U está em forma escada, então os vetores linhas não-nulos formam uma base para o espaço linha de U.

EXERCÍCIOS COM O MATLAB PARA O CAPÍTULO 3

1. (Mudança de Base) Defina

$$U = \texttt{round}(20 * (\texttt{rand}(4) - 0{,}5)), \qquad V = \texttt{round}(10 * \texttt{rand}(4))$$

e $\mathbf{b} = \texttt{ones}(4, 1)$.

(a) Podemos usar a função `rank` do MATLAB para determinar se os vetores colunas de uma matriz são ou não linearmente independentes. Quanto deveria ser o posto se os vetores colunas de U fossem linearmente independentes? Calcule o posto de U e verifique que seus vetores colunas são linearmente independentes e, portanto, formam uma base para R^4. Calcule o posto de V e verifique que seus vetores colunas também formam uma base para R^4.

(b) Use o MATLAB para calcular a matriz mudança de base da base canônica $[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4]$ para a base ordenada $E = [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4]$. [Lembre-se de que a notação do MATLAB para o j-ésimo vetor coluna $\mathbf{u}_j$ é $U(:, j)$.] Use essa matriz mudança de base para calcular o vetor de coordenadas $\boldsymbol{\alpha}$ de $\mathbf{b}$ em relação a E. Verifique que

$$\mathbf{b} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 + \alpha_4\mathbf{u}_4 = U\boldsymbol{\alpha}$$

(c) Use o MATLAB para calcular a matriz mudança de base da base canônica para a base ordenada $F = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4]$ e use essa matriz mudança de base para calcular o vetor de coordenadas $\boldsymbol{\beta}$ de $\mathbf{b}$ em relação a F. Verifique que

$$\mathbf{b} = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \beta_3\mathbf{v}_3 + \beta_4\mathbf{v}_4 = V\boldsymbol{\beta}$$

(d) Use o MATLAB para calcular a matriz mudança de base S de E para F e a matriz mudança de base T de F para E. Qual a relação entre S e T? Verifique que $S\boldsymbol{\alpha} = \boldsymbol{\beta}$ e que $T\boldsymbol{\beta} = \boldsymbol{\alpha}$.

2. (Matrizes de Posto Incompleto) Neste exercício vamos ver como gerar matrizes de determinado posto no MATLAB.

(a) Em geral, se A é uma matriz $m \times n$ de posto r, então $r \leq \min(m, n)$. Por quê? Explique. Se os elementos de A forem números aleatórios, esperaríamos que $r = \min(m, n)$. Por quê? Explique. Verifique isso no MATLAB gerando matrizes aleatórias 6×6, 8×6, 5×8 e verificando seus postos com o comando `rank`. Sempre que o posto de uma matriz $m \times n$ for igual a $\min(m, n)$, diremos que a matriz tem *posto máximo*. Caso contrário, a matriz não tem posto máximo.

(b) Podemos gerar uma matriz aleatória com elementos inteiros multiplicando a matriz por um número $x \geq 10$ e, depois, usando o comando `round`. Por exemplo, o comando

$$A = \texttt{round}(10 * \texttt{rand}(10, 7))$$

gera uma matriz aleatória 10×7 cujos elementos são todos inteiros não-negativos menores ou iguais a 10. Gere matrizes aleatórias 10×7, 8×12, 10×15 dessa maneira e verifique o posto de cada uma delas. Essas matrizes inteiras têm posto máximo?

(c) Suponha que queremos gerar matrizes no MATLAB que não tenham posto máximo. É fácil gerar matrizes de posto 1. Se **x** e **y** são vetores não-nulos em R^m e R^n, respectivamente, então $A = \mathbf{x}\mathbf{y}^T$ é uma matriz $m \times n$ de posto 1. Por quê? Explique. Verifique isso no MATLAB definindo

$$\mathbf{x} = \texttt{round}(10 * \texttt{rand}(8, 1)), \qquad \mathbf{y} = \texttt{round}(10 * \texttt{rand}(6, 1))$$

e usando esses vetores para construir uma matriz A 8×6. Verifique o posto de A com o comando `rank`.

(d) Em geral,

$$\text{posto}(AB) \leq \min(\text{posto}(A), \text{posto}(B)) \tag{1}$$

(Ver Exercício 17 da Seção 6 deste capítulo.) Se A e B forem geradas aleatoriamente, a relação (1) deveria ser uma igualdade. Gere uma matriz A 8×6 definindo

$$X = \texttt{rand}(8, 2), \qquad Y = \texttt{rand}(6, 2), \qquad A = X * Y'$$

Quanto você esperaria que fosse o posto de A? Explique. Teste o posto de A usando o MATLAB.

(e) Use o MATLAB para gerar matrizes A, B e C tais que
   (i) A é 8×8 de posto 3;
   (ii) B é 6×9 de posto 4;
   (iii) C é 10×7 de posto 5.

3. (Espaço Coluna e Forma Escada Reduzida por Linhas)
Faça $B = \texttt{round}(10 * \texttt{rand}(8, 4))$, $X = \texttt{round}(10 * \texttt{rand}(4))$, $C = B * X$ e $A = [B \;\; C]$.

(a) Qual a relação entre os espaços colunas de B e C? (Ver Exercício 17 na Seção 6 deste capítulo.) Quanto você esperaria que fosse o posto de A? Explique. Use o MATLAB para verificar sua resposta.

(b) Quais os vetores colunas de A que deveriam formar uma base para seu espaço coluna? Explique. Se U é a forma escada reduzida por linhas de A, quais deveriam ser as quatro primeiras colunas de U? Explique. E as quatro últimas colunas? Explique. Use o MATLAB para verificar sua resposta calculando U.

(c) Use o MATLAB para construir outra matriz $D = (E \;\; EY)$, onde E é uma matriz aleatória 6×4 e Y é uma matriz aleatória 4×2. Qual deveria ser a forma escada reduzida por linhas de D? Calcule-a usando o MATLAB. Mostre que, em geral, se B é uma matriz $m \times n$ de posto n e se X é uma matriz $n \times k$, a forma escada reduzida por linhas de $(B \;\; BX)$ tem estrutura em bloco da forma

$$(I \quad X) \qquad \text{se} \qquad m = n$$

ou

$$\begin{pmatrix} I & X \\ O & O \end{pmatrix} \qquad \text{se} \qquad m > n$$

4. (Redução de Posto de Sistemas Lineares)

(a) Faça $A = \texttt{round}\,(10 * \texttt{rand}(8))$, $\mathbf{b} = \texttt{round}(10 * \texttt{rand}(8, 1))$ e $M = \textbf{inv}(A)$. Use a matriz M para resolver o sistema $A\mathbf{y} = \mathbf{b}$ para $\mathbf{y}$.

(b) Considere um novo sistema $C\mathbf{x} = \mathbf{b}$, onde C é construída da seguinte maneira:

$$\begin{array}{ll} \mathbf{u} = \texttt{round}(10 * \texttt{rand}(8, 1)), & \mathbf{v} = \texttt{round}(10 * \texttt{rand}(8, 1)) \\ E = \mathbf{u} * \mathbf{v}' & C = A + E \end{array}$$

As matrizes C e A diferem por uma matriz E de posto 1. Use o MATLAB para verificar que a matriz E tem posto 1. A seguir, use o comando `flops`(0) para igualar a 0 a variável flops e resolva o sistema $C\mathbf{x} = \mathbf{b}$ usando o operador "\". Verifique o valor de flops para ver quantas operações foram necessárias.

(c) Vamos agora resolver o sistema $C\mathbf{x} = \mathbf{b}$ por um método que usa o fato de que A e C diferem por uma matriz de posto 1. Esse novo procedimento é chamado de método de *redução de posto*. Iguale novamente a variável flops a 0 e faça

$$\mathbf{z} = M * \mathbf{u}, \quad \alpha = \mathbf{v}' * \mathbf{y}, \quad \beta = \mathbf{v}' * \mathbf{z}, \quad \gamma = \frac{\alpha}{1+\beta}$$

A solução $\mathbf{x}$ é dada por

$$\mathbf{x} = \mathbf{y} - \gamma * \mathbf{z}$$

Verifique o contador flops para o cálculo de $\mathbf{x}$ por esse segundo método. Esse método é mais eficiente do que usando o operador "\"? Use o MATLAB para calcular o vetor residual $\mathbf{b} - C\mathbf{x}$.

(d) Para ver por que o método de redução de posto funciona, use o MATLAB para calcular e comparar

$$C\mathbf{y} \text{ e } \mathbf{b} + \alpha\mathbf{u}$$

Prove que, se todos os cálculos tivessem sido feitos em aritmética exata, esses dois vetores seriam iguais. Calcule, também,

$$C\mathbf{z} \quad \text{e} \quad (1+\beta)\mathbf{u}$$

Prove que, se todos os cálculos tivessem sido feitos em aritmética exata, esses dois vetores seriam iguais. Use essas identidades para provar que $C\mathbf{x} = \mathbf{b}$. Supondo A invertível, o método de redução de posto sempre funciona? Sob que condições ele pode falhar? Explique.

CAPÍTULO 4

TRANSFORMAÇÕES LINEARES

Transformações lineares de um espaço vetorial em outro têm um papel importante em matemática. Neste capítulo, faz-se uma introdução à teoria de tais aplicações. A Seção 1 dá a definição de transformação linear e apresenta alguns exemplos. A Seção 2 mostra que toda transformação linear L de um espaço vetorial V de dimensão n em um espaço vetorial W de dimensão m pode ser representada por uma matriz A $m \times n$. Podemos, portanto, trabalhar com a matriz A no lugar do operador L. No caso em que a transformação linear L vai de V nele mesmo, a matriz que representa L depende da base ordenada escolhida para V. Então, L pode ser representada por uma matriz A em relação a uma base ordenada e por outra matriz B em relação a outra base ordenada. Na Seção 3, vamos considerar a relação entre as diferentes matrizes que representam o mesmo operador linear. Em muitas aplicações é desejável escolher uma base para V de modo que a matriz que representa a transformação linear seja diagonal ou tenha alguma outra forma simples.

1 DEFINIÇÃO E EXEMPLOS

No estudo de espaços vetoriais, as funções mais importantes são as transformações lineares.

Definição. Uma função L de um espaço vetorial V em um espaço vetorial W é chamada de **transformação linear** ou **operador linear** se

(1) $$L(\alpha \mathbf{v}_1 + \beta \mathbf{v}_2) = \alpha L(\mathbf{v}_1) + \beta L(\mathbf{v}_2)$$

para todos os vetores $\mathbf{v}_1, \mathbf{v}_2 \in V$ e para todos os escalares α e β.

Se L é uma transformação linear de V em W, então, de (1), tem-se que

(2) $$L(\mathbf{v}_1 + \mathbf{v}_2) = L(\mathbf{v}_1) + L(\mathbf{v}_2) \qquad (\alpha = \beta = 1)$$

e

(3) $$L(\alpha \mathbf{v}) = \alpha L(\mathbf{v}) \qquad (\mathbf{v} = \mathbf{v}_1, \beta = 0)$$

Por outro lado, se L satisfaz (2) e (3), então

$$\begin{aligned} L(\alpha \mathbf{v}_1 + \beta \mathbf{v}_2) &= L(\alpha \mathbf{v}_1) + L(\beta \mathbf{v}_2) \\ &= \alpha L(\mathbf{v}_1) + \beta L(\mathbf{v}_2) \end{aligned}$$

Portanto, L é um operador linear se e somente se satisfaz (2) e (3).

Notação. Uma aplicação L de um espaço vetorial V em um espaço vetorial W será denotada por

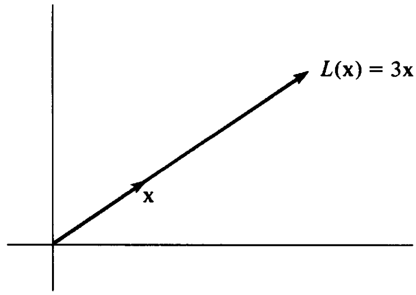


FIG. 4.1.1

$$L: V \to W$$

Ao usarmos essa notação com seta, estaremos supondo que V e W são espaços vetoriais.

Vamos considerar agora alguns exemplos de transformações lineares. Começaremos com operadores lineares de R^2 em si mesmo. Nesse caso é fácil ver o efeito geométrico do operador.

OPERADORES LINEARES EM R^2

EXEMPLO 1. Seja L o operador definido por

$$L(\mathbf{x}) = 3\mathbf{x}$$

para todo $\mathbf{x} \in R^2$. Como

$$L(\alpha\mathbf{x}) = 3(\alpha\mathbf{x}) = \alpha(3\mathbf{x}) = \alpha L(\mathbf{x})$$

e

$$L(\mathbf{x} + \mathbf{y}) = 3(\mathbf{x} + \mathbf{y}) = (3\mathbf{x}) + (3\mathbf{y}) = L(\mathbf{x}) + L(\mathbf{y})$$

L é uma transformação linear. Podemos pensar em L como "esticando" por um fator de 3 (ver Fig. 4.1.1). Em geral, se α é um escalar positivo, a transformação linear $F(\mathbf{x}) = \alpha\mathbf{x}$ representa uma dilatação ou uma compressão por um fator α. □

EXEMPLO 2. Considere a aplicação L definida por

$$L(\mathbf{x}) = x_1\mathbf{e}_1$$

para $\mathbf{x} \in R^2$. Se $\mathbf{x} = (x_1, x_2)^T$, temos $L(\mathbf{x}) = (x_1, 0)^T$. Se $\mathbf{y} = (y_1, y_2)^T$, temos

$$\alpha\mathbf{x} + \beta\mathbf{y} = \begin{pmatrix} \alpha x_1 + \beta y_1 \\ \alpha x_2 + \beta y_2 \end{pmatrix}$$

e, portanto,

$$\begin{aligned} L(\alpha\mathbf{x} + \beta\mathbf{y}) &= (\alpha x_1 + \beta y_1)\mathbf{e}_1 \\ &= \alpha(x_1\mathbf{e}_1) + \beta(y_1\mathbf{e}_1) \\ &= \alpha L(\mathbf{x}) + \beta L(\mathbf{y}) \end{aligned}$$

Logo, L é uma transformação linear. Geometricamente, L é a projeção sobre o eixo dos x_1 (ver Fig. 4.1.2). □

EXEMPLO 3. Seja L o operador definido por

$$L(\mathbf{x}) = (x_1, -x_2)^T$$

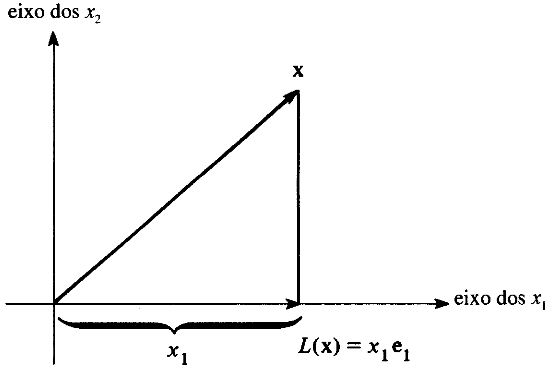


FIG. 4.1.2

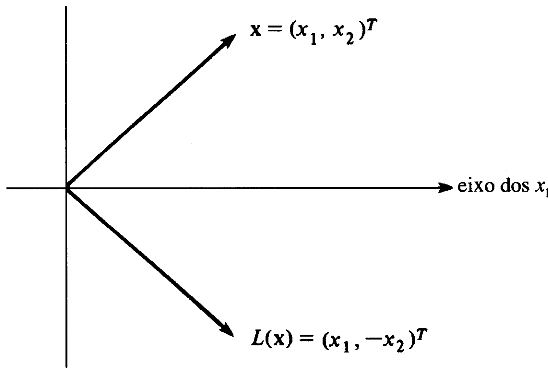


FIG. 4.1.3

para todo $\mathbf{x} = (x_1, x_2)^T$ em R^2. Como

$$
\begin{aligned}
L(\alpha\mathbf{x} + \beta\mathbf{y}) &= \begin{pmatrix} \alpha x_1 + \beta y_1 \\ -(\alpha x_2 + \beta y_2) \end{pmatrix} \\
&= \alpha \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} + \beta \begin{pmatrix} y_1 \\ -y_2 \end{pmatrix} \\
&= \alpha L(\mathbf{x}) + \beta L(\mathbf{y})
\end{aligned}
$$

L é uma transformação linear. O operador L reflete um vetor $\mathbf{x}$ em relação ao eixo dos x_1 (ver Fig. 4.1.3). □

EXEMPLO 4. O operador L definido por

$$L(\mathbf{x}) = (-x_2, x_1)^T$$

é linear, já que

$$
\begin{aligned}
L(\alpha\mathbf{x} + \beta\mathbf{y}) &= \begin{pmatrix} -(\alpha x_2 + \beta y_2) \\ \alpha x_1 + \beta y_1 \end{pmatrix} \\
&= \alpha \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} + \beta \begin{pmatrix} -y_2 \\ y_1 \end{pmatrix} \\
&= \alpha L(\mathbf{x}) + \beta L(\mathbf{y})
\end{aligned}
$$

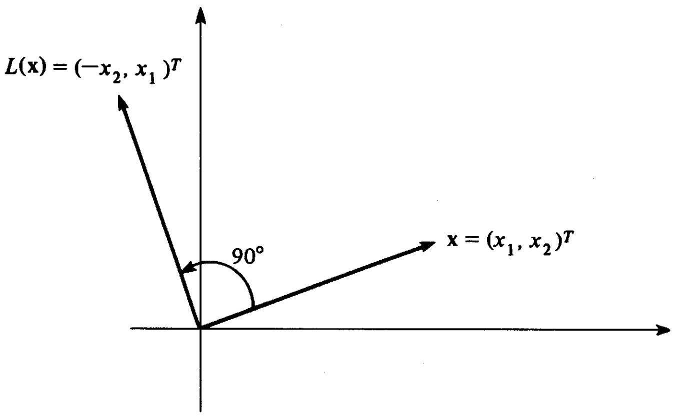


FIG. 4.1.4

O operador L roda cada vetor em R^2 de 90° em torno da origem, no sentido trigonométrico (ver Fig. 4.1.4). □

EXEMPLO 5. Considere a aplicação M definida por

$$M(\mathbf{x}) = (x_1^2 + x_2^2)^{1/2}$$

Como

$$M(\alpha\mathbf{x}) = (\alpha^2 x_1^2 + \alpha^2 x_2^2)^{1/2} = |\alpha| M(\mathbf{x})$$

tem-se que

$$\alpha M(\mathbf{x}) \neq M(\alpha\mathbf{x})$$

sempre que $\alpha < 0$ e $\mathbf{x} \neq \mathbf{0}$. Portanto, M não é uma transformação linear. □

OPERADORES LINEARES DE R^n EM R^m

EXEMPLO 6. A função L: $R^2 \to R^1$ definida por

$$L(\mathbf{x}) = x_1 + x_2$$

é uma transformação linear, já que

$$\begin{aligned} L(\alpha\mathbf{x} + \beta\mathbf{y}) &= (\alpha x_1 + \beta y_1) + (\alpha x_2 + \beta y_2) \\ &= \alpha(x_1 + x_2) + \beta(y_1 + y_2) \\ &= \alpha L(\mathbf{x}) + \beta L(\mathbf{y}) \end{aligned}$$

□

EXEMPLO 7. A aplicação L de R^2 em R^3 definida por

$$L(\mathbf{x}) = (x_2, x_1, x_1 + x_2)^T$$

é linear, pois

$$L(\alpha\mathbf{x}) = (\alpha x_2, \alpha x_1, \alpha x_1 + \alpha x_2)^T = \alpha L(\mathbf{x})$$

e

$$\begin{aligned} L(\mathbf{x} + \mathbf{y}) &= (x_2 + y_2, x_1 + y_1, x_1 + y_1 + x_2 + y_2)^T \\ &= (x_2, x_1, x_1 + x_2)^T + (y_2, y_1, y_1 + y_2)^T \\ &= L(\mathbf{x}) + L(\mathbf{y}) \end{aligned}$$

Observe que, se definirmos a matriz A por

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}$$

então

$$L(\mathbf{x}) = \begin{pmatrix} x_2 \\ x_1 \\ x_1 + x_2 \end{pmatrix} = A\mathbf{x}$$

para todo $\mathbf{x} \in R^2$. □

Em geral, se A é uma matriz $m \times n$, podemos definir um operador linear L_A de R^n em R^m por

$$L_A(\mathbf{x}) = A\mathbf{x}$$

para todo $\mathbf{x} \in R^n$. O operador L_A é linear, pois

$$\begin{aligned} L_A(\alpha\mathbf{x} + \beta\mathbf{y}) &= A(\alpha\mathbf{x} + \beta\mathbf{y}) \\ &= \alpha A\mathbf{x} + \beta A\mathbf{y} \\ &= \alpha L_A(\mathbf{x}) + \beta L_A(\mathbf{y}) \end{aligned}$$

Podemos, então, considerar cada matriz A $m \times n$ como um operador linear de R^n em R^m.

Vimos, no Exemplo 7, que o operador L poderia ter sido definido em termos de uma matriz A. Na próxima seção veremos que isso é verdade para todos os operadores lineares de R^n em R^m.

TRANSFORMAÇÕES LINEARES DE V EM W

Se L é um operador linear de um espaço vetorial V em um espaço vetorial W, então

(i) $L(\mathbf{0}_V) = \mathbf{0}_W$ (onde $\mathbf{0}_V$ e $\mathbf{0}_W$ são os vetores nulos de V e W, respectivamente);

(ii) se $\mathbf{v}_1, \ldots, \mathbf{v}_n$ são elementos de V e se $\alpha_1, \ldots, \alpha_n$ são escalares, temos

$$L(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n) = \alpha_1 L(\mathbf{v}_1) + \alpha_2 L(\mathbf{v}_2) + \cdots + \alpha_n L(\mathbf{v}_n)$$

(iii) $L(-\mathbf{v}) = -L(\mathbf{v})$ para todo $\mathbf{v} \in V$.

A afirmação (i) segue da condição $L(\alpha\mathbf{v}) = \alpha L(\mathbf{v})$ com $\alpha = 0$. A afirmação (ii) pode ser provada facilmente usando indução matemática. Deixamos essa demonstração a cargo do leitor. Para provar (iii), observe que

$$\mathbf{0}_W = L(\mathbf{0}_V) = L(\mathbf{v} + (-\mathbf{v})) = L(\mathbf{v}) + L(-\mathbf{v})$$

Portanto, $L(-\mathbf{v})$ é o inverso aditivo de $L(\mathbf{v})$, isto é,

$$L(-\mathbf{v}) = -L(\mathbf{v})$$

EXEMPLO 8. Se V é um espaço vetorial, então o operador identidade $\mathcal{I}$ é definido pela fórmula

$$\mathcal{I}(\mathbf{v}) = \mathbf{v}$$

para todo $\mathbf{v} \in V$. É claro que $\mathcal{I}$ é uma transformação linear de V em si mesmo.

$$\mathcal{I}(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2 = \alpha\mathcal{I}(\mathbf{v}_1) + \beta\mathcal{I}(\mathbf{v}_2)$$ □

EXEMPLO 9. Seja L a aplicação de $C[a, b]$ em R^1 definida por

$$L(f) = \int_a^b f(x)\,dx$$

Se f e g são vetores quaisquer em $C[a, b]$, então

$$\begin{aligned} L(\alpha f + \beta g) &= \int_a^b (\alpha f + \beta g)(x)\,dx \\ &= \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx \\ &= \alpha L(f) + \beta L(g) \end{aligned}$$

L é, portanto, uma transformação linear. □

EXEMPLO 10. Seja D o operador de $C^1[a, b]$ em $C[a, b]$ definido por

$$D(f) = f' \qquad \text{(a derivada de } f\text{)}$$

D é uma transformação linear, uma vez que

$$D(\alpha f + \beta g) = \alpha f' + \beta g' = \alpha D(f) + \beta D(g)$$ □

A IMAGEM E O NÚCLEO

Seja $L : V \to W$ uma transformação linear. Vamos terminar esta seção estudando o efeito de L nos subespaços de V. A coleção dos vetores de V que são transformados no vetor nulo de W é particularmente importante.

Definição. Seja $L : V \to W$ uma transformação linear. O **núcleo** de L, denotado por ker(L), é definido por

$$\ker(L) = \{\mathbf{v} \in V \,|\, L(\mathbf{v}) = \mathbf{0}_W\}$$

Definição. Seja $L : V \to W$ uma transformação linear e seja S um subespaço de V. A **imagem** de S, denotada por $L(S)$, é definida por

$$L(S) = \{\mathbf{w} \in W \,|\, \mathbf{w} = L(\mathbf{v}) \quad \text{para algum } \mathbf{v} \in S\}$$

A imagem de todo o espaço vetorial, $L(V)$, é chamada de **imagem** de L.

Seja $L : V \to W$ uma transformação linear. É fácil ver que ker(L) é um subespaço de W e que, se S é um subespaço qualquer de V, então $L(S)$ é um subespaço de W. Em particular, $L(V)$ é um subespaço de W. De fato, temos o seguinte teorema.

Teorema 4.1.1. *Se $L : V \to W$ é uma transformação linear e S é um subespaço de V, então:*

(i) ker(L) *é um subespaço de V;*
(ii) $L(S)$ *é um subespaço de W.*

Demonstração. Para provar (i), vamos mostrar que o núcleo de L é fechado sob as operações de soma e multiplicação por escalar. Se $\mathbf{v} \in$ ker(L) e α é um escalar, então

$$L(\alpha\mathbf{v}) = \alpha L(\mathbf{v}) = \alpha\mathbf{0}_W = \mathbf{0}_W$$

Portanto, $\alpha\mathbf{v} \in$ ker(L).

Se $\mathbf{v}_1, \mathbf{v}_2 \in$ ker(L), então

$$L(\mathbf{v}_1 + \mathbf{v}_2) = L(\mathbf{v}_1) + L(\mathbf{v}_2) = \mathbf{0}_W + \mathbf{0}_W = \mathbf{0}_W$$

Portanto, $\mathbf{v}_1 + \mathbf{v}_2 \in$ ker(L) e ker(L) é um subespaço de V.

A demonstração de (ii) é semelhante. Se $\mathbf{w} \in L(S)$, então $\mathbf{w} = L(\mathbf{v})$ para algum $\mathbf{v} \in S$. Para qualquer escalar α, temos

$$\alpha\mathbf{w} = \alpha L(\mathbf{v}) = L(\alpha\mathbf{v})$$

Como $\alpha\mathbf{v} \in S$, $\alpha\mathbf{w} \in L(S)$ e $L(S)$ é fechado sob a multiplicação por escalar. Se $\mathbf{w}_1, \mathbf{w}_2 \in L(S)$, existem $\mathbf{v}_1, \mathbf{v}_2 \in S$ tais que $L(\mathbf{v}_1) = \mathbf{w}_1$ e $L(\mathbf{v}_2) = \mathbf{w}_2$. Logo,

$$\mathbf{w}_1 + \mathbf{w}_2 = L(\mathbf{v}_1) + L(\mathbf{v}_2) = L(\mathbf{v}_1 + \mathbf{v}_2)$$

e $L(S)$ é fechado sob a soma. □

EXEMPLO 11. Seja L a transformação linear de R^2 em R^2 definida por

$$L(\mathbf{x}) = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$$

Um vetor $\mathbf{x}$ pertence ao núcleo de L se e somente se $x_1 = 0$. Logo, $\ker(L)$ é o subespaço unidimensional de R^2 gerado por $\mathbf{e}_2$. Um vetor $\mathbf{y}$ pertence à imagem de L se e somente se $\mathbf{y}$ é um múltiplo de $\mathbf{e}_1$. Logo, $L(R^2)$ é o subespaço unidimensional de R^2 gerado por $\mathbf{e}_1$. □

EXEMPLO 12. Seja $L : R^3 \to R^2$ a transformação linear definida por

$$L(\mathbf{x}) = (x_1 + x_2,\ x_2 + x_3)^T$$

e seja S o subespaço de R^3 gerado por $\mathbf{e}_1$ e $\mathbf{e}_3$.

Se $\mathbf{x} \in \ker(L)$, então

$$x_1 + x_2 = 0 \qquad \text{e} \qquad x_2 + x_3 = 0$$

Fazendo a variável livre $x_3 = a$, obtemos

$$x_2 = -a, \qquad x_1 = a$$

e, portanto, $\ker(L)$ é o subespaço unidimensional de R^3 de todos os vetores da forma $a(1,-1,1)^T$.

Se $\mathbf{x} \in S$, então $\mathbf{x}$ tem que ser da forma $(a, 0, b)^T$, logo $L(\mathbf{x}) = (a, b)^T$. É claro que $L(S) = R^2$. Como a imagem do subespaço S é o R^2 inteiro, a imagem de L tem que ser todo o R^2 [isto é, $L(R^3) = R^2$]. □

EXEMPLO 13. Seja D: $P_3 \to P_3$ o operador derivada, dado por

$$D(p(x)) = p'(x)$$

O núcleo de D consiste em todos os polinômios de grau 0. Logo, $\ker(D) = P_1$. Como a derivada de qualquer polinômio em P_3 é um polinômio em P_2, temos que $D(P_3) = P_2$. □

EXERCÍCIOS

1. Mostre que cada uma das aplicações seguintes é uma transformação linear de R^2 em R^2. Descreva geometricamente o que cada uma delas faz.

(a) $L(\mathbf{x}) = (-x_1, x_2)^T$ (b) $L(\mathbf{x}) = -\mathbf{x}$ (c) $L(\mathbf{x}) = (x_2, x_1)^T$
(d) $L(\mathbf{x}) = \frac{1}{2}\mathbf{x}$ (e) $L(\mathbf{x}) = x_2\mathbf{e}_2$

2. Seja L a transformação linear de R^2 em si mesmo definida por

$$L(\mathbf{x}) = (x_1 \cos\alpha - x_2 \operatorname{sen}\alpha,\ x_1 \operatorname{sen}\alpha + x_2 \cos\alpha)^T$$

Expresse x_1, x_2 e $L(\mathbf{x})$ em coordenadas polares. Descreva geometricamente o efeito dessa transformação linear.

3. Seja $\mathbf{a}$ um vetor fixo não-nulo em R^2. Uma aplicação da forma

$$L(\mathbf{x}) = \mathbf{x} + \mathbf{a}$$

é chamada de *translação*. Mostre que uma translação não é uma transformação linear. Ilustre geometricamente o efeito de uma translação.

4. Determine se as transformações de R^3 em R^2 a seguir são ou não lineares.

(a) $L(\mathbf{x}) = (x_2, x_3)^T$ (b) $L(\mathbf{x}) = (0, 0)^T$
(c) $L(\mathbf{x}) = (1 + x_1, x_2)^T$ (d) $L(\mathbf{x}) = (x_3, x_1 + x_2)^T$

5. Determine se as transformações de R^2 em R^3 a seguir são ou não lineares.
(a) $L(\mathbf{x}) = (x_1, x_2, 1)^T$ (b) $L(\mathbf{x}) = (x_1, x_2, x_1 + 2x_2)^T$
(c) $L(\mathbf{x}) = (x_1, 0, 0)^T$ (d) $L(\mathbf{x}) = (x_1, x_2, x_1^2 + x_2^2)^T$

6. Determine se as transformações de $R^{n\times n}$ em $R^{n\times n}$ a seguir são ou não lineares.
(a) $L(A) = 2A$ (b) $L(A) = A^T$
(c) $L(A) = A + I$ (d) $L(A) = A - A^T$

7. Determine se as transformações de P_2 em P_3 a seguir são ou não lineares.
(a) $L(p(x)) = xp(x)$
(b) $L(p(x)) = x^2 + p(x)$
(c) $L(p(x)) = p(x) + xp(x) + x^2p'(x)$

8. Para cada $f \in C[0, 1]$, defina $L(f) = F$, onde

$$F(x) = \int_0^x f(t)\,dt \qquad 0 \le x \le 1$$

Mostre que L é uma transformação linear de $C[0, 1]$ em $C[0, 1]$. Depois, encontre $L(e^x)$ e $L(x^2)$.

9. Determine se as transformações de $C[0, 1]$ em R^1 a seguir são ou não lineares.
(a) $L(f) = f(0)$ (b) $L(f) = |f(0)|$
(c) $L(f) = [f(0) + f(1)]/2$ (d) $L(f) = \left\{\int_0^1 [f(x)]^2\,dx\right\}^{1/2}$

10. Se L é uma transformação linear de V em W, use indução matemática para provar que

$$L(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n)$$
$$= \alpha_1 L(\mathbf{v}_1) + \alpha_2 L(\mathbf{v}_2) + \cdots + \alpha_n L(\mathbf{v}_n)$$

11. Seja $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ uma base para um espaço vetorial V e sejam L_1 e L_2 duas transformações lineares de V em um espaço vetorial W. Mostre que, se

$$L_1(\mathbf{v}_i) = L_2(\mathbf{v}_i)$$

para cada $i = 1, \ldots, n$, então $L_1 = L_2$ [isto é, mostre que $L_1(\mathbf{v}) = L_2(\mathbf{v})$ para todo $\mathbf{v} \in V$].

12. Seja L uma transformação linear de R^1 em R^1 e seja $a = L(1)$. Mostre que $L(x) = ax$ para todo $x \in R^1$.

13. Seja L um operador linear de um espaço vetorial V nele mesmo. Defina, por recursão, o operador L^n, $n \ge 1$ da seguinte maneira:

$$L^1 = L$$

$$L^{k+1}(\mathbf{v}) = L(L^k(\mathbf{v})) \quad \text{para todo } \mathbf{v} \in V$$

Mostre que L^n é um operador linear para todo $n \ge 1$.

14. Sejam $L_1: U \to V$ e $L_2: V \to W$ transformações lineares e seja $L = L_2 \circ L_1$ a transformação definida por

$$L(\mathbf{u}) = L_2(L_1(\mathbf{u}))$$

para $\mathbf{u} \in U$. Mostre que L é uma transformação linear de U em W.

15. Determine o núcleo e a imagem de cada uma das transformações lineares de R^3 em R^3.
(a) $L(\mathbf{x}) = (x_3, x_2, x_1)^T$
(b) $L(\mathbf{x}) = (x_1, x_2, 0)^T$
(c) $L(\mathbf{x}) = (x_1, x_1, x_1)^T$

16. Seja S o subespaço de R^3 gerado por $\mathbf{e}_1$ e $\mathbf{e}_2$. Para cada um dos operadores lineares no Exercício 15, determine $L(S)$.

17. Determine o núcleo e a imagem de cada uma das transformações lineares de P_3 em P_3 dadas a seguir.
(a) $L(p(x)) = xp'(x)$

(b) $L(p(x)) = p(x) - p'(x)$
(c) $L(p(x)) = p(0)x + p(1)$

18. Seja $L : V \to W$ um transformação linear e seja T um subespaço de W. A *imagem inversa* de T, denotada por $L^{-1}(T)$, é definida por

$$L^{-1}(T) = \{\mathbf{v} \in V \mid L(\mathbf{v}) \in T\}$$

Mostre que $L^{-1}(T)$ é um subespaço de V.

19. Uma transformação linear $L : V \to W$ é dita *injetora* se $L(\mathbf{v}_1) = L(\mathbf{v}_2)$ implica que $\mathbf{v}_1 = \mathbf{v}_2$ (isto é, dois vetores distintos $\mathbf{v}_1, \mathbf{v}_2 \in V$ não podem ser levados no mesmo vetor $\mathbf{w} \in W$). Mostre que L é injetora se e somente se $\ker(L) = \{\mathbf{0}_V\}$.

20. Um operador linear $L : V \to W$ é dito *sobrejetor* se $L(V) = W$. Mostre que o operador $L : R^3 \to R^3$ definido por

$$L(\mathbf{x}) = (x_1, x_1 + x_2, x_1 + x_2 + x_3)^T$$

é sobrejetor.

21. Quais dos operadores no Exercício 15 são injetores? Quais são sobrejetores?

22. Seja A uma matriz 2×2 e seja L_A o operador definido por

$$L_A(\mathbf{x}) = A\mathbf{x}$$

Mostre que:
(a) L_A leva R^2 no espaço coluna de A;
(b) se A é invertível, então L_A é sobrejetora de R^2 em R^2.

23. Seja D o operador derivada em P_3 e seja

$$S = \{p \in P_3 \mid p(0) = 0\}$$

Mostre que:
(a) D de P_3 em P_2 é sobrejetora, mas não é injetora;
(b) $D : S \to P_3$ é injetora, mas não é sobrejetora.

2 REPRESENTAÇÃO MATRICIAL DE TRANSFORMAÇÕES LINEARES

Na Seção 1, mostramos que cada matriz A $m \times n$ define uma transformação linear L_A de R^n em R^m dada por

$$L_A(\mathbf{x}) = A\mathbf{x}$$

para cada $\mathbf{x} \in R^n$. Nesta seção, mostraremos que a cada transformação linear L de R^n em R^m existe uma matriz A $m \times n$ tal que

$$L(\mathbf{x}) = A\mathbf{x}$$

Veremos, também, que qualquer operador linear agindo entre espaços vetoriais de dimensão finita pode ser representado por uma matriz.

Teorema 4.2.1 *Se L é uma transformação linear de R^n em R^m, então existe uma matriz A $m \times n$ tal que*

$$L(\mathbf{x}) = A\mathbf{x}$$

para cada $\mathbf{x} \in R^n$. De fato, o j-ésimo vetor coluna da matriz A é dado por

$$\mathbf{a}_j = L(\mathbf{e}_j) \qquad j = 1, 2, \ldots, n$$

Demonstração. Para $j = 1, \ldots, n$, defina

$$\mathbf{a}_j = (a_{1j}, a_{2j}, \ldots, a_{mj})^T = L(\mathbf{e}_j)$$

Seja

$$A = (a_{ij}) = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$$

Se

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$$

é um elemento arbitrário de R^n, temos

$$\begin{aligned} L(\mathbf{x}) &= x_1 L(\mathbf{e}_1) + x_2 L(\mathbf{e}_2) + \cdots + x_n L(\mathbf{e}_n) \\ &= x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n \\ &= (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &= A\mathbf{x} \end{aligned}$$

□

Mostramos que cada transformação linear de R^n em R^m pode ser representada em termos de uma matriz $m \times n$. O Teorema 4.2.1 mostra como construir a matriz A que representa um determinado operador linear L. Para obter a primeira coluna de A, aplique o operador linear L ao primeiro vetor da base de R^n, $\mathbf{e}_1$. Faça $\mathbf{a}_1 = L(\mathbf{e}_1)$. Para obter a segunda coluna de A, determine o efeito de L em $\mathbf{e}_2$ e faça $\mathbf{a}_2 = L(\mathbf{e}_2)$. E assim por diante. Como usamos os elementos da base canônica de R^n, $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_3$, dizemos que A é a

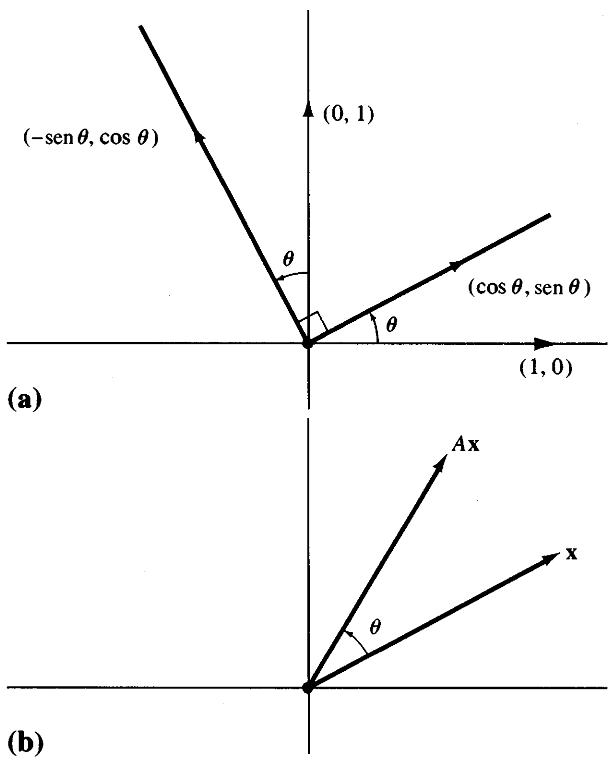


FIG. 4.2.1

representação canônica de L ou a *matriz de L em relação às bases canônicas*. Veremos, mais tarde, como representar um operador linear em relação a outras bases.

EXEMPLO 1. Defina o operador $L : R^3 \to R^2$ por

$$L(\mathbf{x}) = (x_1 + x_2, x_2 + x_3)^T$$

para cada $\mathbf{x} = (x_1, x_2, x_3)^T$ em R^3. É fácil verificar que L é um operador linear. Queremos encontrar uma matriz A tal que $L(\mathbf{x}) = A\mathbf{x}$ para todo $\mathbf{x} \in R^3$. Para fazer isso, precisamos encontrar $L(\mathbf{e}_1)$, $L(\mathbf{e}_2)$ e $L(\mathbf{e}_3)$.

$$L(\mathbf{e}_1) = L((1, 0, 0)^T) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$L(\mathbf{e}_2) = L((0, 1, 0)^T) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$L(\mathbf{e}_3) = L((0, 0, 1)^T) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Escolhemos esses vetores como as colunas da matriz A,

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Para verificar o resultado, calcule $A\mathbf{x}$.

$$A\mathbf{x} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_2 + x_3 \end{pmatrix}$$ □

EXEMPLO 2. Seja L a transformação de R^2 em R^2 que roda cada vetor de um ângulo θ no sentido trigonométrico. Podemos ver, na Fig. 4.2.1(a), que $\mathbf{e}_1$ é transformado em $(\cos\theta, \operatorname{sen}\theta)^T$ e que a imagem de $\mathbf{e}_2$ é $(-\operatorname{sen}\theta, \cos\theta)^T$. A matriz A que representa a transformação tem $(\cos\theta, \operatorname{sen}\theta)^T$ como primeira coluna e $(-\operatorname{sen}\theta, \cos\theta)^T$ como segunda coluna.

$$A = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

Se $\mathbf{x}$ é um vetor arbitrário em R^2, então, para rodar $\mathbf{x}$ em torno da origem de um ângulo θ no sentido trigonométrico, basta multiplicá-lo por A (ver Fig. 4.2.1(b)). □

Agora que vimos como usar matrizes para representar operadores lineares de R^n em R^m, podemos perguntar se é possível encontrar uma representação análoga para operadores lineares de V em W, onde V e W são espaços vetoriais de dimensão n e m, respectivamente. Para ver como fazer isso, sejam $E = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ e $F = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m]$ bases ordenadas para V e W, respectivamente. Seja L um operador linear de V em W. Se $\mathbf{v}$ é um vetor arbitrário em V, então podemos expressar $\mathbf{v}$ em termos da base E:

$$\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n$$

Vamos mostrar que existe uma matriz A $m \times n$ que representa o operador L no seguinte sentido:

$$A\mathbf{x} = \mathbf{y} \text{ se e somente se } L(\mathbf{v}) = y_1\mathbf{w}_1 + y_2\mathbf{w}_2 + \cdots + y_m\mathbf{w}_m$$

A matriz A caracteriza o efeito do operador L. Se $\mathbf{x}$ é o vetor de coordenadas de $\mathbf{v}$ em relação a E, então o vetor de coordenadas de $L(\mathbf{v})$ em relação a F é dado por

$$[L(\mathbf{v})]_F = A\mathbf{x}$$

O procedimento para determinar uma representação matricial A é essencialmente o mesmo que vimos. Para $j = 1, \ldots, n$, seja $\mathbf{a}_j = (a_{1j}, a_{2j}, \ldots, a_{mj})^T$ o vetor de coordenadas de $L(\mathbf{v}_j)$ em relação a $[\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m]$.

$$L(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + a_{2j}\mathbf{w}_2 + \cdots + a_{mj}\mathbf{w}_m \qquad 1 \le j \le n$$

Seja $A = (a_{ij}) = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$. Se

$$\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n$$

então

$$L(\mathbf{v}) = L\left(\sum_{j=1}^{n} x_j\mathbf{v}_j\right) = \sum_{j=1}^{n} x_j L(\mathbf{v}_j) = \sum_{j=1}^{n} x_j\left(\sum_{i=1}^{m} a_{ij}\mathbf{w}_i\right) = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j\right)\mathbf{w}_i$$

Para $i = 1, \ldots, m$, seja

$$y_i = \sum_{j=1}^{n} a_{ij}x_j$$

Logo,

$$\mathbf{y} = (y_1, y_2, \ldots, y_m)^T = A\mathbf{x}$$

é o vetor de coordenadas de $L(\mathbf{v})$ em relação a $[\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m]$. Acabamos de mostrar o teorema abaixo.

Teorema 4.2.2 (Teorema de Representação Matricial). *Se $E = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ e $F = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m]$ são bases ordenadas para V e W, respectivamente, então, a cada transformação linear $L : V \to W$, corresponde uma matriz A $m \times n$ tal que*

$$[L(\mathbf{v})]_F = A[\mathbf{v}]_E \quad \text{para todo } \mathbf{v} \in V$$

A é a matriz de L em relação às bases ordenadas E e F. De fato,

$$\mathbf{a}_j = \left[L(\mathbf{v}_j)\right]_F \qquad j = 1, 2, \ldots, n$$

O Teorema 4.2.2 está ilustrado na Fig. 4.2.2. Se A é a matriz que representa L em relação às bases E e F, e se

$$\mathbf{x} = [\mathbf{v}]_E \qquad \text{(o vetor de coordenadas de } \mathbf{v} \text{ em relação a } E\text{)}$$

$$\mathbf{y} = [\mathbf{w}]_F \qquad \text{(o vetor de coordenadas de } \mathbf{w} \text{ em relação a } F\text{)}$$

então L leva $\mathbf{v}$ em $\mathbf{w}$ se e somente se $A\mathbf{x} = \mathbf{y}$.

EXEMPLO 3. Seja L a transformação linear de R^3 em R^2 definida por

$$L(\mathbf{x}) = x_1\mathbf{b}_1 + (x_2 + x_3)\mathbf{b}_2$$

$$\begin{tikzcd}
\mathbf{v} \in V \arrow[r, "L = L_A"] \arrow[d, leftrightarrow] & \mathbf{w} = L(\mathbf{v}) \in W \arrow[d, leftrightarrow] \\
\mathbf{x} = [\mathbf{v}]_E \in R^n \arrow[r, "A"] & A\mathbf{x} = [\mathbf{w}]_F \in R^m
\end{tikzcd}$$

FIG. 4.2.2

para $\mathbf{x} \in R^3$, onde

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{e} \quad \mathbf{b}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Encontre a matriz A que representa L em relação às bases ordenadas $[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$ e $[\mathbf{b}_1, \mathbf{b}_2]$.

SOLUÇÃO

$$L(\mathbf{e}_1) = 1\mathbf{b}_1 + 0\mathbf{b}_2$$
$$L(\mathbf{e}_2) = 0\mathbf{b}_1 + 1\mathbf{b}_2$$
$$L(\mathbf{e}_3) = 0\mathbf{b}_1 + 1\mathbf{b}_2$$

A i-ésima coluna de A é determinada pelas coordenadas de $L(\mathbf{e}_i)$ em relação à base $[\mathbf{b}_1, \mathbf{b}_2]$ para $i = 1, 2, 3$. Logo,

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$ □

EXEMPLO 4. Seja L a transformação linear de R^2 em si mesmo definida por

$$L(\alpha\mathbf{b}_1 + \beta\mathbf{b}_2) = (\alpha + \beta)\mathbf{b}_1 + 2\beta\mathbf{b}_2$$

onde $[\mathbf{b}_1, \mathbf{b}_2]$ é a base ordenada do Exemplo 3. Encontre a matriz A que representa L em relação a $[\mathbf{b}_1, \mathbf{b}_2]$.

SOLUÇÃO

$$L(\mathbf{b}_1) = 1\mathbf{b}_1 + 0\mathbf{b}_2$$
$$L(\mathbf{b}_2) = 1\mathbf{b}_1 + 2\mathbf{b}_2$$

Logo,

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$$ □

EXEMPLO 5. O operador linear D definido por $D(p) = p'$ leva P_3 em P_2. Dadas as bases ordenadas $[x^2, x, 1]$ e $[x, 1]$ para P_3 e P_2, respectivamente, queremos determinar uma representação matricial para D. Para fazer isso, aplicamos D a cada um dos elementos da base de P_3.

$$D(x^2) = 2x + 0 \cdot 1$$
$$D(x) = 0x + 1 \cdot 1$$
$$D(1) = 0x + 0 \cdot 1$$

Em P_2, os vetores de coordenadas para $D(x^2)$, $D(x)$, $D(1)$ são $(2, 0)^T$, $(0, 1)^T$, $(0, 0)^T$, respectivamente. A matriz A que tem esses vetores por colunas é

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Se $p(x) = ax^2 + bx + c$, o vetor de coordenadas de p em relação à base ordenada de P_3 é $(a, b, c)^T$. Para encontrar o vetor de coordenadas de $D(p)$ em relação à base ordenada de P_2, basta multiplicar:

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 2a \\ b \end{pmatrix}$$

Logo, $D(ax^2 + bx + c) = 2ax + b$. □

Para encontrar a representação matricial A de uma transformação linear $L : R^n \to R^m$ em relação às bases ordenadas $E = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$ e $F = [\mathbf{b}_1, \ldots, \mathbf{b}_m]$, precisamos representar cada vetor $L(\mathbf{u})$ como uma combinação linear de $\mathbf{b}_1, \ldots, \mathbf{b}_m$. O teorema a seguir mostra que encontrar essa representação para $L(\mathbf{u}_j)$ é equivalente a resolver o sistema linear $B\mathbf{x} = L(\mathbf{u}_j)$.

Teorema 4.2.3. *Sejam $E = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$ e $F = [\mathbf{b}_1, \ldots, \mathbf{b}_m]$ bases ordenadas para R^n e R^m, respectivamente. Se $L : R^n \to R^m$ é uma transformação linear e A é a matriz de L em relação a E e F, então*

$$\mathbf{a}_j = B^{-1}L(\mathbf{u}_j) \quad \text{para } j = 1, \ldots, n$$

onde $B = (\mathbf{b}_1, \ldots, \mathbf{b}_m)$.

Demonstração. Se A é a matriz de L em relação a E e F, então, para $j = 1, \ldots, n$,

$$\begin{aligned} L(\mathbf{u}_j) &= a_{1j}\mathbf{b}_1 + a_{2j}\mathbf{b}_2 + \cdots + a_{mj}\mathbf{b}_m \\ &= B\mathbf{a}_j \end{aligned}$$

A matriz B é invertível, já que suas colunas formam uma base para R^m. Portanto,

$$\mathbf{a}_j = B^{-1}L(\mathbf{u}_j) \qquad j = 1, \ldots, n$$ □

Uma conseqüência desse teorema é que podemos determinar a matriz associada a um operador calculando a forma escada reduzida por linhas de uma matriz aumentada. O próximo corolário mostra como fazer isso.

Corolário 4.2.4. *Se A é a matriz do operador linear $L : R^n \to R^m$ em relação às bases $E = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$ e $F = [\mathbf{b}_1, \ldots, \mathbf{b}_m]$, então a forma escada reduzida por linhas de $(\mathbf{b}_1, \ldots, \mathbf{b}_m \mid L(\mathbf{u}_1), \ldots, L(\mathbf{u}_n))$ é $(I \mid A)$.*

Demonstração. Seja $B = (\mathbf{b}_1, \ldots, \mathbf{b}_m)$. A matriz $(B \mid L(\mathbf{u}_1), \ldots, L(\mathbf{u}_n))$ é equivalente por linhas a

$$\begin{aligned} B^{-1}(B \mid L(\mathbf{u}_1), \ldots, L(\mathbf{u}_n)) &= (I \mid B^{-1}L(\mathbf{u}_1), \ldots, B^{-1}L(\mathbf{u}_n)) \\ &= (I \mid \mathbf{a}_1, \ldots, \mathbf{a}_n) \\ &= (I \mid A) \end{aligned}$$ □

EXEMPLO 6. Seja $L: R^2 \to R^3$ a transformação linear definida por

$$L(\mathbf{x}) = (x_2, x_1 + x_2, x_1 - x_2)^T$$

Encontre a matriz de L em relação às bases ordenadas $[\mathbf{u}_1, \mathbf{u}_2]$ e $[\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3]$, onde

$$\mathbf{u}_1 = (1, 2)^T, \qquad \mathbf{u}_2 = (3, 1)^T$$

e

$$\mathbf{b}_1 = (1, 0, 0)^T, \qquad \mathbf{b}_2 = (1, 1, 0)^T, \qquad \mathbf{b}_3 = (1, 1, 1)^T$$

SOLUÇÃO. Precisamos calcular $L(\mathbf{u}_1)$, $L(\mathbf{u}_2)$ e depois colocar a matriz $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3 \mid L(\mathbf{u}_1), L(\mathbf{u}_2))$ em forma escada reduzida por linhas.

$$L(\mathbf{u}_1) = (2, 3, -1)^T \qquad \text{e} \qquad L(\mathbf{u}_2) = (1, 4, 2)^T$$

$$\left(\begin{array}{ccc|cc} 1 & 1 & 1 & 2 & 1 \\ 0 & 1 & 1 & 3 & 4 \\ 0 & 0 & 1 & -1 & 2 \end{array}\right) \to \left(\begin{array}{ccc|cc} 1 & 0 & 0 & -1 & -3 \\ 0 & 1 & 0 & 4 & 2 \\ 0 & 0 & 1 & -1 & 2 \end{array}\right)$$

A matriz de L em relação às bases ordenadas dadas é

$$A = \begin{pmatrix} -1 & -3 \\ 4 & 2 \\ -1 & 2 \end{pmatrix}$$

O leitor pode verificar que

$$L(\mathbf{u}_1) = -\mathbf{b}_1 + 4\mathbf{b}_2 - \mathbf{b}_3$$

$$L(\mathbf{u}_2) = -3\mathbf{b}_1 + 2\mathbf{b}_2 + 2\mathbf{b}_3$$

□

EXERCÍCIOS

1. Para cada uma das transformações lineares L no Exercício 1 da Seção 1, encontre a matriz A que representa L.

2. Para cada uma das transformações lineares L de R^3 em R^2 a seguir, encontre uma matriz A tal que $L(\mathbf{x}) = A\mathbf{x}$ para todo $\mathbf{x}$ em R^3.
(a) $L((x_1, x_2, x_3)^T) = (x_1 + x_2, 0)^T$
(b) $L((x_1, x_2, x_3)^T) = (x_1, x_2)^T$
(c) $L((x_1, x_2, x_3)^T) = (x_2 - x_1, x_3 - x_2)^T$

3. Para cada uma das transformações lineares L de R^3 em R^3 a seguir, encontre uma matriz A tal que $L(\mathbf{x}) = A\mathbf{x}$ para todo $\mathbf{x}$ em R^3.
(a) $L((x_1, x_2, x_3)^T) = (x_3, x_2, x_1)^T$
(b) $L((x_1, x_2, x_3)^T) = (x_1, x_1 + x_2, x_1 + x_2 + x_3)^T$
(c) $L((x_1, x_2, x_3)^T) = (2x_3, x_2 + 3x_1, 2x_1 - x_3)^T$

4. Seja L a transformação linear de R^3 em R^3 definida por
$L(\mathbf{x}) = (2x_1 - x_2 - x_3, 2x_2 - x_1 - x_3, 2x_3 - x_1 - x_2)^T$
Determine a matriz A de L em relação à base canônica e use-a para encontrar $L(\mathbf{x})$ para cada um dos vetores $\mathbf{x}$ a seguir.
(a) $\mathbf{x} = (1, 1, 1)^T$ (b) $\mathbf{x} = (2, 1, 1)^T$ (c) $\mathbf{x} = (-5, 3, 2)^T$

5. Encontre a representação matricial canônica para cada um dos operadores lineares L em R^2 descritos a seguir.
(a) L roda cada vetor $\mathbf{x}$ de 45° no sentido antitrigonométrico.
(b) L reflete cada vetor $\mathbf{x}$ em relação ao eixo dos x_1 e depois roda o vetor refletido de 90° no sentido trigonométrico.
(c) L dobra o comprimento de $\mathbf{x}$ e depois roda o vetor obtido de 30° no sentido trigonométrico.
(d) L reflete cada vetor $\mathbf{x}$ em relação à reta $x_1 = x_2$ e depois projeta o vetor refletido sobre o eixo dos x_1.

6. Sejam

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{b}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{b}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

e seja L a transformação linear de R^2 em R^3 definida por

$$L(\mathbf{x}) = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + (x_1 + x_2)\mathbf{b}_3$$

Encontre a matriz A de L em relação às bases $[\mathbf{e}_1, \mathbf{e}_2]$ e $[\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3]$.

7. Sejam

$$\mathbf{y}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{y}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{y}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

e seja $\mathcal{I}$ o operador identidade em R^3.
(a) Encontre as coordenadas de $\mathcal{I}(\mathbf{e}_1)$, $\mathcal{I}(\mathbf{e}_2)$, $\mathcal{I}(\mathbf{e}_3)$ em relação a $[\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3]$.

(b) Encontre uma matriz A tal que $A\mathbf{x}$ é o vetor de coordenadas de $\mathbf{x}$ em relação a $[\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3]$.

8. Sejam $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$ como no Exercício 7 e seja L a transformação linear de R^3 em R^3 definida por

$$L(c_1\mathbf{y}_1 + c_2\mathbf{y}_2 + c_3\mathbf{y}_3) = (c_1 + c_2 + c_3)\mathbf{y}_1 + (2c_1 + c_3)\mathbf{y}_2 - (2c_2 + c_3)\mathbf{y}_3$$

(a) Encontre a matriz de L em relação à base ordenada $[\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3]$.
(b) Escreva cada um dos vetores $\mathbf{x}$ a seguir como uma combinação linear de $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$ e use a matriz encontrada em (a) para determinar $L(\mathbf{x})$.

(i) $\mathbf{x} = (7, 5, 2)^T$ (ii) $\mathbf{x} = (3, 2, 1)^T$ (iii) $\mathbf{x} = (1, 2, 3)^T$

9. Seja L o operador linear de P_2 em R^2 definido por

$$L(p(x)) = \begin{pmatrix} \int_0^1 p(x)\,dx \\ p(0) \end{pmatrix}$$

Encontre uma matriz A tal que

$$L(\alpha + \beta x) = A\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

10. O operador linear definido por

$$L(p(x)) = p'(x) + p(0)$$

vai de P_3 em P_2. Encontre a matriz de L em relação às bases ordenadas $[x^2, x, 1]$ e $[2, 1 - x]$. Para cada um dos vetores $p(x)$ em P_3 a seguir, encontre as coordenadas de $L(p(x))$ em relação à base ordenada $[2, 1 - x]$.

(a) $x^2 + 2x - 3$ (b) $x^2 + 1$ (c) $3x$ (d) $4x^2 + 2x$

11. Seja S o subespaço de $C[a, b]$ gerado por e^x, xe^x e x^2e^x. Seja D o operador derivada em S. Encontre a matriz de D em relação à base $[e^x, xe^x, x^2e^x]$.

12. Seja L uma transformação linear de R^n em R^n. Suponha que $L(\mathbf{x}) = 0$ para alguma $\mathbf{x} \neq \mathbf{0}$. Seja A a matriz de L em relação à base canônica $[\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n]$. Mostre que A é singular.

13. Seja L um operador linear de um espaço vetorial V em si mesmo. Seja A a matriz de L em relação à base ordenada $[\mathbf{v}_1, \ldots, \mathbf{v}_n]$ [isto é, $L(\mathbf{v}_j) = \sum_{i=1}^{n} a_{ij}\mathbf{v}_i, j = 1, \ldots, n$]. Mostre que A^m é a matriz de L^m em relação a $[\mathbf{v}_1, \ldots, \mathbf{v}_n]$.

14. Sejam $E = [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3]$ e $F = [\mathbf{b}_1, \mathbf{b}_2]$, onde

$$\mathbf{u}_1 = (1, 0, -1)^T, \quad \mathbf{u}_2 = (1, 2, 1)^T, \quad \mathbf{u}_3 = (-1, 1, 1)^T$$

e

$$\mathbf{b}_1 = (1, -1)^T, \quad \mathbf{b}_2 = (2, -1)^T$$

Para cada uma das transformações lineares L de R^3 em R^2 a seguir, encontre a matriz de L em relação às bases ordenadas E e F.

(a) $L(\mathbf{x}) = (x_3, x_1)^T$
(b) $L(\mathbf{x}) = (x_1 + x_2, x_1 - x_3)^T$
(c) $L(\mathbf{x}) = (2x_2, -x_1)^T$

15. Suponha que $L_1 : V \to W$ e $L_2 : W \to Z$ são transformações lineares e que E, F e G são bases ordenadas para V, W e Z, respectivamente. Mostre que, se A é a matriz de L_1 em relação às bases E e F e se B é a matriz de L_2 em relação às bases F e G, então a matriz $C = BA$ é a matriz de $L_2 \circ L_1$: $V \to Z$ em relação a E e G.

[***Sugestão:*** Mostre que $BA[\mathbf{v}]_E = [(L_2 \circ L_1)(\mathbf{v})]_G$ para todo $\mathbf{v} \in V$.]

16. Sejam V e W espaços vetoriais com bases ordenadas E e F, respectivamente. Se $L : V \to W$ é uma transformação linear e A é sua matriz em relação a E e F, mostre que:
(a) $\mathbf{v} \in \ker(L)$ se e somente se $[\mathbf{v}]_E \in N(A)$;
(b) $\mathbf{w} \in L(V)$ se e somente se $[\mathbf{w}]_F$ pertence ao espaço coluna de A.

3 SEMELHANÇA

Se L é uma transformação linear de um espaço vetorial V de dimensão n em si mesmo, a representação matricial de L depende da base ordenada escolhida para V. Usando bases diferentes, é possível representar L por matrizes diferentes $n \times n$. Nesta seção, vamos considerar representações matriciais diferentes de operadores lineares e caracterizar a relação entre matrizes associadas ao mesmo operador linear.

Vamos começar considerando um exemplo em R^2. Seja L a transformação linear de R^2 em si mesmo definida por

$$L(\mathbf{x}) = (2x_1, x_1 + x_2)^T$$

Como

$$L(\mathbf{e}_1) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \text{e} \quad L(\mathbf{e}_2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

a matriz de L em relação a $[\mathbf{e}_1, \mathbf{e}_2]$ é

$$A = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$$

Usando uma base diferente, a matriz de L muda. Por exemplo, se usarmos como base

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{e} \quad \mathbf{u}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

então, para determinar a matriz de L em relação a $[\mathbf{u}_1, \mathbf{u}_2]$, precisamos determinar $L(\mathbf{u}_1)$, $L(\mathbf{u}_2)$ e escrever esses vetores como uma combinação linear de $\mathbf{u}_1$ e $\mathbf{u}_2$. Podemos usar a matriz A para encontrar $L(\mathbf{u}_1)$ e $L(\mathbf{u}_2)$.

$$L(\mathbf{u}_1) = A\mathbf{u}_1 = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

$$L(\mathbf{u}_2) = A\mathbf{u}_2 = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

Para expressar esses vetores em termos de $\mathbf{u}_1$ e $\mathbf{u}_2$, use a matriz mudança de base de $[\mathbf{e}_1, \mathbf{e}_2]$ para $[\mathbf{u}_1, \mathbf{u}_2]$. Vamos primeiro calcular a matriz mudança de base de $[\mathbf{u}_1, \mathbf{u}_2]$ para $[\mathbf{e}_1, \mathbf{e}_2]$. Ela é, simplesmente,

$$U = (\mathbf{u}_1, \mathbf{u}_2) = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

Logo, a matriz mudança de base de $[\mathbf{e}_1, \mathbf{e}_2]$ para $[\mathbf{u}_1, \mathbf{u}_2]$ é

$$U^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Para determinar as coordenadas de $L(\mathbf{u}_1)$, $L(\mathbf{u}_2)$ em relação a $[\mathbf{u}_1, \mathbf{u}_2]$, multiplicamos esses vetores por U^{-1}.

$$U^{-1}L(\mathbf{u}_1) = U^{-1}A\mathbf{u}_1 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

$$U^{-1}L(\mathbf{u}_2) = U^{-1}A\mathbf{u}_2 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Então,

$$\begin{aligned} L(\mathbf{u}_1) &= 2\mathbf{u}_1 + 0\mathbf{u}_2 \\ L(\mathbf{u}_2) &= -1\mathbf{u}_1 + 1\mathbf{u}_2 \end{aligned}$$

e a matriz de L em relação a $[\mathbf{u}_1, \mathbf{u}_2]$ é

$$B = \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix}$$

Qual a relação entre A e B? Observe que as colunas de B são

$$\begin{pmatrix} 2 \\ 0 \end{pmatrix} = U^{-1}A\mathbf{u}_1 \qquad \text{e} \qquad \begin{pmatrix} -1 \\ 1 \end{pmatrix} = U^{-1}A\mathbf{u}_2$$

Logo,

$$B = (U^{-1}A\mathbf{u}_1, U^{-1}A\mathbf{u}_2) = U^{-1}A(\mathbf{u}_1, \mathbf{u}_2) = U^{-1}AU$$

Portanto, se

(i) B é a matriz de L em relação a $[\mathbf{u}_1, \mathbf{u}_2]$,
(ii) A é a matriz de L em relação a $[\mathbf{e}_1, \mathbf{e}_2]$,
(iii) U é a matriz mudança de base de $[\mathbf{u}_1, \mathbf{u}_2]$ para $[\mathbf{e}_1, \mathbf{e}_2]$,

temos

(1) $$B = U^{-1}AU$$

Os resultados que estabelecemos para esse operador linear particular em R^2 são típicos do que acontece em um contexto muito mais geral. Vamos mostrar a seguir que a relação (1) é válida para duas representações matriciais quaisquer de um operador linear de um espaço vetorial de dimensão n em si mesmo.

Teorema 4.3.1. *Sejam $E = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ e $F = [\mathbf{w}_1, \ldots, \mathbf{w}_n]$ duas bases ordenadas para um espaço vetorial V e seja L um operador linear de V em si mesmo. Seja S a matriz mudança de base de F para E. Se A é a matriz de L em relação a E e B é a matriz de L em relação a F, então $B = S^{-1}AS$.*

Demonstração. Seja $\mathbf{x}$ qualquer vetor em R^n e seja

$$\mathbf{v} = x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + \cdots + x_n\mathbf{w}_n$$

Defina

(2) $$\mathbf{y} = S\mathbf{x}, \qquad \mathbf{t} = A\mathbf{y}, \qquad \mathbf{z} = B\mathbf{x}$$

Pela definição de S, temos que $\mathbf{y} = [\mathbf{v}]_E$ e, portanto,

$$\mathbf{v} = y_1\mathbf{v}_1 + \cdots + y_n\mathbf{v}_n$$

Como A é a matriz de L em relação a E e B é a matriz de L em relação a F, tem-se

$$\mathbf{t} = [L(\mathbf{v})]_E \qquad \text{e} \qquad \mathbf{z} = [L(\mathbf{v})]_F$$

A matriz mudança de base de E para F é S^{-1}, logo

(3) $$S^{-1}\mathbf{t} = \mathbf{z}$$

De (2) e (3), tem-se que

$$S^{-1}AS\mathbf{x} = S^{-1}A\mathbf{y} = S^{-1}\mathbf{t} = \mathbf{z} = B\mathbf{x}$$

$$
\begin{tikzcd}
\mathbf{y} \arrow[r, "A"] & \mathbf{t} \arrow[d, "S^{-1}"] \\
\mathbf{x} \arrow[u, "S"] \arrow[r, "B"'] & \mathbf{z}
\end{tikzcd}
$$

FIG. 4.3.1

(ver Fig. 4.3.1). Então,

$$S^{-1}AS\mathbf{x} = B\mathbf{x}$$

para todo $\mathbf{x} \in R^n$ e, portanto, $S^{-1}AS = B$. □

Uma outra maneira de ver o Teorema 4.3.1 é considerar S a matriz do operador identidade $\mathcal{I}$ em relação às bases ordenadas $E = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ e $F = [\mathbf{w}_1, \ldots, \mathbf{w}_n]$. Então, se

S é a matriz de $\mathcal{I}$ em relação a E e F,
A é a matriz de L em relação a E,
S^{-1} é a matriz de $\mathcal{I}$ em relação a F e E,

L pode ser escrito como um operador composto $\mathcal{I} \circ L \circ \mathcal{I}$, e a representação de uma composição é o produto das matrizes associadas a cada componente. Logo, a matriz de $\mathcal{I} \circ L \circ \mathcal{I}$ em relação a F é $S^{-1}AS$. Se B é a matriz de L em relação a F, então B tem que ser igual a $S^{-1}AS$ (ver Fig. 4.3.2).

Definição. Sejam A e B matrizes $n \times n$. Dizemos que B é **semelhante** a A se existe uma matriz invertível S tal que $B = S^{-1}AS$.

Observe que, se B é semelhante a A, então A é semelhante a B. Logo, podemos dizer simplesmente que A e B são semelhantes.

Pelo Teorema 4.3.1, se A e B são duas matrizes representando o mesmo operador L, então A e B são semelhantes. Por outro lado, suponha que A é a matriz de L em relação à base ordenada $[\mathbf{v}_1, \ldots, \mathbf{v}_n]$ e que $B = S^{-1}AS$ para alguma matriz invertível S. Se $\mathbf{w}_1, \ldots, \mathbf{w}_n$ são definidos por

$$
\begin{aligned}
\mathbf{w}_1 &= s_{11}\mathbf{v}_1 + s_{21}\mathbf{v}_2 + \cdots + s_{n1}\mathbf{v}_n \\
\mathbf{w}_2 &= s_{12}\mathbf{v}_1 + s_{22}\mathbf{v}_2 + \cdots + s_{n2}\mathbf{v}_n \\
&\vdots \\
\mathbf{w}_n &= s_{1n}\mathbf{v}_1 + s_{2n}\mathbf{v}_2 + \cdots + s_{nn}\mathbf{v}_n
\end{aligned}
$$

$$
\begin{tikzcd}
\text{Base } E\text{:} & V \arrow[r, "L", "A"'] & V \arrow[d, "\mathcal{I}", "S^{-1}"'] \\
\text{Base } F\text{:} & V \arrow[u, "\mathcal{I}", "S"'] \arrow[r, "B", "L"'] & V
\end{tikzcd}
$$

FIG. 4.3.2

então $[\mathbf{w}_1, \ldots, \mathbf{w}_n]$ é uma base ordenada de V e B é a matriz de L em relação a $[\mathbf{w}_1, \ldots, \mathbf{w}_n]$.

EXEMPLO 1. Seja D o operador derivada em P_3. Encontre a matriz B de D em relação a $[1, x, x^2]$ e a matriz A de D em relação a $[1, 2x, 4x^2 - 2]$.

SOLUÇÃO

$$D(1) = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2$$
$$D(x) = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2$$
$$D(x^2) = 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2$$

Logo, a matriz B é dada por

$$B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Aplicando D a 1, $2x$ e $4x^2 - 2$, obtemos

$$D(1) = 0 \cdot 1 + 0 \cdot 2x + 0 \cdot (4x^2 - 2)$$
$$D(2x) = 2 \cdot 1 + 0 \cdot 2x + 0 \cdot (4x^2 - 2)$$
$$D(4x^2 - 2) = 0 \cdot 1 + 4 \cdot 2x + 0 \cdot (4x^2 - 2)$$

Portanto,

$$A = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}$$

A matriz mudança de base S de $[1, 2x, 4x^2 - 2]$ para $[1, x, x^2]$ e sua inversa são dadas por

$$S = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \quad \text{e} \quad S^{-1} = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} \end{pmatrix}$$

(Ver Exemplo 7 na Seção 5 do Cap. 3.) O leitor pode verificar que $A = S^{-1}BS$. □

EXEMPLO 2. Seja L o operador linear de R^3 em R^3 definido por $L(\mathbf{x}) = A\mathbf{x}$, onde

$$A = \begin{pmatrix} 2 & 2 & 0 \\ 1 & 1 & 2 \\ 1 & 1 & 2 \end{pmatrix}$$

A matriz A, portanto, é a matriz de L em relação a $[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$. Encontre a matriz de L em relação a $[\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3]$, onde

$$\mathbf{y}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \mathbf{y}_2 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{y}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

SOLUÇÃO

$$L(\mathbf{y}_1) = A\mathbf{y}_1 = \mathbf{0} = 0\mathbf{y}_1 + 0\mathbf{y}_2 + 0\mathbf{y}_3$$
$$L(\mathbf{y}_2) = A\mathbf{y}_2 = \mathbf{y}_2 = 0\mathbf{y}_1 + 1\mathbf{y}_2 + 0\mathbf{y}_3$$
$$L(\mathbf{y}_3) = A\mathbf{y}_3 = 4\mathbf{y}_3 = 0\mathbf{y}_1 + 0\mathbf{y}_2 + 4\mathbf{y}_3$$

Logo, a matriz de L em relação a $[\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3]$ é

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

□

Poderíamos ter encontrado D usando a matriz mudança de base $Y = (\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3)$ e calculando

$$D = Y^{-1}AY$$

Isso não foi necessário devido à simplicidade da ação de L na base $[\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3]$.

No Exemplo 2, o operador linear L é representado por uma matriz diagonal D em relação à base $[\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3]$. É muito mais simples trabalhar com D do que com A. Por exemplo, é mais fácil calcular $D\mathbf{x}$ e $D^n\mathbf{x}$ do que $A\mathbf{x}$ e $A^n\mathbf{x}$. De modo geral, é desejável encontrar a representação matricial mais simples possível para um operador linear. Em particular, se o operador puder ser representado por uma matriz diagonal, essa é, normalmente, a representação preferida. O problema de encontrar uma matriz diagonal associada a um operador linear será estudado no Cap. 6.

EXERCÍCIOS

1. Para cada uma das transformações lineares L de R^2 em R^2 a seguir, determine a matriz A que representa L em relação a $[\mathbf{e}_1, \mathbf{e}_2]$ (ver Exercício 1 da Seção 2) e a matriz B que representa L em relação a $[\mathbf{u}_1 = (1, 1)^T, \mathbf{u}_2 = (-1, 1)^T$.

(a) $L(\mathbf{x}) = (-x_1, x_2)^T$ (b) $L(\mathbf{x}) = -\mathbf{x}$ (c) $L(\mathbf{x}) = (x_2, x_1)^T$

(d) $L(\mathbf{x}) = \frac{1}{2}\mathbf{x}$ (e) $L(\mathbf{x}) = x_2\mathbf{e}_2$

2. Sejam $[\mathbf{u}_1, \mathbf{u}_2]$ e $[\mathbf{v}_1, \mathbf{v}_2]$ bases ordenadas de R^2, onde

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \quad \text{e} \quad \mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Seja L a transformação linear definida por

$$L(\mathbf{x}) = (-x_1, x_2)^T$$

e seja B a matriz de L em relação a $[\mathbf{u}_1, \mathbf{u}_2]$ [do Exercício 1(a)].

(a) Encontre a matriz mudança de base S de $[\mathbf{u}_1, \mathbf{u}_2]$ para $[\mathbf{v}_1, \mathbf{v}_2]$.

(b) Encontre a matriz A que representa L em relação a $[\mathbf{v}_1, \mathbf{v}_2]$ calculando SBS^{-1}.

(c) Verifique que

$$L(\mathbf{v}_1) = a_{11}\mathbf{v}_1 + a_{21}\mathbf{v}_2$$

$$L(\mathbf{v}_2) = a_{12}\mathbf{v}_1 + a_{22}\mathbf{v}_2$$

3. Seja L a transformação linear em R^3 definida por

$$L(\mathbf{x}) = (2x_1 - x_2 - x_3, 2x_2 - x_1 - x_3, 2x_3 - x_1 - x_2)^T$$

e seja A a matriz de L em relação a $[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$ (ver Exercício 4 da Seção 3). Se $\mathbf{u}_1 = (1, 1, 0)^T$, $\mathbf{u}_2 = (1, 0, 1)^T$ e $\mathbf{u}_3 = (0, 1, 1)^T$, então $[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3]$ é uma base ordenada para R^3.

(a) Encontre a matriz mudança de base U de $[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3]$ para $[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$.

(b) Determine a matriz B que representa L em relação a $[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3]$ calculando $U^{-1}AU$.

4. Seja L o operador linear de R^3 em R^3 definido por $L(\mathbf{x}) = A\mathbf{x}$, onde

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

e sejam

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \qquad \mathbf{v}_3 = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}$$

Encontre a matriz mudança de base V de $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$ para $[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$ e use-a para encontrar a matriz B que representa L em relação a $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$.

5. Seja L o operador em P_3 definido por

$$L(p(x)) = xp'(x) + p''(x)$$

(a) Encontre a matriz A que representa L em relação a $[1, x, x^2]$.
(b) Encontre a matriz B que representa L em relação a $[1, x, 1 + x^2]$.
(c) Encontre a matriz S tal que $B = S^{-1}AS$.
(d) Se $p(x) = a_0 + a_1x + a_2(1 + x^2)$, calcule $L^n(p(x))$.

6. Seja V o subespaço de $C[a, b]$ gerado por $1, e^x, e^{-x}$ e seja D o operador derivada em V.
(a) Encontre a matriz mudança de base S que corresponde à mudança das coordenadas em relação a $[1, e^x, e^{-x}]$ para $[1, \cosh x, \operatorname{senh} x]$. [$\cosh x = (e^x + e^{-x})/2$, $\operatorname{senh} x = (e^x - e^{-x})/2$.]
(b) Encontre a matriz A que representa D em relação a $[1, \cosh x, \operatorname{senh} x]$.
(c) Encontre a matriz B que representa D em relação a $[1, e^x, e^{-x}]$.
(d) Verifique que $B = S^{-1}AS$.

7. Prove que, se A é semelhante a B e se B é semelhante a C, então A é semelhante a C.

8. Suponha que $A = S\Lambda S^{-1}$, onde Λ é uma matriz diagonal com elementos diagonais $\lambda_1, \lambda_2, \ldots, \lambda_n$.
(a) Mostre que $A\mathbf{s}_i = \lambda_i\mathbf{s}_i$, $i = 1, \ldots, n$.
(b) Mostre que, se $\mathbf{x} = \alpha_1\mathbf{s}_1 + \alpha_2\mathbf{s}_2 + \cdots + \alpha_n\mathbf{s}_n$, então

$$A^k\mathbf{x} = \alpha_1\lambda_1^k\mathbf{s}_1 + \alpha_2\lambda_2^k\mathbf{s}_2 + \cdots + \alpha_n\lambda_n^k\mathbf{s}_n$$

(c) Suponha que $|\lambda_i| < 1$ para $i = 1, \ldots, n$. O que acontece com $A^k\mathbf{x}$ quando $k \to \infty$? Explique.

9. Suponha que $A = ST$, onde S é invertível. Seja $B = TS$. Mostre que B é semelhante a A.

10. Sejam A e B matrizes $n \times n$. Mostre que, se A é semelhante a B, então existem matrizes S e T $n \times n$, com S invertível, tais que

$$A = ST \qquad \text{e} \qquad B = TS$$

11. Mostre que, se A e B são matrizes semelhantes, então $\det(A) = \det(B)$.

12. Sejam A e B matrizes semelhantes. Mostre que:
(a) A^T e B^T são semelhantes;
(b) A^k e B^k são semelhantes para todo inteiro positivo k.

13. Mostre que, se A é semelhante a B e se A é invertível, então B também é invertível e A^{-1} e B^{-1} também são semelhantes.

14. O *traço* de uma matriz A $n \times n$, denotado por $\operatorname{tr}(A)$, é a soma de seus elementos diagonais, isto é,

$$\operatorname{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$$

Mostre que:
(a) $\operatorname{tr}(AB) = \operatorname{tr}(BA)$;
(b) se A é semelhante a B, então $\operatorname{tr}(A) = \operatorname{tr}(B)$.

15. Sejam A e B matrizes semelhantes e seja λ um escalar arbitrário. Mostre que:
(a) $A - \lambda I$ e $B - \lambda I$ são semelhantes;
(b) $\det(A - \lambda I) = \det(B - \lambda I)$.

EXERCÍCIOS COM O MATLAB PARA O CAPÍTULO 4

1. Use o MATLAB para gerar uma matriz W e um vetor $\mathbf{x}$ digitando

$$W = \texttt{triu(ones(5))} \qquad \text{e} \qquad \mathbf{x} = [1:5]'$$

As colunas de W podem ser usadas para se formar uma base ordenada

$$F = [\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4, \mathbf{w}_5]$$

Seja $L : R^5 \to R^5$ um operador linear tal que

$$L(\mathbf{w}_1) = \mathbf{w}_2, \qquad L(\mathbf{w}_2) = \mathbf{w}_3, \qquad L(\mathbf{w}_3) = \mathbf{w}_4$$

e

$$L(\mathbf{w}_4) = 4\mathbf{w}_1 + 3\mathbf{w}_2 + 2\mathbf{w}_3 + \mathbf{w}_4$$

$$L(\mathbf{w}_5) = \mathbf{w}_1 + \mathbf{w}_2 + \mathbf{w}_3 + 3\mathbf{w}_4 + \mathbf{w}_5$$

(a) Determine a matriz A que representa L em relação a F e coloque-a no MATLAB.
(b) Use o MATLAB para calcular o vetor de coordenadas $\mathbf{y} = W^{-1}\mathbf{x}$ de $\mathbf{x}$ em relação a F.
(c) Use A para calcular o vetor de coordenadas $\mathbf{z}$ de $L(\mathbf{x})$ em relação a F.
(d) W é a matriz mudança de base de F para a base canônica de R^5. Use W para calcular o vetor de coordenadas de $L(\mathbf{x})$ em relação à base canônica.

2. Faça A = `triu(ones(5)). * tril(ones(5))`. Se L é o operador linear definido por $L(\mathbf{x}) = A\mathbf{x}$ para todo $\mathbf{x}$ em R^n, então A é a matriz de L em relação à base canônica de R^5. Construa uma matriz U 5×5 fazendo

$$U = \texttt{hankel(ones}(5, 1), 1 : 5)$$

Use a função `rank` do MATLAB para verificar que os vetores colunas de U são linearmente independentes. Logo, $E = [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5]$ é uma base ordenada para R^5. A matriz U é a matriz mudança de base de E para a base canônica.

(a) Use o MATLAB para calcular a matriz B que representa L em relação a E. (A matriz B deve ser calculada em termos de A, U e U^{-1}.)
(b) Gere outra matriz fazendo V = `toeplitz`([1, 0, 1, 1, 1]). Use o MATLAB para verificar que V é invertível. Como os vetores colunas de V são linearmente independentes, eles formam uma base ordenada F para R^5. Use o MATLAB para encontrar a matriz C que representa L em relação a F. (A matriz C deve ser calculada em termos de A, V, V^{-1}.)
(c) As matrizes A e B nos itens (a) e (b) deveriam ser semelhantes. Por quê? Explique. Use o MATLAB para calcular a matriz mudança de base S de F para E. Calcule a matriz C em termos de B, S e S^{-1}. Compare seu resultado com o encontrado em (b).

3. Faça

$$A = \texttt{toeplitz}(1 : 7), \qquad S = \texttt{compan(ones}(8, 1))$$

e $B = S^{-1}$`*A*S`. As matrizes A e B são semelhantes. Use o MATLAB para verificar as seguintes propriedades nessas duas matrizes:

(a) $\det(B) = \det(A)$
(b) $B^T = S^T A^T (S^T)^{-1}$
(c) $B^{-1} = S^{-1} A^{-1} S$
(d) $B^9 = S^{-1} A^9 S$
(e) $\operatorname{tr}(B) = \operatorname{tr}(A)$ (Note que o traço de uma matriz pode ser calculado usando-se o comando `trace` do MATLAB.)
(f) $B - 3I = S^{-1}(A - 3I)S$
(g) $\det(B - 3I) = \det(A - 3I)$

Essas propriedades são válidas em geral para qualquer par de matrizes semelhantes. Ver Exercícios de 11 a 15 na Seção 3.

CAPÍTULO 5

ORTOGONALIDADE

Podemos adicionar à estrutura de espaço vetorial um produto escalar ou produto interno. Tal produto não é uma multiplicação vetorial verdadeira, pois associa a cada par de vetores um escalar, e não um terceiro vetor. Por exemplo, em R^2, podemos definir o produto escalar de dois vetores $\mathbf{x}$ e $\mathbf{y}$ como sendo $\mathbf{x}^T\mathbf{y}$. Podemos imaginar os vetores em R^2 como segmentos de reta orientados com ponto inicial na origem. Não é difícil mostrar que o ângulo entre dois segmentos orientados é um ângulo reto se e somente se o produto escalar dos vetores correspondentes é nulo. Em geral, se V é um espaço vetorial munido de um produto escalar, então dois vetores em V são ditos *ortogonais* se seu produto escalar é nulo.

Podemos pensar em ortogonalidade como uma generalização do conceito de *perpendicularidade* em um espaço vetorial arbitrário munido de um produto interno. Para compreender o significado disso, considere o seguinte problema: dados uma reta l contendo a origem e um ponto Q não pertencente a l, encontre o ponto P em l mais próximo de Q. A solução P desse problema é caracterizada pela condição de que QP é perpendicular a OP (ver a Fig. 5.0.1). Se pensarmos na reta l como um subespaço de R^2 e em $\mathbf{v} = OQ$ como um vetor em R^2, então o problema é encontrar um vetor no subespaço o "mais perto" possível de $\mathbf{v}$. A solução $\mathbf{p}$ é caracterizada pela propriedade de que $\mathbf{p}$ é ortogonal a $\mathbf{v} - \mathbf{p}$ (ver Fig. 5.0.1). No contexto de espaços vetoriais munidos de um produto interno, podemos considerar problemas gerais de "mínimos quadráticos". Em tais problemas é dado um vetor $\mathbf{v}$ em V e um subespaço W, e queremos encontrar um vetor em W o "mais próximo" possível de $\mathbf{v}$. Uma solução $\mathbf{p}$ tem que ser ortogonal a $\mathbf{v} - \mathbf{p}$. Essa condição de ortogonalidade nos dá a chave para resolver o problema de mínimos quadráticos. Problemas de mínimos quadráticos aparecem em muitas aplicações estatísticas envolvendo a escolha de curvas que melhor se adaptam a um conjunto de dados.

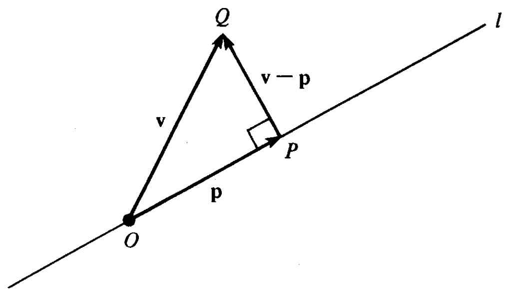


FIG. 5.0.1

1 O PRODUTO ESCALAR EM R^n

Dois vetores **x** e **y** em R^n podem ser considerados como matrizes $n \times 1$. Podemos, então, formar o produto matricial $\mathbf{x}^T\mathbf{y}$. O resultado desse produto é uma matriz 1×1 que pode ser considerada como um vetor em R^1 ou, mais simplesmente, um número real. O produto $\mathbf{x}^T\mathbf{y}$ é chamado de *produto escalar* de **x** e **y**. Em particular, se $\mathbf{x} = (x_1, \ldots, x_n)^T$ e $\mathbf{y} = (y_1, \ldots, y_n)^T$, então

$$\mathbf{x}^T\mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

EXEMPLO 1. Se

$$\mathbf{x} = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} \quad \text{e} \quad \mathbf{y} = \begin{pmatrix} 4 \\ 3 \\ 2 \end{pmatrix}$$

então

$$\mathbf{x}^T\mathbf{y} = (3, -2, 1)\begin{pmatrix} 4 \\ 3 \\ 2 \end{pmatrix} = 3 \cdot 4 - 2 \cdot 3 + 1 \cdot 2 = 8$$

□

O PRODUTO ESCALAR EM R^2 E R^3

Para compreender o significado geométrico do produto escalar, vamos restringir nossa atenção a R^2 e R^3. Vetores em R^2 e R^3 podem ser representados por segmentos de reta orientados. Dado um vetor **x** em R^2 ou R^3, seu comprimento euclidiano pode ser definido em termos do produto escalar.

$$\|\mathbf{x}\| = (\mathbf{x}^T\mathbf{x})^{1/2} = \begin{cases} \sqrt{x_1^2 + x_2^2} & \text{se } \mathbf{x} \in R^2 \\ \sqrt{x_1^2 + x_2^2 + x_3^2} & \text{se } \mathbf{x} \in R^3 \end{cases}$$

Dados dois vetores não-nulos **x** e **y**, podemos imaginá-los como segmentos orientados com o mesmo ponto inicial. O ângulo entre esses dois vetores é definido, então, como o ângulo entre os segmentos.

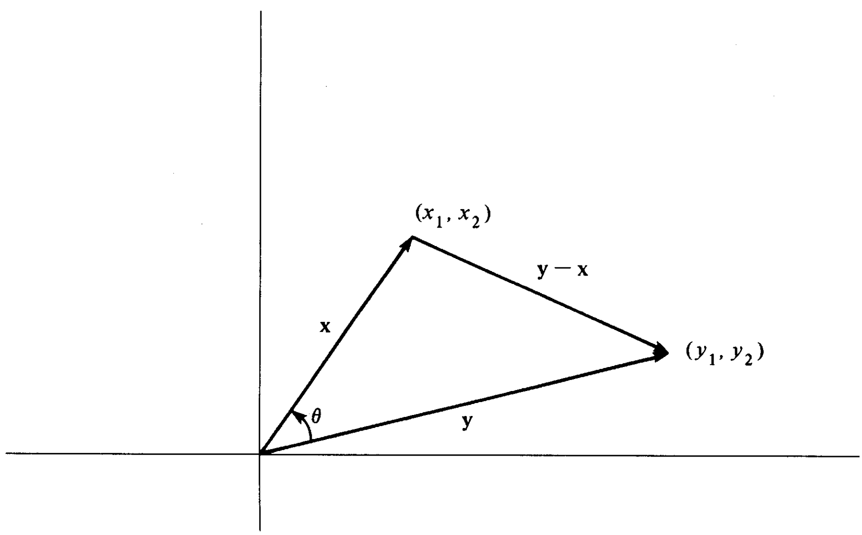


FIG. 5.1.1

Teorema 5.1.1. *Se* **x** *e* **y** *são dois vetores não-nulos em* R^2 *ou* R^3 *e se* θ *é o ângulo entre eles, então*

(1) $$\mathbf{x}^T\mathbf{y} = \|\mathbf{x}\|\,\|\mathbf{y}\|\cos\theta$$

Demonstração. Vamos provar o resultado para R^2. A demonstração para R^3 é análoga. Podemos usar os vetores **x**, **y** e **y** - **x** para formar um triângulo como na Fig. 5.1.1. Pela lei dos co-senos,

$$\|\mathbf{y}-\mathbf{x}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\,\|\mathbf{y}\|\cos\theta$$

ou

$$\begin{aligned}\|\mathbf{x}\|\,\|\mathbf{y}\|\cos\theta &= \tfrac{1}{2}(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - \|\mathbf{y}-\mathbf{x}\|^2)\\ &= \tfrac{1}{2}\left[x_1^2 + x_2^2 + y_1^2 + y_2^2 - (x_1-y_1)^2 - (x_2-y_2)^2\right]\\ &= x_1y_1 + x_2y_2\\ &= \mathbf{x}^T\mathbf{y}\end{aligned}$$

□

Corolário 5.1.2 (Desigualdade de Cauchy-Schwarz). *Se* **x** *e* **y** *são vetores em* R^2 *ou* R^3, *então*

(2) $$|\mathbf{x}^T\mathbf{y}| \leq \|\mathbf{x}\|\,\|\mathbf{y}\|$$

e a igualdade é válida se e somente se um dos vetores é igual a **0** *ou se um dos vetores é um múltiplo do outro.*

Demonstração. A desigualdade segue de (1). Se um dos vetores é igual a **0**, então os dois lados de (2) são iguais a 0. Se ambos os vetores são não-nulos, então, por (1), a igualdade é válida se e somente se $\cos\theta = \pm 1$. Mas isso significa que os vetores ou são iguais ou têm sentidos opostos, logo um dos vetores tem que ser um múltiplo do outro. □

Se $\mathbf{x}^T\mathbf{y} = 0$, pelo Teorema 5.1.1, ou um dos vetores é nulo ou $\cos\theta = 0$. Se $\cos\theta = 0$, o ângulo entre os vetores é um ângulo reto.

Definição. Os vetores **x** e **y** em R^2 (ou R^3) são ditos **ortogonais** se $\mathbf{x}^T\mathbf{y} = 0$.

EXEMPLO 2

(a) O vetor **0** é ortogonal a todos os vetores em R^2.

(b) Os vetores $\begin{pmatrix}3\\2\end{pmatrix}$ e $\begin{pmatrix}-4\\6\end{pmatrix}$ são ortogonais em R^2.

(c) Os vetores $\begin{pmatrix}2\\-3\\1\end{pmatrix}$ e $\begin{pmatrix}1\\1\\1\end{pmatrix}$ são ortogonais em R^3.

□

PROJEÇÕES ESCALARES E VETORIAIS

O produto escalar pode ser usado para se encontrar a componente de um vetor na direção de outro. Sejam **x** e **y** vetores não-nulos em R^2 ou R^3. Gostaríamos de escrever **x** como uma soma da forma **p** + **z**, onde **p** tem a mesma direção que **y** e **z** é ortogonal a **p** (ver Fig. 5.1.2). Para fazer isso, defina $\mathbf{u} = (1/\|\mathbf{y}\|)\mathbf{y}$. Então, **u** é um vetor unitário (comprimento 1) com mesma direção e mesmo sentido que **y**. Queremos encontrar α tal que $\mathbf{p} = \alpha\mathbf{u}$ seja ortogonal a $\mathbf{z} = \mathbf{x} - \alpha\mathbf{u}$. Para que **p** e **z** sejam ortogonais, o escalar α tem que satisfazer

$$\begin{aligned}\alpha &= \|\mathbf{x}\|\cos\theta\\ &= \frac{\|\mathbf{x}\|\,\|\mathbf{y}\|\cos\theta}{\|\mathbf{y}\|}\\ &= \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{y}\|}\end{aligned}$$

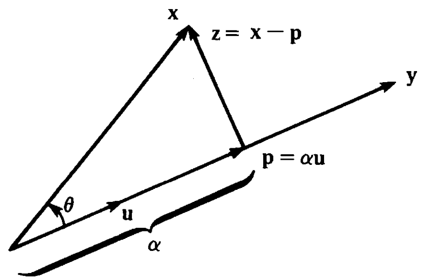


FIG. 5.1.2

O escalar α é chamado de *projeção escalar* de **x** sobre **y**, e o vetor **p** é chamado de *projeção vetorial* de **x** sobre **y**.

Projeção escalar de **x** sobre **y**:

$$\alpha = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{y}\|}$$

Projeção vetorial de **x** sobre **y**:

$$\mathbf{p} = \alpha\mathbf{u} = \alpha\frac{1}{\|\mathbf{y}\|}\mathbf{y} = \frac{\mathbf{x}^T\mathbf{y}}{\mathbf{y}^T\mathbf{y}}\mathbf{y}$$

EXEMPLO 3. Encontre o ponto mais próximo do ponto (1, 4) que pertence à reta $y = (1/3)x$ (ver Fig. 5.1.3).

SOLUÇÃO. O vetor $\mathbf{w} = (3, 1)^T$ é um vetor na direção da reta $y = (1/3)x$. Seja $\mathbf{v} = (1, 4)^T$. Se Q é o ponto desejado, então Q^T é a projeção vetorial de **v** sobre **w**.

$$Q^T = \left(\frac{\mathbf{v}^T\mathbf{w}}{\mathbf{w}^T\mathbf{w}}\right)\mathbf{w} = \frac{7}{10}\begin{pmatrix}3\\1\end{pmatrix} = \begin{pmatrix}2{,}1\\0{,}7\end{pmatrix}$$

Logo, $Q = (2{,}1, 0{,}7)$ é o ponto mais próximo. □

Notação. Se P_1 e P_2 são dois pontos no espaço tridimensional, denotaremos por $\overline{P_1P_2}$ o vetor de P_1 a P_2.

Se **N** é um vetor não-nulo e P_0 é um ponto fixo, o conjunto de pontos P tais que $\overline{P_0P}$ é ortogonal a **N** forma um plano π no espaço tridimensional que contém P_0. O vetor **N** e o plano π são ditos *normais* entre si. Um ponto $P = (x, y, z)$ pertence a π se e somente se

$$(\overline{P_0P})^T\mathbf{N} = 0$$

Se $\mathbf{N} = (a, b, c)^T$ e $P_0 = (x_0, y_0, z_0)$, essa equação pode ser colocada na forma

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

EXEMPLO 4. Encontre a equação do plano contendo o ponto (2, −1, 3) e normal ao vetor $\mathbf{N} = (2, 3, 4)^T$.

SOLUÇÃO. $\overline{P_0P} = (x - 2, y + 1, z - 3)^T$. A equação é $(\overline{P_0P})^T\mathbf{N} = 0$, ou

$$2(x - 2) + 3(y + 1) + 4(z - 3) = 0$$ □

EXEMPLO 5. Encontre a distância do ponto (2, 0, 0) ao plano $x + 2y + 2z = 0$.

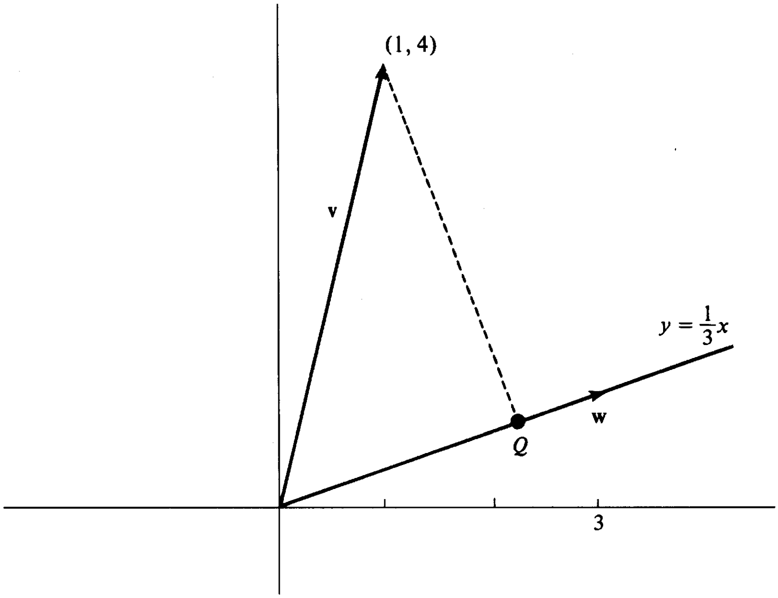


FIG. 5.1.3

SOLUÇÃO. O vetor $\mathbf{N} = (1, 2, 2)^T$ é normal ao plano e o plano contém a origem. Seja $\mathbf{v} = (2, 0, 0)^T$. A distância d de $(2, 0, 0)$ ao plano é, simplesmente, o valor absoluto da projeção escalar de $\mathbf{v}$ sobre $\mathbf{N}$. Logo,

$$d = \frac{|\mathbf{v}^T\mathbf{N}|}{\|\mathbf{N}\|} = \frac{2}{3}$$ □

ORTOGONALIDADE EM R^n

Todas as definições que foram dadas para R^2 e R^3 podem ser generalizadas para R^n. De fato, se $\mathbf{x} \in R^n$, o comprimento euclidiano de $\mathbf{x}$ é definido por

$$\|\mathbf{x}\| = (\mathbf{x}^T\mathbf{x})^{1/2} = (x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2}$$

e o ângulo θ entre dois vetores não-nulos $\mathbf{x}$ e $\mathbf{y}$ é dado por

$$\cos\theta = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\ \|\mathbf{y}\|}$$

Os vetores $\mathbf{x}$ e $\mathbf{y}$ são ditos *ortogonais* se $\mathbf{x}^T\mathbf{y} = 0$. O símbolo "$\perp$" é utilizado muitas vezes para indicar ortogonalidade. Então, se $\mathbf{x}$ e $\mathbf{y}$ são ortogonais, escreveremos $\mathbf{x} \perp \mathbf{y}$. As projeções vetoriais e escalares são definidas em R^n da mesma maneira que em R^2. Uma das aplicações principais desses conceitos é a solução de problemas de mínimos quadráticos. Estudaremos problemas de mínimos quadráticos na Seção 4.

EXERCÍCIOS

1. Encontre o ângulo entre cada par de vetores $\mathbf{v}$ e $\mathbf{w}$ a seguir.

(a) $\mathbf{v} = (2, 1, 3)^T$, $\quad \mathbf{w} = (6, 3, 9)^T$
(b) $\mathbf{v} = (2, -3)^T$, $\quad \mathbf{w} = (3, 2)^T$
(c) $\mathbf{v} = (4, 1)^T$, $\quad \mathbf{w} = (3, 2)^T$
(d) $\mathbf{v} = (-2, 3, 1)^T$, $\quad \mathbf{w} = (1, 2, 4)^T$

2. Para cada par de vetores no Exercício 1, encontre a projeção escalar de $\mathbf{v}$ sobre $\mathbf{w}$. Encontre, também, a projeção vetorial de $\mathbf{v}$ sobre $\mathbf{w}$.
3. Para cada par de vetores $\mathbf{x}$ e $\mathbf{y}$ a seguir, encontre a projeção $\mathbf{p}$ de $\mathbf{x}$ sobre $\mathbf{y}$ e verifique que $\mathbf{p}$ e $\mathbf{x} - \mathbf{p}$ são ortogonais.

   (a) $\mathbf{x} = (3, 4)^T$, $\quad \mathbf{y} = (1, 0)^T$
   (b) $\mathbf{x} = (3, 5)^T$, $\quad \mathbf{y} = (1, 1)^T$
   (c) $\mathbf{x} = (2, 4, 3)^T$, $\quad \mathbf{y} = (1, 1, 1)^T$
   (d) $\mathbf{x} = (2, -5, 4)^T$, $\quad \mathbf{y} = (1, 2, -1)^T$

4. Encontre o ponto mais próximo de (5, 2) que pertence à reta $y = 2x$.
5. Encontre o ponto mais próximo de (5, 2) que pertence à reta $y = 2x + 1$.
6. Encontre a distância do ponto (1, 2) à reta $4x - 3y = 0$.
7. Em cada um dos itens a seguir, encontre a equação do plano normal ao vetor $\mathbf{N}$ dado que contém o ponto P_0.

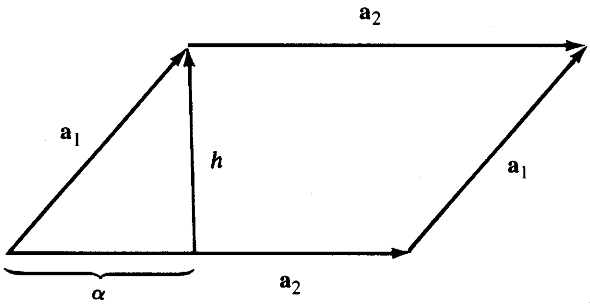
   (a) $\mathbf{N} = (2, 4, 3)^T$, $\quad P_0 = (0, 0, 0)$
   (b) $\mathbf{N} = (-3, 6, 2)^T$, $\quad P_0 = (4, 2, -5)$
   (c) $\mathbf{N} = (0, 0, 1)^T$, $\quad P_0 = (3, 2, 4)$

8. Encontre a distância do ponto (1, 1, 1) ao plano $2x + 2y + z = 0$.
9. Encontre a distância do ponto (2, 1, −2) ao plano $6(x - 1) + 2(y - 3) + 3(z + 4) = 0$.
10. Se $\mathbf{x} = (x_1, x_2)^T$, $\mathbf{y} = (y_1, y_2)^T$ e $\mathbf{z} = (z_1, z_2)^T$ são vetores arbitrários em R^2, prove que:

   (a) $\mathbf{x}^T\mathbf{x} \geq 0$ $\qquad$ (b) $\mathbf{x}^T\mathbf{y} = \mathbf{y}^T\mathbf{x}$
   (c) $\mathbf{x}^T(\mathbf{y} + \mathbf{z}) = \mathbf{x}^T\mathbf{y} + \mathbf{x}^T\mathbf{z}$

11. Se $\mathbf{u}$ e $\mathbf{v}$ são dois vetores quaisquer em R^2, mostre que $\|\mathbf{u} + \mathbf{v}\|^2 \leq (\|\mathbf{u}\| + \|\mathbf{v}\|)^2$ e, portanto, $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$. Quando a igualdade é válida? Interprete geometricamente essa desigualdade.
12. Sejam l_1 a reta $y = m_1x + b_1$, $m_1 \neq 0$, e l a reta $y = mx + b$. Mostre que l é perpendicular a l_1 se e somente se $m = -1/m_1$.
13. Sejam $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ vetores em R^3. Se $\mathbf{x}_1 \perp \mathbf{x}_2$ e $\mathbf{x}_2 \perp \mathbf{x}_3$, é necessariamente verdade que $\mathbf{x}_1 \perp \mathbf{x}_3$? Prove.
14. Seja A uma matriz 2×2 com vetores colunas $\mathbf{a}_1$ e $\mathbf{a}_2$ linearmente independentes. Se $\mathbf{a}_1$ e $\mathbf{a}_2$ são usados para formar um paralelogramo P de altura h (ver a figura seguir), mostre que:

   (a) $h^2\|\mathbf{a}_2\|^2 = \|\mathbf{a}_1\|^2\|\mathbf{a}_2\|^2 - (\mathbf{a}_1^T\mathbf{a}_2)^2$
   (b) Área de $P = |\det(A)|$



2 SUBESPAÇOS ORTOGONAIS

Seja A uma matriz $m \times n$ e seja $\mathbf{x} \in N(A)$. Como $A\mathbf{x} = \mathbf{0}$, temos

(1) $$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = 0$$

para $i = 1, \ldots, m$. A Equação (1) diz que $\mathbf{x}$ é ortogonal à i-ésima coluna de A^T para $i = 1, \ldots, m$. Como $\mathbf{x}$ é ortogonal a cada coluna de A^T, ele é ortogonal a qualquer combinação linear dos vetores colunas de A^T. Então, se $\mathbf{y}$ é um vetor arbitrário no espaço coluna de A^T, $\mathbf{x}^T\mathbf{y} = 0$. Logo, cada vetor em $N(A)$ é ortogonal a todos os vetores no espaço coluna de A^T. Quando dois subespaços de R^n têm essa propriedade, dizemos que eles são ortogonais.

Definição. Dois subespaços X e Y de R^n são ditos **ortogonais** se $\mathbf{x}^T\mathbf{y} = 0$, quaisquer que sejam $\mathbf{x} \in X$ e $\mathbf{y} \in Y$. Se X e Y são ortogonais, escrevemos $X \perp Y$.

EXEMPLO 1. Seja X o subespaço de R^3 gerado por $\mathbf{e}_1$ e seja Y o subespaço gerado por $\mathbf{e}_2$. Se $\mathbf{x} \in X$ e $\mathbf{y} \in Y$, esses vetores têm que ser da forma

$$\mathbf{x} = \begin{pmatrix} x_1 \\ 0 \\ 0 \end{pmatrix} \quad \text{e} \quad \mathbf{y} = \begin{pmatrix} 0 \\ y_2 \\ 0 \end{pmatrix}$$

Logo,

$$\mathbf{x}^T\mathbf{y} = x_1 \cdot 0 + 0 \cdot y_2 + 0 \cdot 0 = 0$$

Portanto, $X \perp Y$. □

O conceito de espaços ortogonais nem sempre coincide com a nossa idéia intuitiva de perpendicularidade. Por exemplo, o chão e a parede da sala de aula "parecem" ortogonais, mas os planos xy e yz não são subespaços ortogonais. De fato, podemos considerar os vetores $\mathbf{x}_1 = (1, 1, 0)^T$ e $\mathbf{x}_2 = (0, 1, 1)^T$ como pertencentes aos planos xy e yz, respectivamente. Como

$$\mathbf{x}_1^T\mathbf{x}_2 = 1 \cdot 0 + 1 \cdot 1 + 0 \cdot 1 = 1$$

esses subespaços não são ortogonais. O próximo exemplo mostra que o subespaço correspondente ao eixo dos z é ortogonal ao subespaço correspondente ao plano xy.

EXEMPLO 2. Seja X o subespaço de R^3 gerado por $\mathbf{e}_1$ e $\mathbf{e}_2$, e seja Y o subespaço gerado por $\mathbf{e}_3$. Se $\mathbf{x} \in X$ e $\mathbf{y} \in Y$, então

$$\mathbf{x}^T\mathbf{y} = x_1 \cdot 0 + x_2 \cdot 0 + 0 \cdot y_3 = 0$$

Logo, $X \perp Y$. Além disso, se $\mathbf{z}$ é um vetor qualquer em R^3 que é ortogonal a todos os vetores em Y, então $\mathbf{z} \perp \mathbf{e}_3$ e, portanto,

$$z_3 = \mathbf{z}^T\mathbf{e}_3 = 0$$

Mas, se $z_3 = 0$, então $\mathbf{z} \in X$. Então, X é o conjunto de todos os vetores em R^3 que são ortogonais a todos os vetores em Y (ver Fig. 5.2.1). □

Definição. Seja Y um subespaço de R^n. O conjunto de todos os vetores em R^n que são ortogonais a todos os vetores em Y será denotado por $Y^\perp$. Então

$$Y^\perp = \left\{\mathbf{x} \in R^n \mid \mathbf{x}^T\mathbf{y} = 0 \quad \text{para todo} \quad \mathbf{y} \in Y\right\}$$

O conjunto $Y^\perp$ é chamado de **complemento ortogonal** de Y.

Nota. Os subespaços $X = [\{\mathbf{e}_1\}]$ e $Y = [\{\mathbf{e}_2\}]$ de R^3 dados no Exemplo 1 são ortogonais, mas nenhum deles é complemento ortogonal do outro. De fato,

$$X^\perp = [\{\mathbf{e}_2, \mathbf{e}_3\}] \quad \text{e} \quad Y^\perp = [\{\mathbf{e}_1, \mathbf{e}_3\}].$$

Observações

1. Se X e Y são subespaços ortogonais de R^n, então $X \cap Y = \{\mathbf{0}\}$.
2. Se Y é um subespaço de R^n, então $Y^\perp$ também é um subespaço de R^n.

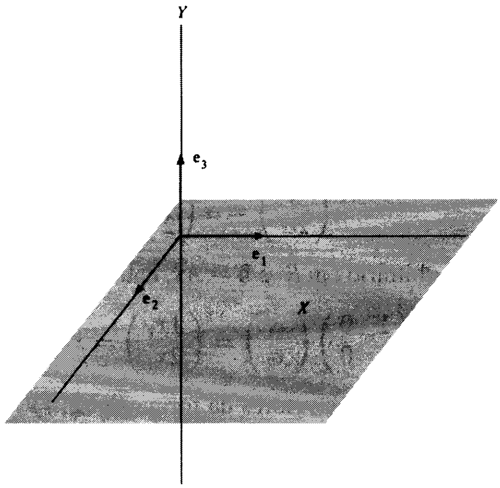


FIG. 5.2.1

Demonstração de (1). Se $\mathbf{x} \in X \cap Y$ e $X \perp Y$, então $\|\mathbf{x}\|^2 = \mathbf{x}^T\mathbf{x} = 0$ e, portanto, $\mathbf{x} = \mathbf{0}$. □

Demonstração de (2). Se $\mathbf{x} \in Y^\perp$ e α é um escalar, então, para todo $\mathbf{y} \in Y$,

$$(\alpha\mathbf{x})^T\mathbf{y} = \alpha(\mathbf{x}^T\mathbf{y}) = \alpha \cdot 0 = 0$$

Logo, $\alpha\mathbf{x} \in Y^\perp$. Se $\mathbf{x}_1$ e $\mathbf{x}_2$ são elementos de $Y^\perp$, então

$$(\mathbf{x}_1 + \mathbf{x}_2)^T\mathbf{y} = \mathbf{x}_1^T\mathbf{y} + \mathbf{x}_2^T\mathbf{y} = 0 + 0 = 0$$

para todo $\mathbf{y} \in Y$. Logo, $\mathbf{x}_1 + \mathbf{x}_2 \in Y^\perp$. Portanto, $Y^\perp$ é um subespaço de R^n. □

SUBESPAÇOS FUNDAMENTAIS

Seja A uma matriz $m \times n$. Vimos, no Cap. 3, que um vetor $\mathbf{b} \in R^m$ está no espaço coluna de A se e somente se $\mathbf{b} = A\mathbf{x}$ para algum $\mathbf{x} \in R^n$. Se pensarmos em A como um operador de R^n em R^m, então o espaço coluna de A é a mesma coisa que a imagem de A. Vamos denotar a imagem de A por $I(A)$. Então,

$$\begin{aligned} I(A) &= \{\mathbf{b} \in R^m \mid \mathbf{b} = A\mathbf{x} \text{ para algum } \mathbf{x} \in R^n\} \\ &= \text{o espaço coluna de } A. \end{aligned}$$

O espaço coluna de A^T, $I(A^T)$, é um subespaço de R^n:

$$I(A^T) = \{\mathbf{y} \in R^n \mid \mathbf{y} = A^T\mathbf{x} \text{ para algum } \mathbf{x} \in R^m\}.$$

O espaço coluna $I(A^T)$ é, essencialmente, o mesmo que o espaço linha de A, exceto que ele é formado por vetores em R^n (matrizes $n \times 1$), em vez de n-uplas. Logo, $\mathbf{y} \in I(A^T)$ se e somente se $\mathbf{y}^T$ pertence ao espaço linha de A. Vimos que $I(A^T) \perp N(A)$. O próximo teorema diz que, de fato, $N(A)$ é o complemento ortogonal de $I(A^T)$.

Teorema 5.2.1. *Se A é uma matriz $m \times n$, então $N(A) = I(A^T)^\perp$ e $N(A^T) = I(A)^\perp$.*

Demonstração. Já vimos que $N(A) \perp I(A^T)$ e isso implica que $N(A) \subset I(A^T)^\perp$. Por outro lado, se $\mathbf{x}$ é um vetor qualquer em $I(A^T)^\perp$, então $\mathbf{x}$ é ortogonal a cada uma das colunas de A^T e, portanto, $A\mathbf{x} = \mathbf{0}$. Logo, $\mathbf{x}$ tem que ser um elemento de $N(A)$, o que mostra que $N(A) = I(A^T)^\perp$. Essa demonstração não depende das dimensões de A. Em particular, o resultado também é válido para a matriz $B = A^T$. Logo,

$$N(A^T) = N(B) = R(B^T)^\perp = R(A)^\perp$$ □

EXEMPLO 3. Seja

$$A = \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}$$

O espaço coluna de A consiste em todos os vetores da forma

$$\begin{pmatrix} \alpha \\ 2\alpha \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Observe que, se $\mathbf{x}$ é um vetor arbitrário em R^n e $\mathbf{b} = A\mathbf{x}$, então

$$\mathbf{b} = \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1x_1 \\ 2x_1 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

O núcleo de A^T é o conjunto de todos os vetores da forma $\beta(-2, 1)^T$. Como $(1, 2)^T$ e $(-2, 1)^T$ são ortogonais, todo vetor em $I(A)$ vai ser ortogonal a todos os vetores em $N(A^T)$. A mesma relação é válida para $I(A^T)$ e $N(A)$. $I(A^T)$ é o conjunto dos vetores da forma $\alpha\mathbf{e}_1$ e $N(A)$ é o conjunto dos vetores da forma $\beta\mathbf{e}_2$. Como $\mathbf{e}_1$ e $\mathbf{e}_2$ são ortogonais, cada vetor em $I(A^T)$ é ortogonal a todos os vetores em $N(A)$. □

O Teorema 5.2.1 é um dos teoremas mais importantes neste capítulo. Na Seção 4, veremos que o resultado $N(A^T) = I(A)^\perp$ nos fornece a chave para resolver problemas de mínimos quadráticos. No momento, vamos usar o Teorema 5.2.1 para provar o teorema a seguir, o qual, por sua vez, será usado para se obterem dois resultados importantes sobre subespaços ortogonais.

Teorema 5.2.2. *Se S é um subespaço de R^n, então* $\dim S + \dim S^\perp = n$. *Além disso, se $\{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$ é uma base para S e se $\{\mathbf{x}_{r+1}, \ldots, \mathbf{x}_n\}$ é uma base para $S^\perp$, então $\{\mathbf{x}_1, \ldots, \mathbf{x}_r, \mathbf{x}_{r+1}, \ldots, \mathbf{x}_n\}$ é uma base para R^n.*

Demonstração. Se $S = \{\mathbf{0}\}$, então $S^\perp = R^n$ e

$$\dim S + \dim S^\perp = 0 + n = n$$

Se $S \neq \{\mathbf{0}\}$, seja $\{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$ uma base para S e defina uma matriz X $r \times n$ cuja i-ésima linha é $\mathbf{x}_i^T$ para cada i. Por construção, a matriz X tem posto r e $I(X^T) = S$. Pelo Teorema 5.2.1,

$$S^\perp = R(X^T)^\perp = N(X)$$

Pelo Teorema 3.6.4, temos que

$$\dim S^\perp = \dim N(X) = n - r$$

Para mostrar que $\{\mathbf{x}_1, \ldots, \mathbf{x}_r, \mathbf{x}_{r+1}, \ldots, \mathbf{x}_n\}$ é uma base para R^n, basta mostrar que os n vetores são linearmente independentes. Suponha que

$$c_1\mathbf{x}_1 + \cdots + c_r\mathbf{x}_r + c_{r+1}\mathbf{x}_{r+1} + \cdots + c_n\mathbf{x}_n = \mathbf{0}$$

Sejam $\mathbf{y} = c_1\mathbf{x}_1 + \ldots + c_r\mathbf{x}_r$ e $\mathbf{z} = c_{r+1}\mathbf{x}_{r+1} + \ldots + c_n\mathbf{x}_n$. Temos

$$\mathbf{y} + \mathbf{z} = \mathbf{0}$$

$$\mathbf{y} = -\mathbf{z}$$

Então ambos, $\mathbf{y}$ e $\mathbf{z}$, são elementos de $S \cap S^\perp$. Mas $S \cap S^\perp = \{\mathbf{0}\}$, logo

$$c_1\mathbf{x}_1 + \cdots + c_r\mathbf{x}_r = \mathbf{0}$$

$$c_{r+1}\mathbf{x}_{r+1} + \cdots + c_n\mathbf{x}_n = \mathbf{0}$$

Como $\mathbf{x}_1, \ldots, \mathbf{x}_r$ são linearmente independentes,

$$c_1 = c_2 = \cdots = c_r = 0$$

Analogamente, $\mathbf{x}_{r+1}, \ldots, \mathbf{x}_n$ são linearmente independentes e

$$c_{r+1} = c_{r+2} = \cdots = c_n = 0$$

Logo, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ são linearmente independentes e, portanto, formam uma base para R^n. □

Dado um subespaço S de R^n, vamos usar o Teorema 5.2.2 para provar que cada $\mathbf{x} \in R^n$ pode ser expresso de maneira única como uma soma $\mathbf{y} + \mathbf{z}$, onde $\mathbf{y} \in S$ e $\mathbf{z} \in S^\perp$.

Definição. Se U e V são subespaços de um espaço vetorial W e se cada $\mathbf{w} \in W$ pode ser expresso de maneira única como uma soma $\mathbf{u} + \mathbf{v}$, onde $\mathbf{u} \in U$ e $\mathbf{v} \in V$, dizemos que W é a **soma direta** de U e V e escrevemos $W = U \oplus V$.

Teorema 5.2.3. *Se S é um subespaço de R^n, então*

$$R^n = S \oplus S^\perp$$

Demonstração. O resultado é trivial se $S = \{\mathbf{0}\}$ ou $S = R^n$. No caso em que $\dim S = r$, $0 < r < n$, segue, do Teorema 5.2.2, que cada vetor $\mathbf{x} \in R^n$ pode ser representado na forma

$$\mathbf{x} = c_1\mathbf{x}_1 + \cdots + c_r\mathbf{x}_r + c_{r+1}\mathbf{x}_{r+1} + \cdots + c_n\mathbf{x}_n$$

onde $\{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$ é uma base para S e $\{\mathbf{x}_{r+1}, \ldots, \mathbf{x}_n\}$ é uma base para $S^\perp$. Definindo

$$\mathbf{u} = c_1\mathbf{x}_1 + \cdots + c_r\mathbf{x}_r \qquad \text{e} \qquad \mathbf{v} = c_{r+1}\mathbf{x}_{r+1} + \cdots + c_n\mathbf{x}_n$$

temos que $\mathbf{u} \in S$, $\mathbf{v} \in S^\perp$ e $\mathbf{x} = \mathbf{u} + \mathbf{v}$. Para mostrar a unicidade, suponha que $\mathbf{x}$ também pode ser escrito como uma soma $\mathbf{y} + \mathbf{z}$, onde $\mathbf{y} \in S$ e $\mathbf{z} \in S^\perp$. Então,

$$\mathbf{u} + \mathbf{v} = \mathbf{x} = \mathbf{y} + \mathbf{z}$$

$$\mathbf{u} - \mathbf{y} = \mathbf{z} - \mathbf{v}$$

Mas $\mathbf{u} - \mathbf{y} \in S$ e $\mathbf{z} - \mathbf{v} \in S^\perp$, logo ambos pertencem a $S \cap S^\perp$. Como

$$S \cap S^\perp = \{\mathbf{0}\}$$

temos que

$$\mathbf{u} = \mathbf{y} \qquad \text{e} \qquad \mathbf{v} = \mathbf{z}$$ □

Teorema 5.2.4. *Se S é um subespaço de R^n, então $(S^\perp)^\perp = S$.*

Demonstração. Se $\mathbf{x} \in S$, então $\mathbf{x}$ é ortogonal a todos os $\mathbf{y}$ em $S^\perp$, logo $\mathbf{x} \in (S^\perp)^\perp$. Portanto, $S \subset (S^\perp)^\perp$. Por outro lado, suponha que $\mathbf{z}$ é um elemento arbitrário de $(S^\perp)^\perp$. Pelo Teorema 5.2.3, podemos escrever $\mathbf{z}$ como uma soma $\mathbf{u} + \mathbf{v}$, onde $\mathbf{u} \in S$ e $\mathbf{v} \in S^\perp$. Como $\mathbf{v} \in S^\perp$, $\mathbf{v}$ é ortogonal a ambos $\mathbf{u}$ e $\mathbf{z}$. Temos, então, que

$$0 = \mathbf{v}^T\mathbf{z} = \mathbf{v}^T\mathbf{u} + \mathbf{v}^T\mathbf{v} = \mathbf{v}^T\mathbf{v}$$

e, em conseqüência, $\mathbf{v} = \mathbf{0}$. Portanto, $\mathbf{z} = \mathbf{u} \in S$ e $S = (S^\perp)^\perp$. □

Pelo Teorema 5.2.4, se T é o complemento ortogonal de S, então S é o complemento ortogonal de T, e podemos dizer simplesmente que S e T são complementos ortogonais um do outro. Em particular, pelo Teorema 5.2.1., $N(A)$ e $I(A^T)$ são complementos ortogonais um do outro, assim como $N(A^T)$ e $I(A)$. Podemos, então, escrever

$$N(A)^\perp = R(A^T) \qquad \text{e} \qquad N(A^T)^\perp = R(A)$$

Lembre-se de que o sistema $A\mathbf{x} = \mathbf{b}$ é compatível se e somente se $\mathbf{b} \in I(A)$. Como $I(A) = N(A^T)^\perp$, temos o seguinte resultado, que pode ser considerado como um corolário do Teorema 5.2.1.

Corolário 5.2.5. *Se A é uma matriz $m \times n$ e $\mathbf{b} \in R^n$, então, ou existe um vetor $\mathbf{x} \in R^n$ tal que $A\mathbf{x} = \mathbf{b}$, ou existe um vetor $\mathbf{y} \in R^m$ tal que $A^T\mathbf{y} = \mathbf{0}$ e $\mathbf{y}^T\mathbf{b} \neq 0$.*

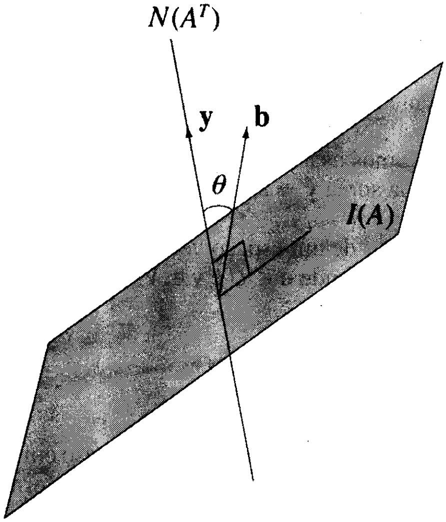


FIG. 5.2.2

O Corolário 5.2.5 está ilustrado na Fig. 5.2.2 para o caso em que $I(A)$ é um subespaço bidimensional de R^3. O ângulo θ na figura vai ser um ângulo reto se e somente se $\mathbf{b} \in I(A)$.

EXEMPLO 4. Seja

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

Encontre bases para $N(A)$, $I(A^T)$, $N(A^T)$ e $I(A)$.

SOLUÇÃO. Podemos encontrar bases para $N(A)$ e $I(A^T)$ colocando A em sua forma escada reduzida por linhas.

$$\begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 1 & 3 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Como $(1, 0, 1)$ e $(0, 1, 1)$ formam uma base para o espaço linha de A, temos que $(1, 0, 1)^T$ e $(0, 1, 1)^T$ formam uma base para $I(A^T)$. Se $\mathbf{x} \in N(A)$, da forma escada reduzida por linhas de A temos que

$$x_1 + x_3 = 0$$

$$x_2 + x_3 = 0$$

Logo,

$$x_1 = x_2 = -x_3$$

Fazendo $x_3 = \alpha$, vemos que $N(A)$ é formado por todos os vetores da forma $\alpha(-1, -1, 1)^T$. Observe que $(-1, -1, 1)^T$ é ortogonal a $(1, 0, 1)^T$ e a $(0, 1, 1)^T$.

Para encontrar bases para $I(A)$ e $N(A^T)$, coloque A^T em sua forma escada reduzida por linhas.

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 3 \\ 2 & 1 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Logo, $(1, 0, 1)^T$ e a $(0, 1, 2)^T$ formam uma base para $I(A)$. Se $\mathbf{x} \in N(A^T)$, então $x_1 = -x_3$, $x_2 = -2x_3$. Logo, $N(A^T)$ é o subespaço de R^3 gerado por $(-1, -2, 1)^T$. Observe que $(-1, -2, 1)^T$ é ortogonal a $(1, 0, 1)^T$ e a $(0, 1, 2)^T$. □

Vimos, no Cap. 3, que os espaços linha e coluna têm a mesma dimensão. Se A tem posto r, então

$$\dim I(A) = \dim I(A^T) = r$$

De fato, A pode ser usada para se estabelecer uma bijeção entre $I(A^T)$ e $I(A)$.

Podemos considerar uma matriz A $m \times n$ como uma transformação linear de R^n em R^m.

$$\mathbf{x} \in R^n \rightarrow A\mathbf{x} \in R^m$$

Como $I(A^T)$ e $N(A)$ são complementos ortogonais em R^n, temos

$$R^n = I(A^T) \oplus N(A)$$

Cada vetor em R^n pode ser escrito como uma soma

$$\mathbf{x} = \mathbf{y} + \mathbf{z}, \qquad \mathbf{y} \in I(A^T), \qquad \mathbf{z} \in N(A)$$

Logo,

$$A\mathbf{x} = A\mathbf{y} + A\mathbf{z} = A\mathbf{y} \quad \text{para cada } \mathbf{x} \in R^n$$

e, portanto,

$$I(A) = \{A\mathbf{x} \mid \mathbf{x} \in R^n\} = \{A\mathbf{y} \mid \mathbf{y} \in I(A^T)\}$$

Então, restringindo o domínio de A a $I(A^T)$, obtemos uma aplicação sobrejetora de $I(A^T)$ em $I(A)$. Além disso, essa aplicação é injetora. De fato, se $\mathbf{x}_1, \mathbf{x}_2 \in I(A^T)$ e

$$A\mathbf{x}_1 = A\mathbf{x}_2$$

então

$$A(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0}$$

e, portanto,

$$\mathbf{x}_1 - \mathbf{x}_2 \in I(A^T) \cap N(A)$$

Como $I(A^T) \cap N(A) = \{\mathbf{0}\}$, temos que $\mathbf{x}_1 = \mathbf{x}_2$. Logo, A determina uma bijeção entre $I(A^T)$ e $I(A)$. Como cada $\mathbf{b} \in I(A)$ corresponde a exatamente um $\mathbf{y} \in I(A^T)$, podemos definir uma transformação inversa de $I(A)$ a $I(A^T)$. De fato, toda matriz A $m \times n$ é invertível quando considerada como uma transformação linear de $I(A^T)$ para $I(A)$.

EXEMPLO 5. Seja $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}$. $I(A^T)$ é gerado por $\mathbf{e}_1$ e $\mathbf{e}_2$ e $N(A)$ é gerado por $\mathbf{e}_3$. Qualquer vetor $\mathbf{x} \in R^3$ pode ser escrito como uma soma

$$\mathbf{x} = \mathbf{y} + \mathbf{z}$$

onde

$$\mathbf{y} = (x_1, x_2, 0)^T \in I(A^T) \quad \text{e} \quad \mathbf{z} = (0, 0, x_3)^T \in N(A)$$

Restringindo-nos a vetores $\mathbf{y} \in I(A^T)$, temos

$$\mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \rightarrow A\mathbf{y} = \begin{pmatrix} 2x_1 \\ 3x_2 \end{pmatrix}$$

Nesse caso $I(A) = R^2$ e a transformação inversa de $I(A)$ em $I(A^T)$ é definida por

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \to \begin{pmatrix} \frac{1}{2}b_1 \\ \frac{1}{3}b_2 \\ 0 \end{pmatrix}$$

□

EXERCÍCIOS

1. Para cada uma das matrizes a seguir, determine uma base para cada um dos subespaços $I(A^T)$, $N(A)$, $I(A)$ e $N(A^T)$.

(a) $A = \begin{pmatrix} 3 & 4 \\ 6 & 8 \end{pmatrix}$ (b) $A = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 4 & 0 \end{pmatrix}$

(c) $A = \begin{pmatrix} 4 & -2 \\ 1 & 3 \\ 2 & 1 \\ 3 & 4 \end{pmatrix}$ (d) $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 2 & 2 \end{pmatrix}$

2. Seja S o subespaço de R^3 gerado por $\mathbf{x} = (1, -1, 1)^T$.
(a) Encontre uma base para $S^\perp$.
(b) Descreva geometricamente S e $S^\perp$.

3. (a) Seja S o subespaço de R^3 gerado pelos vetores $\mathbf{x}_1 = (x_1, x_2, x_3)^T$ e $\mathbf{y} = (y_1, y_2, y_3)^T$. Seja

$$A = \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}$$

Mostre que $S^\perp = N(A)$.
(b) Encontre o complemento ortogonal do subespaço de R^3 gerado por $(1, 2, 1)^T$ e $(1, -1, 2)^T$.

4. Seja S o subespaço de R^4 gerado por $\mathbf{x}_1 = (1, 0, -2, 1)^T$ e $\mathbf{x}_2 = (0, 1, 3, -2)^T$. Encontre uma base para $S^\perp$.

5. Sejam $P_1 = (1, 1, 1)$, $P_2 = (2, 4, -1)$ e $P_3 = (0, -1, 5)$.
(a) Encontre um vetor não-nulo $\mathbf{N}$ que é ortogonal a $\overline{P_1P_2}$ e $\overline{P_1P_3}$.
(b) Encontre a equação do plano determinado por esses três pontos.

6. É possível o espaço linha de uma matriz conter o vetor (3, 1, 2) e seu núcleo conter o vetor $(2, 1, 1)^T$? Explique.

7. Seja $\mathbf{a}_j$ um vetor coluna não-nulo de uma matriz A $m \times n$. É possível $\mathbf{a}_j$ pertencer a $N(A^T)$? Explique.

8. Seja S o subespaço de R^n gerado pelos vetores $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$. Mostre que $\mathbf{y} \in S^\perp$ se e somente se $\mathbf{y} \perp \mathbf{x}_i$ para $i = 1, \ldots, k$.

9. Se A é uma matriz $m \times n$ de posto r, quais são as dimensões de $N(A)$ e de $N(A^T)$? Explique.

10. Prove o Corolário 5.2.5.

11. Prove que, se A é uma matriz $m \times n$ e se $\mathbf{x} \in R^n$, então, ou $A\mathbf{x} = \mathbf{0}$ ou existe um $\mathbf{y} \in I(A^T)$ tal que $\mathbf{x}^T\mathbf{y} \neq 0$. Desenhe uma imagem semelhante à da Fig. 2 para ilustrar esse resultado geometricamente no caso em que $N(A)$ é um subespaço bidimensional de R^3.

12. Seja A uma matriz $m \times n$. Explique por que as afirmações a seguir são verdadeiras.
(a) Qualquer vetor $\mathbf{x}$ em R^n pode ser escrito de maneira única como uma soma $\mathbf{y} + \mathbf{z}$, onde $\mathbf{y} \in N(A)$ e $\mathbf{z} \in I(A^T)$.
(b) Qualquer vetor $\mathbf{b} \in R^m$ pode ser escrito de maneira única como uma soma $\mathbf{u} + \mathbf{v}$, onde $\mathbf{u} \in N(A^T)$ e $\mathbf{v} \in I(A)$.

13. Seja A uma matriz $m \times n$. Mostre que:
(a) Se $\mathbf{x} \in N(A^TA)$, então $A\mathbf{x}$ pertence a ambos $I(A)$ e $N(A^T)$;
(b) $N(A^TA) = N(A)$;

(c) A e A^TA têm o mesmo posto;
(d) Se A tem colunas linearmente independentes, então A^TA é invertível.

14. Sejam $A \in R^{m \times n}$, $B \in R^{n \times r}$ e $C = AB$. Mostre que:
(a) $N(B)$ é um subespaço de $N(C)$;
(b) $N(C)^{\perp}$ é um subespaço de $N(B)^{\perp}$ e, conseqüentemente, $I(C^T)$ é um subespaço de $I(B^T)$.

15. Sejam U e V subespaços de um espaço vetorial W. Se $W = U \oplus V$, mostre que $U \cap V = \{\mathbf{0}\}$.

16. Seja A uma matriz $m \times n$ de posto r e seja $\{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$ uma base para $I(A^T)$. Mostre que $\{A\mathbf{x}_1, \ldots, A\mathbf{x}_r\}$ é uma base para $I(A)$.

3 ESPAÇOS MUNIDOS DE PRODUTO INTERNO

Produtos escalares são úteis não apenas em R^n, mas em uma ampla variedade de contextos. Para generalizar esse conceito para outros espaços vetoriais, vamos começar com a definição a seguir.

DEFINIÇÃO E EXEMPLOS

Definição. Um **produto interno** em um espaço vetorial V é uma operação que associa a cada par de vetores $\mathbf{x}$ e $\mathbf{y}$ em V um número real $<\mathbf{x}, \mathbf{y}>$ satisfazendo as seguintes condições:

(i) $<\mathbf{x}, \mathbf{x}> \geq 0$ e a igualdade é válida se e somente se $\mathbf{x} = \mathbf{0}$.
(ii) $<\mathbf{x}, \mathbf{y}> = <\mathbf{y}, \mathbf{x}>$ quaisquer que sejam $\mathbf{x}$ e $\mathbf{y}$ em V.
(iii) $<\alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z}> = \alpha <\mathbf{x}, \mathbf{z}> + \beta <\mathbf{y}, \mathbf{z}>$ quaisquer que sejam $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ em V e α, β escalares.

O Espaço Vetorial R^n. O produto interno canônico em R^n é o produto escalar

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$$

Dado um vetor $\mathbf{w}$ com coeficientes positivos, poderíamos, também, definir um produto interno em R^n por

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i w_i \tag{1}$$

Os coeficientes w_i são chamados de *pesos*.

O Espaço Vetorial $R^{m \times n}$. Dados A e B em $R^{m \times n}$, podemos definir um produto interno por

$$\langle A, B \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ij} \tag{2}$$

Deixamos a cargo do leitor verificar que (2) define, de fato, um produto interno em $R^{m \times n}$.

O Espaço Vetorial $C[a, b]$. Em $C[a, b]$ podemos definir um produto interno por

$$\langle f, g \rangle = \int_a^b f(x) g(x)\, dx \tag{3}$$

Observe que

$$\langle f, f \rangle = \int_a^b (f(x))^2\, dx \geq 0$$

Se $f(x_0) \neq 0$ para algum x_0 em $[a, b]$, então, como $(f(x))^2$ é contínua, existe um subintervalo I de $[a, b]$ contendo x_0 tal que $(f(x))^2 \geq (f(x_0))^2/2$ para todo x em I. Se denotarmos por p o comprimento de I, temos que

$$\langle f, f \rangle = \int_a^b (f(x))^2\, dx \geq \int_I (f(x))^2\, dx \geq \frac{(f(x_0))^2 p}{2} > 0$$

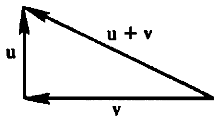


FIG. 5.3.1

Então, se $<f, f> = 0$, $f(x)$ tem que ser identicamente nula em $[a, b]$. Deixamos a cargo do leitor verificar que (3) satisfaz as outras duas condições especificadas na definição de produto interno.

Se $w(x)$ é uma função contínua e positiva em $[a, b]$, então

$$\langle f, g \rangle = \int_a^b f(x)g(x)w(x)\,dx \tag{4}$$

também define um produto interno em $C[a, b]$. A função $w(x)$ é chamada de *peso*. É possível, portanto, definir muitos produtos internos diferentes em $C[a, b]$.

O Espaço Vetorial P_n. Sejam $x_1, x_2, \ldots, x_n$ números reais distintos. Para cada par p, q de polinômios em P_n, defina

$$\langle p, q \rangle = \sum_{i=1}^{n} p(x_i)q(x_i) \tag{5}$$

É fácil verificar que (5) satisfaz as condições (ii) e (iii) da definição de produto interno. Para mostrar que (i) é válida, observe que

$$\langle p, p \rangle = \sum_{i=1}^{n} (p(x_i))^2 \geq 0$$

Se $<p, p> = 0$, então $x_1, x_2, \ldots, x_n$ têm que ser raízes de $p(x) = 0$. Como $p(x)$ tem grau menor do que n, ele tem que ser o polinômio nulo.

Se $w(x)$ é uma função positiva, então

$$\langle p, q \rangle = \sum_{i=1}^{n} p(x_i)q(x_i)w(x_i)$$

também define um produto interno em P_n.

PROPRIEDADES BÁSICAS DE ESPAÇOS COM PRODUTO INTERNO

Os resultados apresentados na Seção 1 para produtos escalares em R^n podem ser generalizados para espaços com produto interno. Em particular, se $\mathbf{v}$ é um vetor em um espaço vetorial V munido de um produto interno, o *comprimento* ou *norma* de $\mathbf{v}$ é definido por

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

Dois vetores $\mathbf{u}$ e $\mathbf{v}$ são ditos *ortogonais* se $<\mathbf{u}, \mathbf{v}> = 0$. O próximo teorema nos dá uma propriedade interessante de vetores ortogonais.

Teorema 5.3.1 (Teorema de Pitágoras). *Se $\mathbf{u}$ e $\mathbf{v}$ são vetores ortogonais em um espaço vetorial V munido de um produto interno, então*

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

Demonstração

$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|^2 &= \langle \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}\rangle \\ &= \langle \mathbf{u}, \mathbf{u}\rangle + 2\langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{v}\rangle \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2\end{aligned}$$ □

Em R^2, esse é simplesmente o teorema de Pitágoras usual, como ilustrado na Fig. 5.3.1.

EXEMPLO 1. Considere o espaço vetorial $C[-1, 1]$ com o produto interno definido por (3). Os vetores 1 e x são ortogonais, já que

$$\langle 1, x\rangle = \int_{-1}^{1} 1 \cdot x\, dx = 0$$

Para determinar o comprimento desses vetores, calculamos

$$\begin{aligned}\langle 1, 1\rangle &= \textstyle\int_{-1}^{1} 1 \cdot 1\, dx &= 2 \\ \langle x, x\rangle &= \textstyle\int_{-1}^{1} x^2\, dx &= \frac{2}{3}\end{aligned}$$

Temos que

$$\begin{aligned}\|1\| &= (\langle 1, 1\rangle)^{1/2} &= \sqrt{2} \\ \|x\| &= (\langle x, x\rangle)^{1/2} &= \frac{\sqrt{6}}{3}\end{aligned}$$

Como 1 e x são ortogonais, eles satisfazem o teorema de Pitágoras,

$$\|1 + x\|^2 = \|1\|^2 + \|x\|^2 = 2 + \frac{2}{3} = \frac{8}{3}$$

O leitor pode verificar que

$$\|1 + x\|^2 = \langle 1 + x, 1 + x\rangle = \int_{-1}^{1} (1 + x)^2\, dx = \frac{8}{3}$$ □

EXEMPLO 2. Para o espaço vetorial $C[-\pi, \pi]$, se usarmos como peso a função constante $w(x) = 1/\pi$ para definir o produto interno

(6) $$\langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)\, dx$$

obteremos

$$\begin{aligned}\langle \cos x, \operatorname{sen} x\rangle &= \frac{1}{\pi}\textstyle\int_{-\pi}^{\pi} \cos x \operatorname{sen} x\, dx &= 0 \\ \langle \cos x, \cos x\rangle &= \frac{1}{\pi}\textstyle\int_{-\pi}^{\pi} \cos x \cos x\, dx &= 1 \\ \langle \operatorname{sen} x, \operatorname{sen} x\rangle &= \frac{1}{\pi}\textstyle\int_{-\pi}^{\pi} \operatorname{sen} x \operatorname{sen} x\, dx &= 1\end{aligned}$$

Logo, $\cos x$ e $\operatorname{sen} x$ são vetores unitários ortogonais em relação a esse produto interno. Do teorema de Pitágoras, obtemos

$$\|\cos x + \operatorname{sen} x\| = \sqrt{2}$$ □

Para o espaço vetorial $R^{m\times n}$, a norma definida pelo produto interno é chamada de *norma de Frobenius* e é denotada por $\|\cdot\|_F$. Logo, se $A \in R^{m\times n}$, temos

$$\|A\|_F = (\langle A, A\rangle)^{1/2} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right)^{1/2}$$

EXEMPLO 3. Se

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 3 & 3 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} -1 & 1 \\ 3 & 0 \\ -3 & 4 \end{pmatrix}$$

então

$$\langle A, B\rangle = 1\cdot -1 + 1\cdot 1 + 1\cdot 3 + 2\cdot 0 + 3\cdot -3 + 3\cdot 4 = 6$$

Logo A não é ortogonal a B. As normas dessas duas matrizes são dadas por

$$\|A\|_F = (1+1+1+4+9+9)^{1/2} = 5$$

$$\|B\|_F = (1+1+9+0+9+16)^{1/2} = 6$$

□

EXEMPLO 4. Em P_5, defina um produto interno por (5) com $x_i = (i-1)/4$ para $i = 1, 2, \ldots, 5$. O comprimento da função $p(x) = 4x$ é dado por

$$\|4x\| = (\langle 4x, 4x\rangle)^{1/2} = \left(\sum_{i=1}^{5} 16x_i^2\right)^{1/2} = \left(\sum_{i=1}^{5} (i-1)^2\right)^{1/2} = \sqrt{30}$$

□

Definição. Se **u** e **v** são vetores em um espaço vetorial V com produto interno e se $\mathbf{v} \neq \mathbf{0}$, então a **projeção escalar** de **u** sobre **v** é definida por

$$\alpha = \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{v}\|}$$

e a **projeção vetorial** de **u** sobre **v** é dada por

$$\mathbf{p} = \alpha\left(\frac{1}{\|\mathbf{v}\|}\mathbf{v}\right) = \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\langle \mathbf{v}, \mathbf{v}\rangle}\mathbf{v} \tag{7}$$

Observações. Se $\mathbf{v} \neq \mathbf{0}$ e se **p** é a projeção vetorial de **u** sobre **v**, então

I. $\mathbf{u} - \mathbf{p}$ e **p** são ortogonais;
II. $\mathbf{u} = \mathbf{p}$ se e somente se **u** é um múltiplo escalar de **v**.

Demonstração de I. Como

$$\langle \mathbf{p}, \mathbf{p}\rangle = \left\langle \frac{\alpha}{\|\mathbf{v}\|}\mathbf{v}, \frac{\alpha}{\|\mathbf{v}\|}\mathbf{v}\right\rangle = \left(\frac{\alpha}{\|\mathbf{v}\|}\right)^2 \langle \mathbf{v}, \mathbf{v}\rangle = \alpha^2$$

e

$$\langle \mathbf{u}, \mathbf{p}\rangle = \frac{(\langle \mathbf{u}, \mathbf{v}\rangle)^2}{\langle \mathbf{v}, \mathbf{v}\rangle} = \alpha^2$$

temos que

$$\langle \mathbf{u} - \mathbf{p}, \mathbf{p}\rangle = \langle \mathbf{u}, \mathbf{p}\rangle - \langle \mathbf{p}, \mathbf{p}\rangle = \alpha^2 - \alpha^2 = 0$$

Portanto, $\mathbf{u} - \mathbf{p}$ e **p** são ortogonais. □

Demonstração de II. Se $\mathbf{u} = \beta\mathbf{v}$, então a projeção vetorial de $\mathbf{u}$ sobre $\mathbf{v}$ é dada por

$$\mathbf{p} = \frac{\langle \beta\mathbf{v}, \mathbf{v}\rangle}{\langle \mathbf{v}, \mathbf{v}\rangle}\mathbf{v} = \beta\mathbf{v} = \mathbf{u}$$

Reciprocamente, se $\mathbf{u} = \mathbf{p}$, segue da Equação (7) que

$$\mathbf{u} = \beta\mathbf{v} \qquad \text{onde} \quad \beta = \frac{\alpha}{\|\mathbf{v}\|}$$

□

As observações I e II são úteis para se obter o teorema a seguir.

Teorema 5.3.2 (Desigualdade de Cauchy-Schwarz). *Se* $\mathbf{u}$ *e* $\mathbf{v}$ *são dois vetores quaisquer em um espaço vetorial V munido de um produto interno, então*

(8) $$|\langle \mathbf{u}, \mathbf{v}\rangle| \le \|\mathbf{u}\|\,\|\mathbf{v}\|$$

A igualdade é válida se e somente se $\mathbf{u}$ *e* $\mathbf{v}$ *são linearmente dependentes.*

Demonstração. Se $\mathbf{v} = \mathbf{0}$, então

$$|\langle \mathbf{u}, \mathbf{v}\rangle| = 0 = \|\mathbf{u}\|\,\|\mathbf{v}\|$$

Se $\mathbf{v} \neq \mathbf{0}$, seja $\mathbf{p}$ a projeção vetorial de $\mathbf{u}$ sobre $\mathbf{v}$. Como $\mathbf{p}$ é ortogonal a $\mathbf{u} - \mathbf{p}$, pelo teorema de Pitágoras temos que

$$\|\mathbf{p}\|^2 + \|\mathbf{u} - \mathbf{p}\|^2 = \|\mathbf{u}\|^2$$

Logo,

$$\frac{(\langle \mathbf{u}, \mathbf{v}\rangle)^2}{\|\mathbf{v}\|^2} = \|\mathbf{p}\|^2 = \|\mathbf{u}\|^2 - \|\mathbf{u} - \mathbf{p}\|^2$$

e, portanto,

(9) $$(\langle \mathbf{u}, \mathbf{v}\rangle)^2 = \|\mathbf{u}\|^2\|\mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{p}\|^2\|\mathbf{v}\|^2 \le \|\mathbf{u}\|^2\|\mathbf{v}\|^2$$

Então,

$$|\langle \mathbf{u}, \mathbf{v}\rangle| \le \|\mathbf{u}\|\,\|\mathbf{v}\|$$

A igualdade na Equação (9) é válida se e somente se $\mathbf{u} = \mathbf{p}$. Pela observação II, a igualdade em (8) é válida se e somente se $\mathbf{v} = \mathbf{0}$ ou $\mathbf{u}$ é um múltiplo de $\mathbf{v}$. Colocado de maneira mais simples, a igualdade é válida se e somente se $\mathbf{u}$ e $\mathbf{v}$ são linearmente dependentes. □

Uma conseqüência da desigualdade de Cauchy-Schwarz é que, se $\mathbf{u}$ e $\mathbf{v}$ são vetores não-nulos, então

$$-1 \le \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{u}\|\,\|\mathbf{v}\|} \le 1$$

e, portanto, existe um único ângulo θ em $[0, \pi]$ tal que

(10) $$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{u}\|\,\|\mathbf{v}\|}$$

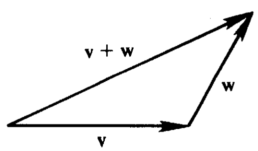


FIG. 5.3.2

A equação (10) pode, então, ser usada para se definir o ângulo θ entre dois vetores não-nulos **u** e **v**.

NORMAS

A palavra *norma*, em matemática, tem significado próprio, independentemente da existência de produto interno, e devemos justificar sua utilização aqui.

Definição. Um espaço vetorial V é dito um **espaço vetorial normado** se a cada vetor $\mathbf{v} \in V$ está associado um número real $\|\mathbf{v}\|$, chamado de **norma** de **v**, satisfazendo as seguintes propriedades:

(i) $\|\mathbf{v}\| \geq 0$ e a igualdade vale se e somente se $\mathbf{v} = \mathbf{0}$;
(ii) $\|\alpha\mathbf{v}\| = |\alpha| \, \|\mathbf{v}\|$ para todo escalar α;
(iii) $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$ quaisquer que sejam $\mathbf{v}, \mathbf{w} \in V$.

A terceira condição é chamada de *desigualdade triangular* (ver Fig. 5.3.2).

Teorema 5.3.3. *Se V é um espaço munido de produto interno, então*

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} \quad \textit{para todo} \quad \mathbf{v} \in V$$

define uma norma em V.

Demonstração. É fácil ver que as condições (i) e (ii) da definição são satisfeitas. Vamos deixar a cargo do leitor essa verificação e mostrar como provar (iii).

$$\begin{aligned}
\|\mathbf{u} + \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle \\
&= \langle \mathbf{u}, \mathbf{u} \rangle + 2\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \\
&\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\| \, \|\mathbf{v}\| + \|\mathbf{v}\|^2 \quad \text{(Cauchy–Schwarz)} \\
&= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2
\end{aligned}$$

Portanto,

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}| \qquad \square$$

É possível definir muitas normas diferentes em um espaço vetorial dado. Por exemplo, em R^n poderíamos definir

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i|$$

para todo $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$. É fácil verificar que $\| \cdot \|_1$ define uma norma em R^n. Uma outra norma importante em R^n é a *norma uniforme*, definida por

$$\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

Mais geralmente, poderíamos definir uma norma em R^n por

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

para qualquer número real $p \geq 1$. Em particular, se $p = 2$, temos

$$\|\mathbf{x}\|_2 = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

A norma $\|\cdot\|_2$ é a norma em R^n proveniente do produto interno. Se $p \neq 2$, $\|\cdot\|_p$ não corresponde a nenhum produto interno. O teorema de Pitágoras não é válido para normas que não correspondem a nenhum produto interno. Por exemplo,

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \text{e} \quad \mathbf{x}_2 = \begin{pmatrix} -4 \\ 2 \end{pmatrix}$$

são ortogonais; no entanto,

$$\|\mathbf{x}_1\|_\infty^2 + \|\mathbf{x}_2\|_\infty^2 = 4 + 16 = 20$$

enquanto

$$\|\mathbf{x}_1 + \mathbf{x}_2\|_\infty^2 = 16$$

Por outro lado, se usarmos a norma $\|\cdot\|_2$, temos

$$\|\mathbf{x}_1\|_2^2 + \|\mathbf{x}_2\|_2^2 = 5 + 20 = 25 = \|\mathbf{x}_1 + \mathbf{x}_2\|_2^2$$

EXEMPLO 5. Seja $\mathbf{x}$ o vetor $(4, -5, 3)^T$ em R^3. Calcule $\|\mathbf{x}\|_1$, $\|\mathbf{x}\|_2$ e $\|\mathbf{x}\|_\infty$.

$$\|\mathbf{x}\|_1 = |4| + |-5| + |3| = 12$$
$$\|\mathbf{x}\|_2 = \sqrt{16 + 25 + 9} = 5\sqrt{2}$$
$$\|\mathbf{x}\|_\infty = \max(|4|, |-5|, |3|) = 5$$

□

É possível, também, definir normas diferentes em $R^{m\times n}$. No Cap. 7, vamos estudar outros tipos de normas matriciais que são úteis para se determinar a sensibilidade de um sistema linear.

Em geral, uma norma nos dá um modo de medir a distância entre vetores.

Definição. Seja $\mathbf{x}$ e $\mathbf{y}$ vetores em um espaço vetorial normado. A distância entre $\mathbf{x}$ e $\mathbf{y}$ é definida pelo número $\|\mathbf{x} - \mathbf{y}\|$.

EXEMPLO 6. Considere $\mathbf{x} = (x_1, x_2)^T$ e $\mathbf{y} = (y_1, y_2)^T$ no espaço R^2 com a norma $\|\cdot\|_2$. A distância entre $\mathbf{x}$ e $\mathbf{y}$ é o comprimento de $\mathbf{x} - \mathbf{y}$ (ver Fig. 5.3.3).

$$\|\mathbf{x} - \mathbf{y}\|_2 = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

Essa é a fórmula padrão utilizada em geometria analítica para a distância entre dois pontos no plano.

□

Muitas aplicações envolvem a procura de um único vetor em um subespaço S que seja o mais próximo possível de um dado vetor $\mathbf{v}$ no espaço vetorial V. Se a norma utilizada em V está associada a um produto interno, então o vetor mais próximo pode ser calculado como a projeção vetorial de $\mathbf{v}$ sobre o subespaço S. Esse tipo de problema de aproximação será discutido em cada uma das três seções seguintes deste capítulo.

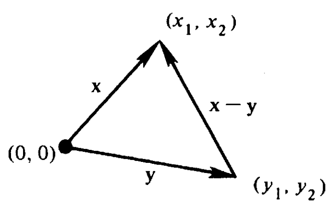


FIG. 5.3.3

EXERCÍCIOS

1. Sejam $\mathbf{x} = (-1, -1, 1, 1)^T$ e $\mathbf{y} = (1, 1, 5, -3)^T$. Mostre que $\mathbf{x} \perp \mathbf{y}$. Calcule $\|\mathbf{x}\|_2$, $\|\mathbf{y}\|_2$, $\|\mathbf{x} + \mathbf{y}\|_2$ e verifique a validade do teorema de Pitágoras.

2. Seja $\mathbf{x} = (1, 1, 1, 1)^T$ e $\mathbf{y} = (8, 2, 2, 0)^T$.
(a) Encontre o ângulo θ entre $\mathbf{x}$ e $\mathbf{y}$.
(b) Encontre a projeção vetorial $\mathbf{p}$ de $\mathbf{x}$ sobre $\mathbf{y}$.
(c) Verifique que $\mathbf{x} - \mathbf{p}$ é ortogonal a $\mathbf{p}$.
(d) Calcule $\|\mathbf{x} - \mathbf{p}\|_2$, $\|\mathbf{p}\|_2$, $\|\mathbf{x}\|_2$ e verifique a validade do teorema de Pitágoras.

3. Seja $\mathbf{w} = (1/4, 1/2, 1/4)^T$ e use a Equação (1) para definir um produto interno em R^3. Sejam $\mathbf{x} = (1, 1, 1)^T$ e $\mathbf{y} = (-5, 1, 3)^T$.
(a) Mostre que $\mathbf{x}$ e $\mathbf{y}$ são ortogonais em relação a esse produto interno com peso.
(b) Calcule os valores de $\|\mathbf{x}\|$ e $\|\mathbf{y}\|$ para esse produto interno.

4. Considere

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 0 & 2 \\ 3 & 1 & 1 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} -4 & 1 & 1 \\ -3 & 3 & 2 \\ 1 & -2 & -2 \end{pmatrix}$$

Determine o valor de cada uma das expressões a seguir.
(a) $\langle A, B \rangle$ (b) $\|A\|_F$ (c) $\|B\|_F$ (d) $\|A + B\|_F$

5. Mostre que a Equação (2) define um produto interno em $R^{m\times n}$.

6. Mostre que o produto interno definido pela Equação (3) satisfaz as duas últimas condições na definição de produto interno.

7. Para $C[0, 1]$ com produto interno definido por (3), calcule:

$$\text{(a) } \langle e^x, e^{-x} \rangle \qquad \text{(b) } \langle x, \operatorname{sen} \pi x \rangle \qquad \text{(c) } \langle x^2, x^3 \rangle$$

8. Para $C[0, 1]$ com produto interno definido por (3), considere os vetores 1 e x.
(a) Encontre o ângulo θ entre 1 e x.
(b) Determine a projeção vetorial $\mathbf{p}$ de 1 sobre x e verifique que $1 - \mathbf{p}$ é ortogonal a $\mathbf{p}$.
(c) Calcule $\|1 - \mathbf{p}\|$, $\|\mathbf{p}\|$, $\|1\|$ e verifique a validade do teorema de Pitágoras.

9. Para $C[-\pi, \pi]$ com produto interno definido por (6), mostre que cos mx e sen nx são ortogonais e que ambos são vetores unitários. Determine a distância entre os dois vetores.

10. Mostre que as funções x e x^2 são ortogonais em P_5 em relação ao produto interno definido por (5), onde $x_i = (i - 3)/2$ para $i = 1, \ldots, 5$.

11. Considere em P_5 o produto interno como no Exercício 10 e a norma definida por

$$\|p\| = \sqrt{\langle p, p \rangle} = \left\{ \sum_{i=1}^{5} [p(x_i)]^2 \right\}^{1/2}$$

Calcule:
(a) $\|x\|$ (b) $\|x^2\|$ (c) a distância entre x e x^2

12. Se V é um espaço munido de um produto interno, mostre que

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

satisfaz as duas primeiras propriedades na definição de norma.

13. Mostre que

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i|$$

define uma norma em R^n.

14. Mostre que

$$\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

define uma norma em R^n.

15. Calcule $\|\mathbf{x}\|_1$, $\|\mathbf{x}\|_2$, $\|\mathbf{x}\|_\infty$ para cada um dos vetores a seguir pertencentes a R^3.

(a) $\mathbf{x} = (-3, 4, 0)^T$ (b) $\mathbf{x} = (-1, -1, 2)^T$ (c) $\mathbf{x} = (1, 1, 1)^T$

16. Sejam $\mathbf{x} = (5, 2, 4)^T$ e $\mathbf{y} = (3, 3, 2)^T$. Calcule $\|\mathbf{x} - \mathbf{y}\|_1$, $\|\mathbf{x} - \mathbf{y}\|_2$ e $\|\mathbf{x} - \mathbf{y}\|_\infty$. Para qual dessas normas os dois vetores estão mais próximos? Para qual eles estão mais longe?

17. Sejam $\mathbf{x}$ e $\mathbf{y}$ dois vetores em um espaço com produto interno. Mostre que, se $\mathbf{x} \perp \mathbf{y}$, então a distância entre $\mathbf{x}$ e $\mathbf{y}$ é

$$\left(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2\right)^{1/2}$$

18. Considere R^n com o produto interno

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$$

Encontre uma fórmula para a distância entre dois vetores $\mathbf{x} = (x_1, \ldots, x_n)^T$ e $\mathbf{y}$ $(y_1, \ldots, y_n)^T$.

19. Seja $\mathbf{x} \in R^n$. Mostre que $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2$.

20. Seja $\mathbf{x} \in R^2$. Mostre que $\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1$.

[***Sugestão:*** Escreva $\mathbf{x}$ na forma $x_1\mathbf{e}_1 + x_2\mathbf{e}_2$ e use a desigualdade triangular.]

21. Dê um exemplo de um vetor não-nulo $\mathbf{x} \in R^2$ para o qual

$$\|\mathbf{x}\|_\infty = \|\mathbf{x}\|_2 = \|\mathbf{x}\|_1$$

22. Mostre que, em qualquer espaço vetorial normado,

$$\|-\mathbf{v}\| = \|\mathbf{v}\|$$

23. Mostre que, quaisquer que sejam $\mathbf{u}$ e $\mathbf{v}$ em um espaço vetorial normado,

$$\|\mathbf{u} + \mathbf{v}\| \geq | \, \|\mathbf{u}\| - \|\mathbf{v}\| \, |$$

24. Mostre que, quaisquer que sejam $\mathbf{u}$ e $\mathbf{v}$ em um espaço vetorial munido de um produto interno,

$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$$

Interprete geometricamente esse resultado para o espaço vetorial R^2.

25. O resultado do Exercício 24 não é válido para normas que não estão associadas a produtos internos. Dê um exemplo disso em R^2 para a norma $\|\cdot\|_1$.

26. Determine se as expressões a seguir definem ou não normas em $C[a, b]$.

(a) $\|f\| = |f(a)| + |f(b)|$
(b) $\|f\| = \int_a^b |f(x)|\, dx$
(c) $\|f\| = \max_{a \leq x \leq b} |f(x)|$

27. Seja $\mathbf{x} \in R^n$. Mostre que:

(a) $\|\mathbf{x}\|_1 \leq n\|\mathbf{x}\|_\infty$ (b) $\|\mathbf{x}\|_2 \leq \sqrt{n}\, \|\mathbf{x}\|_\infty$

Dê exemplos de vetores em R^n para os quais as igualdades nos itens (a) e (b) são válidas.

28. Desenhe o conjunto de pontos $(x_1, x_2) = x^T$ em R^2 para os quais:

(a) $\|\mathbf{x}\|_2 = 1$ (b) $\|\mathbf{x}\|_1 = 1$ (c) $\|\mathbf{x}\|_\infty = 1$

29. Considere o espaço vetorial R^n com o produto interno $<\mathbf{x}, \mathbf{y}> = \mathbf{x}^T\mathbf{y}$. Mostre que, qualquer que seja a matriz A $m \times n$, temos:

(a) $\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A^T\mathbf{y} \rangle$ (b) $\langle A^T A\mathbf{x}, \mathbf{x} \rangle = \|A\mathbf{x}\|^2$

4 PROBLEMAS DE MÍNIMOS QUADRÁTICOS

Até agora, preocupamo-nos essencialmente com sistemas de equações lineares compatíveis. Nesta seção, vamos considerar sistemas envolvendo mais equações do que incógnitas. Tais sistemas são, muitas vezes, incompatíveis. Então, dado um sistema $m \times n$ $A\mathbf{x} = \mathbf{b}$ com $m > n$, não podemos esperar, em geral, encontrar um vetor $\mathbf{x} \in R^n$ para o qual $A\mathbf{x}$ seja igual a $\mathbf{b}$. Em vez disso, vamos procurar um vetor $\mathbf{x}$ para o qual $A\mathbf{x}$ está "o mais próximo possível" de $\mathbf{b}$. Como poderíamos esperar, a ortogonalidade tem um papel importante na escolha de tal $\mathbf{x}$.

Seja A uma matriz $m \times n$ com $m > n$. Para cada $\mathbf{b} \in R^m$, defina

$$\|\mathbf{b}\| = \sqrt{\langle \mathbf{b}, \mathbf{b} \rangle} = \sqrt{\mathbf{b}^T\mathbf{b}}$$

Considere o sistema de equações $A\mathbf{x} = \mathbf{b}$. Para cada $\mathbf{x} \in R^n$, podemos formar o *resíduo*

$$r(\mathbf{x}) = \mathbf{b} - A\mathbf{x}$$

A distância entre $\mathbf{b}$ e $A\mathbf{x}$ é dada por

$$\|\mathbf{b} - A\mathbf{x}\| = \|r(\mathbf{x})\|$$

Queremos encontrar um vetor $\mathbf{x} \in R^n$ para o qual $\|r(\mathbf{x})\|$ é mínimo. Minimizar $\|r(\mathbf{x})\|$ é equivalente a minimizar $\|r(\mathbf{x})\|^2$. Um vetor $\hat{\mathbf{x}}$ que faz isso é dito uma *solução de mínimos quadráticos* para o sistema $A\mathbf{x} = \mathbf{b}$.

Se $\hat{\mathbf{x}}$ é uma solução de mínimos quadráticos para o sistema $A\mathbf{x} = \mathbf{b}$ e $\mathbf{p} = A\hat{\mathbf{x}}$, então $\mathbf{p}$ é o vetor no espaço coluna de A mais próximo de $\mathbf{b}$. O próximo teorema garante não só a existência do vetor mais próximo $\mathbf{p}$, mas também sua unicidade. Além disso, o teorema nos dá uma caracterização importante do vetor mais próximo.

Teorema 5.4.1. *Seja S um subespaço de R^m. Para cada $\mathbf{b} \in R^m$ existe um único elemento $\mathbf{p}$ de S que está o mais próximo possível de $\mathbf{b}$, isto é,*

$$\|\mathbf{b} - \mathbf{y}\| > \|\mathbf{b} - \mathbf{p}\|$$

para todo $\mathbf{y} \neq \mathbf{p}$ em S. Além disso, um determinado vetor $\mathbf{p}$ em S está o mais próximo possível de um dado vetor $\mathbf{b} \in R^m$ se e somente se $\mathbf{b} - \mathbf{p} \in S^\perp$.

Demonstração. Como $R^m = S \oplus S^\perp$, cada elemento $\mathbf{b}$ em R^m pode ser expresso de maneira única como uma soma

$$\mathbf{b} = \mathbf{p} + \mathbf{z}$$

onde $\mathbf{p} \in S$ e $\mathbf{z} \in S^\perp$. Se $\mathbf{y}$ é outro elemento qualquer de S, então

$$\|\mathbf{b} - \mathbf{y}\|^2 = \|(\mathbf{b} - \mathbf{p}) + (\mathbf{p} - \mathbf{y})\|^2$$

Como $\mathbf{p} - \mathbf{y} \in S$ e $\mathbf{b} - \mathbf{p} = \mathbf{z} \in S^\perp$, pelo teorema de Pitágoras temos que

$$\|\mathbf{b} - \mathbf{y}\|^2 = \|\mathbf{b} - \mathbf{p}\|^2 + \|\mathbf{p} - \mathbf{y}\|^2$$

Portanto,

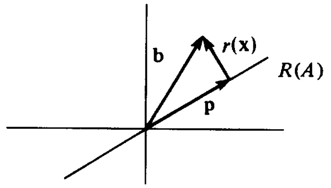
$$\|\mathbf{b} - \mathbf{y}\| > \|\mathbf{b} - \mathbf{p}\|$$

Logo, se $\mathbf{p} \in S$ e $\mathbf{b} - \mathbf{p} \in S^\perp$, então $\mathbf{p}$ é o elemento de S mais próximo de $\mathbf{b}$. Reciprocamente, se $\mathbf{q} \in S$ e $\mathbf{b} - \mathbf{q} \notin S^\perp$, então $\mathbf{q} \neq \mathbf{p}$ e, pelo argumento anterior (com $\mathbf{y} = \mathbf{q}$), temos que

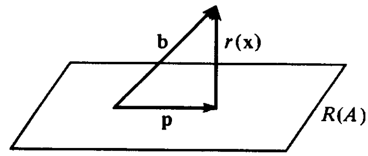
$$\|\mathbf{b} - \mathbf{q}\| > \|\mathbf{b} - \mathbf{p}\| \qquad \square$$

No caso especial em que $\mathbf{b}$ pertence ao subespaço S, temos

$$\mathbf{b} = \mathbf{p} + \mathbf{z} \qquad \mathbf{p} \in S, \quad \mathbf{z} \in S^\perp$$



(a) b $\in R^2$ e A é uma matriz 2×1 de posto 1.



(b) b $\in R^3$ e A é uma matriz 3×1 de posto 2.

FIG. 5.4.1

e

$$\mathbf{b} = \mathbf{b} + \mathbf{0}$$

Pela unicidade da representação em soma direta, temos

$$\mathbf{p} = \mathbf{b} \qquad \text{e} \qquad \mathbf{z} = \mathbf{0}$$

Um vetor $\hat{\mathbf{x}}$ é uma solução para o problema de mínimos quadráticos $A\mathbf{x} = \mathbf{b}$ se e somente se $\mathbf{p} = A\hat{\mathbf{x}}$ é o vetor em $I(A)$ mais próximo de $\mathbf{b}$. O vetor $\mathbf{p}$ é a *projeção ortogonal de* $\mathbf{b}$ *sobre* $I(A)$. Pelo Teorema 5.4.1, temos que

$$\mathbf{b} - \mathbf{p} = \mathbf{b} - A\hat{\mathbf{x}} = r(\hat{\mathbf{x}})$$

tem que ser um elemento de $I(A)^{\perp}$. Logo, $\hat{\mathbf{x}}$ é a solução para o problema de mínimos quadráticos se e somente se

(1)
$$r(\hat{\mathbf{x}}) \in I(A)^{\perp}$$

(ver a Fig. 5.4.1). Como encontrar um vetor $\hat{\mathbf{x}}$ satisfazendo (1)? A chave da solução do problema de mínimos quadráticos é dada pelo Teorema 5.2.1, que diz que

$$I(A)^{\perp} = N(A^T)$$

Um vetor $\hat{\mathbf{x}}$ é uma solução para o problema de mínimos quadráticos $A\mathbf{x} = \mathbf{b}$ se e somente se

$$r(\hat{\mathbf{x}}) \in N(A^T)$$

ou, equivalentemente,

$$\mathbf{0} = A^T r(\hat{\mathbf{x}}) = A^T(\mathbf{b} - A\hat{\mathbf{x}})$$

Logo, para resolver o problema de mínimos quadráticos $A\mathbf{x} = \mathbf{b}$, precisamos resolver

(2)
$$A^T A\mathbf{x} = A^T\mathbf{b}$$

A Equação (2) representa um sistema $n \times n$ de equações lineares. Essas equações são chamadas de *equações normais*. Em geral, é possível que as equações normais tenham mais de uma solução; no entanto, se $\hat{\mathbf{x}}$ e $\hat{\mathbf{y}}$ são ambos soluções, como a projeção ortogonal $\mathbf{p}$ de $\mathbf{b}$ sobre $I(A)$ é única,

$$A\hat{\mathbf{x}} = A\hat{\mathbf{y}} = \mathbf{p}$$

O próximo teorema determina condições sob as quais o problema de mínimos quadráticos $A\mathbf{x} = \mathbf{b}$ tem uma única solução.

Teorema 5.4.2. *Se A é uma matriz $m \times n$ de posto n, então as equações normais*

$$A^T A\mathbf{x} = A^T\mathbf{b}$$

têm uma única solução

$$\hat{\mathbf{x}} = (A^T A)^{-1} A^T\mathbf{b}$$

e $\hat{\mathbf{x}}$ é a única solução de mínimos quadráticos para o sistema $A\mathbf{x} = \mathbf{b}$.

Demonstração. Vamos mostrar primeiro que A^TA é invertível. Para provar isso, seja $\mathbf{z}$ uma solução de

(3) $$A^TA\mathbf{x} = \mathbf{0}$$

Então, $A\mathbf{z} \in N(A^T)$. É claro que $A\mathbf{z} \in I(A) = N(A^T)^\perp$. Como $N(A^T) \cap N(A^T)^\perp = \{\mathbf{0}\}$, temos que $A\mathbf{z} = \mathbf{0}$. Se A tem posto n, as colunas de A são linearmente independentes e, conseqüentemente, $A\mathbf{x} = \mathbf{0}$ tem apenas a solução trivial. Logo, $\mathbf{z} = \mathbf{0}$ e (3) tem apenas a solução trivial. Portanto, pelo Teorema 1.4.3, A^TA é invertível. Temos, então, que $\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$ é a única solução para as equações normais e, portanto, a única solução de mínimos quadráticos para o sistema $A\mathbf{x} = \mathbf{b}$. □

O vetor projeção

$$\mathbf{p} = A\hat{\mathbf{x}} = A(A^TA)^{-1}A^T\mathbf{b}$$

é o elemento de $I(A)$ mais próximo de $\mathbf{b}$ no sentido de mínimos quadráticos. A matriz $P = A(A^TA)^{-1}A^T$ é chamada de *matriz de projeção*.

APLICAÇÃO

A lei de Hooke diz que a força aplicada a uma mola é proporcional ao comprimento de seu alongamento. Então, se F é a força aplicada e x o alongamento da mola, então $F = kx$. A constante de proporcionalidade k é a *constante da mola*.

Alguns alunos de física querem determinar a constante da mola para uma determinada mola. Eles aplicam forças de 13, 22 e 36 newtons, que alongam a mola em 10, 18 e 28 centímetros, respectivamente. Usando a lei de Hooke, eles obtêm o seguinte sistema de equações:

$$0{,}10k = 13$$

$$0{,}18k = 22$$

$$0{,}28k = 36$$

Esse sistema é claramente incompatível, já que cada uma das equações corresponde a um valor diferente de k. Em vez de usar um desses valores, os alunos decidem calcular a solução de mínimos quadráticos para o sistema.

$$(0{,}10,\ 0{,}18,\ 0{,}28)\begin{pmatrix} 0{,}10 \\ 0{,}18 \\ 0{,}28 \end{pmatrix}(k) = (0{,}10,\ 0{,}18,\ 0{,}28)\begin{pmatrix} 13 \\ 22 \\ 36 \end{pmatrix}$$

$$0{,}1208k = 15{,}34$$

$$k \approx 127$$

EXEMPLO 1. Encontre a solução de mínimos quadráticos para o sistema

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

SOLUÇÃO. As equações normais para esse sistema são

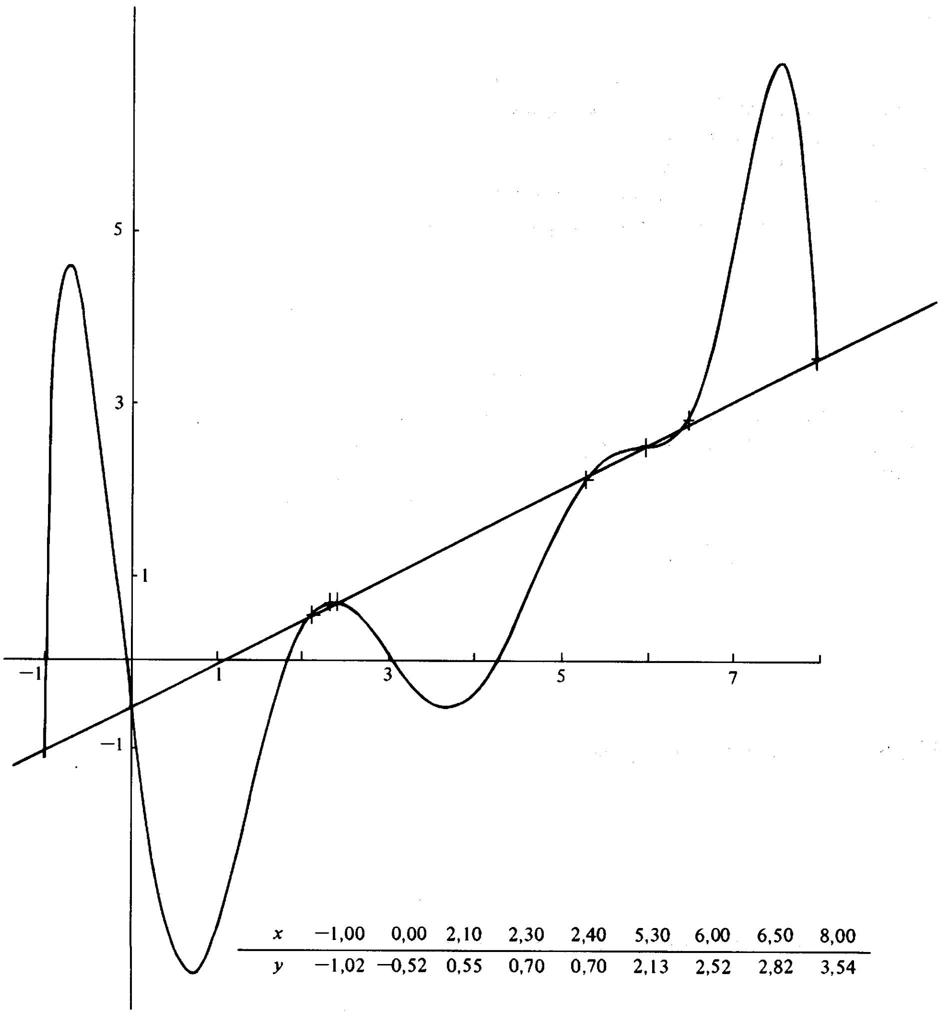
$$\begin{pmatrix} 1 & -2 & 2 \\ 1 & 3 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -2 & 3 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 2 \\ 1 & 3 & -1 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$$

Obtemos, então, o sistema 2 × 2

$$\begin{pmatrix} 9 & -7 \\ -7 & 11 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$$

A solução desse sistema 2 × 2 é $(83/50, 71/50)^T$. □

Cientistas, com freqüência, coletam dados e tentam encontrar uma relação entre as variáveis. Por exemplo, os dados podem envolver as temperaturas $T_0, T_1, \ldots, T_n$ de um líquido medidas nos instantes $t_0, t_1, \ldots, t_n$, respectivamente. Se a temperatura T pode ser representada por uma função do tempo t, essa função pode ser usada para se prever os valores da temperatura em instantes futuros. Se os dados consistem em $n + 1$ pontos no plano, é possível encontrar um polinômio de grau menor ou igual a n cujo gráfico contém todos os pontos. Tal polinômio é chamado de *polinômio interpolador*. De fato, como os dados geralmente envolvem erros experimentais, não há necessidade de exigir que o gráfico contenha todos os pontos. De fato, polinômios de grau menor cujos gráficos não contêm todos os pontos exatamente fornecem, em geral, uma versão mais próxima da relação exata entre as variáveis. Se, por exem-



x	−1,00	0,00	2,10	2,30	2,40	5,30	6,00	6,50	8,00
y	−1,02	−0,52	0,55	0,70	0,70	2,13	2,52	2,82	3,54

FIG. 5.4.2

plo, a relação entre as variáveis é de fato linear e os dados envolvem erros pequenos, seria desastroso usar um polinômio interpolador (ver Fig. 5.4.2).

Dada uma tabela de dados

x	x_1	x_2	$\cdots$	x_m
y	y_1	y_2	$\cdots$	y_m

queremos encontrar uma função linear

$$y = c_0 + c_1 x$$

que melhor se ajusta aos dados, no sentido de mínimos quadráticos. Colocando a condição

$$y_i = c_0 + c_1 x_i \qquad \text{para} \quad i = 1, \dots, m$$

obtemos um sistema de m equações em duas incógnitas.

(4)
$$\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

A função linear cujos coeficientes correspondem à solução de mínimos quadráticos para (4) é o melhor ajuste, por uma função linear, de mínimos quadráticos para os dados.

EXEMPLO 2. Dada a tabela de dados

x	0	3	6
y	1	4	5

encontre o melhor ajuste de mínimos quadráticos por uma função linear.

SOLUÇÃO. Para esse exemplo, o sistema (4) fica

$$A\mathbf{c} = \mathbf{y}$$

onde

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 3 \\ 1 & 6 \end{pmatrix}, \qquad \mathbf{c} = \begin{pmatrix} c_0 \\ c_1 \end{pmatrix}, \qquad \text{e} \qquad \mathbf{y} = \begin{pmatrix} 1 \\ 4 \\ 5 \end{pmatrix}$$

As equações normais

$$A^T A\mathbf{c} = A^T\mathbf{y}$$

ficam

(5)
$$\begin{pmatrix} 3 & 9 \\ 9 & 45 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} 10 \\ 42 \end{pmatrix}$$

A solução desse sistema é (4/3, 2/3). Logo, o melhor ajuste de mínimos quadráticos é dado por

$$y = \tfrac{4}{3} + \tfrac{2}{3}x$$

□

O Exemplo 2 também poderia ser resolvido usando-se Cálculo. O resíduo $r(\mathbf{c})$ é dado por

$$r(\mathbf{c}) = \mathbf{y} - A\mathbf{c}$$

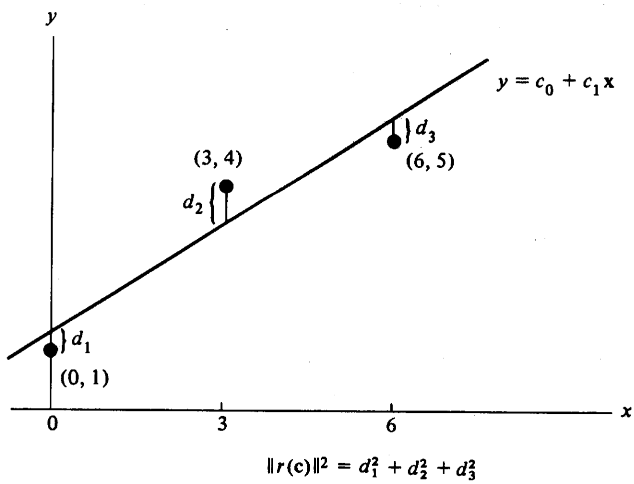


FIG. 5.4.3

e

$$\begin{aligned}\|r(\mathbf{c})\|^2 &= \|\mathbf{y} - A\mathbf{c}\|^2\\ &= [1 - (c_0 + 0c_1)]^2 + [4 - (c_0 + 3c_1)]^2 + [5 - (c_0 + 6c_1)]^2\\ &= f(c_0, c_1)\end{aligned}$$

Logo, $\|r(\mathbf{c})\|^2$ pode ser considerada como uma função de duas variáveis, $f(c_0, c_1)$. O mínimo dessa função será atingido quando suas derivadas parciais se anularem.

$$\frac{\partial f}{\partial c_0} = -2(10 - 3c_0 - 9c_1) = 0$$

$$\frac{\partial f}{\partial c_1} = -6(14 - 3c_0 - 15c_1) = 0$$

Dividindo ambas as equações por 2, obtemos o mesmo sistema que em (5) (ver Fig. 5.4.3).

Se os dados não parecem representar uma função linear, pode-se usar um polinômio de grau maior. Para encontrar os coeficientes $c_0, c_1, \ldots, c_n$ do melhor ajuste de mínimos quadráticos para os dados

x	x_1	x_2	$\cdots$	x_m
y	y_1	y_2	$\cdots$	y_m

por um polinômio de grau n, precisamos encontrar a solução de mínimos quadráticos para o sistema

$$(6)\qquad \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & & & & \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

EXEMPLO 3. Encontre o melhor ajuste de mínimos quadráticos por um polinômio de grau dois para os dados

x	0	1	2	3
y	3	2	4	4

SOLUÇÃO. Para esse exemplo, o sistema (6) fica

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 4 \\ 4 \end{pmatrix}$$

As equações normais são, então,

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 4 & 9 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 4 & 9 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 4 \\ 4 \end{pmatrix}$$

Ou seja,

$$\begin{pmatrix} 4 & 6 & 14 \\ 6 & 14 & 36 \\ 14 & 36 & 98 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 13 \\ 22 \\ 54 \end{pmatrix}$$

A solução desse sistema é $(2{,}75, -0{,}25, 0{,}25)$. O polinômio de grau dois que corresponde ao melhor ajuste de mínimos quadráticos para esses dados é

$$p(x) = 2{,}75 - 0{,}25x + 0{,}25x^2$$

□

EXERCÍCIOS

1. Encontre a solução de mínimos quadráticos para cada um dos sistemas a seguir.

(a)
$$\begin{aligned} x_1 + x_2 &= 3 \\ 2x_1 - 3x_2 &= 1 \\ 0x_1 + 0x_2 &= 2 \end{aligned}$$

(b)
$$\begin{aligned} -x_1 + x_2 &= 10 \\ 2x_1 + x_2 &= 5 \\ x_1 - 2x_2 &= 20 \end{aligned}$$

(c)
$$\begin{aligned} x_1 + x_2 + x_3 &= 4 \\ -x_1 + x_2 + x_3 &= 0 \\ - x_2 + x_3 &= 1 \\ x_1 \qquad + x_3 &= 2 \end{aligned}$$

2. Para cada uma das soluções $\hat{\mathbf{x}}$ encontradas no Exercício 1:
(a) Determine a projeção ortogonal $\mathbf{p} = A\hat{\mathbf{x}}$;
(b) Calcule o resíduo $r(\hat{\mathbf{x}})$;
(c) Verifique que $r(\hat{\mathbf{x}}) \in N(A^T)$.

3. Para cada um dos sistemas $A\mathbf{x} = \mathbf{b}$ a seguir, encontre todas as soluções de mínimos quadráticos.

(a) $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ -1 & -2 \end{pmatrix}$, $\quad \mathbf{b} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$

(b) $A = \begin{pmatrix} 1 & 1 & 3 \\ -1 & 3 & 1 \\ 1 & 2 & 4 \end{pmatrix}$, $\quad \mathbf{b} = \begin{pmatrix} -2 \\ 0 \\ 8 \end{pmatrix}$

4. Para cada um dos sistemas no Exercício 3, determine a projeção ortogonal $\mathbf{p}$ de $\mathbf{b}$ sobre $I(A)$ e verifique que $\mathbf{b} - \mathbf{p}$ é ortogonal a cada uma das colunas de A.

5. (a) Encontre o melhor ajuste de mínimos quadráticos por uma função linear para os dados

x	-1	0	1	2
y	0	1	3	9

(b) Coloque em um sistema de coordenadas a função linear encontrada no item (a) junto com os dados.

6. Encontre o melhor ajuste de mínimos quadráticos por um polinômio de grau dois para os dados no Exercício 5. Coloque os pontos correspondentes a $x = -1, 0, 1, 2$ e desenhe o gráfico da função.

7. Seja A uma matriz $m \times n$ de posto r e seja $P = A(A^TA)^{-1}A^T$.
(a) Mostre que $P\mathbf{b} = \mathbf{b}$ para todo $\mathbf{b} \in I(A)$. Explique essa propriedade em termos de projeções.
(b) Se $\mathbf{b} \in I(A)^{\perp}$, mostre que $P\mathbf{b} = \mathbf{0}$.
(c) Ilustre geometricamente os itens (a) e (b) no caso em que $I(A)$ é um plano contendo a origem em R^3.

8. Seja $P = A(A^TA)^{-1}A^T$, onde A é uma matriz $m \times n$ de posto n.
(a) Mostre que $P^2 = P$.
(b) Prove que $P^k = P$ para $k = 1, 2, \ldots$.
(c) Mostre que P é simétrica.

[***Lembre:*** Se B é invertível, então $(B^{-1})^T = (B^T)^{-1}$.]

9. Mostre que, se

$$\begin{pmatrix} A & I \\ O & A^T \end{pmatrix}\begin{pmatrix} \hat{\mathbf{x}} \\ \mathbf{r} \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ \mathbf{0} \end{pmatrix}$$

então $\hat{\mathbf{x}}$ é a solução de mínimos quadráticos para o sistema $A\mathbf{x} = \mathbf{b}$ e $\mathbf{r}$ é o vetor resíduo.

10. Seja $A \in R^{m\times n}$ e seja $\hat{\mathbf{x}}$ uma solução para o problema de mínimos quadráticos para $A\mathbf{x} = \mathbf{b}$. Mostre que um vetor $\mathbf{y} \in R^n$ é também solução se e somente se $\mathbf{y} = \hat{\mathbf{x}} + \mathbf{z}$ para algum $\mathbf{z} \in N(A)$.

[***Sugestão:*** $N(A^TA) = N(A)$.]

5 CONJUNTOS ORTONORMAIS

É geralmente mais conveniente usar a base canônica $\{\mathbf{e}_1, \mathbf{e}_2\}$ para R^2 do que alguma outra base, como $\{(2, 1)^T, (3, 5)^T\}$. Por exemplo, é mais fácil encontrar as coordenadas de $(x_1, x_2)^T$ em relação à base canônica. Os elementos da base canônica são vetores unitários ortogonais. Ao trabalhar em um espaço V munido de produto interno, é geralmente desejável usar uma base com vetores unitários ortogonais dois a dois. Isso é conveniente não apenas para encontrar coordenadas de vetores, mas para resolver problemas de mínimos quadráticos.

Definição. Sejam $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ vetores em um espaço V munido de produto interno. Se $\langle\mathbf{v}_i, \mathbf{v}_j\rangle = 0$ sempre que $i \neq j$, então $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ é dito um **conjunto ortogonal** de vetores.

EXEMPLO 1. O conjunto $\{(1, 1, 1)^T, (2, 1, -3)^T, (4, -5, 1)^T\}$ é um conjunto ortogonal em R^3, já que

$$(1, 1, 1)(2, 1, -3)^T = 0$$
$$(1, 1, 1)(4, -5, 1)^T = 0$$
$$(2, 1, -3)(4, -5, 1)^T = 0 \qquad \square$$

Teorema 5.5.1. *Se $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ é um conjunto ortogonal de vetores não-nulo em um espaço V munido de um produto interno, então $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são linearmente independentes.*

Demonstração. Suponha que $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são vetores não-nulos ortogonais dois a dois e que

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0} \tag{1}$$

Se $1 \leq j \leq n$, então, fazendo o produto interno de $\mathbf{v}_j$ com os dois lados da Equação (1), vemos que

$$c_1\langle\mathbf{v}_j, \mathbf{v}_1\rangle + c_2\langle\mathbf{v}_j, \mathbf{v}_2\rangle + \cdots + c_n\langle\mathbf{v}_j, \mathbf{v}_n\rangle = 0$$
$$c_j\|\mathbf{v}_j\|^2 = 0$$

e, portanto, todos os escalares $c_1, c_2, \ldots, c_n$ têm que ser iguais a 0. □

Definição. Um conjunto **ortonormal** de vetores é um conjunto ortogonal de vetores unitários.
O conjunto $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ é ortonormal se e somente se

$$\langle\mathbf{u}_i, \mathbf{u}_j\rangle = \delta_{ij}$$

onde

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

Dado qualquer conjunto ortogonal de vetores não-nulos $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$, é possível formar um conjunto ortonormal definindo

$$\mathbf{u}_i = \left(\frac{1}{\|\mathbf{v}_i\|}\right)\mathbf{v}_i \quad \text{para} \quad i = 1, 2, \ldots, n$$

O leitor pode verificar que $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ é um conjunto ortonormal.

EXEMPLO 2. Vimos, no Exemplo 1, que, se $\mathbf{v}_1 = (1, 1, 1)^T$, $\mathbf{v}_2 = (2, 1, -3)^T$ e $\mathbf{v}_3 = (4, -5, 1)^T$, então $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ é um conjunto ortogonal em R^3. Para formar um conjunto ortonormal, defina

$$\mathbf{u}_1 = \left(\frac{1}{\|\mathbf{v}_1\|}\right)\mathbf{v}_1 = \frac{1}{\sqrt{3}}(1, 1, 1)^T$$
$$\mathbf{u}_2 = \left(\frac{1}{\|\mathbf{v}_2\|}\right)\mathbf{v}_2 = \frac{1}{\sqrt{14}}(2, 1, -3)^T$$
$$\mathbf{u}_3 = \left(\frac{1}{\|\mathbf{v}_3\|}\right)\mathbf{v}_3 = \frac{1}{\sqrt{42}}(4, -5, 1)^T$$

□

EXEMPLO 3. Em $C[-\pi, \pi]$ com o produto interno

$$\langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)\,dx \tag{2}$$

o conjunto $\{1, \cos x, \cos 2x, \ldots, \cos nx\}$ é um conjunto ortogonal de vetores, já que, quaisquer que sejam os inteiros positivos j e k, temos

$$\langle 1, \cos kx\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos kx\,dx = 0$$
$$\langle \cos jx, \cos kx\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos jx \cos kx\,dx = 0 \qquad (j \neq k)$$

As funções $\cos x, \cos 2x, \ldots, \cos nx$ já são vetores unitários, pois

$$\langle \cos kx, \cos kx\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos^2 kx\,dx = 1 \quad \text{para} \quad k = 1, 2, \ldots, n$$

Para formar um conjunto ortonormal, basta encontrar um vetor unitário na direção de 1.

$$\|1\|^2 = \langle 1, 1\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} 1\,dx = 2$$

Logo, $1/\sqrt{2}$ é um vetor unitário e, portanto, $\{1/\sqrt{2}, \cos x, \cos 2x, \ldots, \cos nx\}$ é um conjunto ortonormal de vetores. □

Pelo Teorema 5.5.1, se $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ é um conjunto ortonormal em um espaço V munido de um produto interno, então B é uma base para um subespaço S de V. Dizemos que B é uma *base ortonormal* para S. É, em geral, muito mais fácil trabalhar com uma base ortonormal do que com uma base qualquer. Em particular, é muito mais fácil calcular as coordenadas de um vetor dado $\mathbf{v}$ em relação a uma base ortonormal. Uma vez determinadas essas coordenadas, elas podem ser usadas para calcular $\|\mathbf{v}\|$.

Teorema 5.5.2. *Seja $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ uma base ortonormal para um espaço V munido de um produto interno. Se $\mathbf{v} = \sum_{i=1}^{n} a_i\mathbf{u}_i$, então $a_i = <\mathbf{u}_i, \mathbf{v}>$.*

Demonstração

$$\langle \mathbf{u}_i, \mathbf{v}\rangle = \left\langle \mathbf{u}_i, \sum_{j=1}^{n} a_j\mathbf{u}_j \right\rangle = \sum_{j=1}^{n} a_j\langle \mathbf{u}_i, \mathbf{u}_j\rangle = \sum_{j=1}^{n} a_j\delta_{ij} = a_i$$ □

Como conseqüência do Teorema 5.5.2, temos mais dois resultados importantes.

Corolário 5.5.3. *Seja $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ uma base ortonormal para um espaço V munido de um produto interno. Se $\mathbf{u} = \sum_{i=1}^{n} a_i\mathbf{u}_i$ e $\mathbf{v} = \sum_{i=1}^{n} b_i\mathbf{u}_i$, então*

$$\langle \mathbf{u}, \mathbf{v}\rangle = \sum_{i=1}^{n} a_i b_i$$

Demonstração. Pelo Teorema 5.5.2,

$$\langle \mathbf{u}_i, \mathbf{v}\rangle = b_i \qquad i = 1, \ldots, n$$

Portanto,

$$\langle \mathbf{u}, \mathbf{v}\rangle = \left\langle \sum_{i=1}^{n} a_i\mathbf{u}_i, \mathbf{v} \right\rangle = \sum_{i=1}^{n} a_i\,\langle \mathbf{u}_i, \mathbf{v}\rangle = \sum_{i=1}^{n} a_i b_i$$ □

Corolário 5.5.4 (Fórmula de Parseval). *Se $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ é uma base ortonormal para um espaço V munido de um produto interno e se $\mathbf{v} = \sum_{i=1}^{n} a_i\mathbf{u}_i$, então*

$$\|\mathbf{v}\|^2 = \sum_{i=1}^{n} a_i^2$$

Demonstração. Se $\mathbf{v} = \sum_{i=1}^{n} a_i\mathbf{u}_i$, então, pelo Corolário 5.5.3,

$$\|\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{v}\rangle = \sum_{i=1}^{n} a_i^2$$ □

EXEMPLO 4. Os vetores

$$\mathbf{u}_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T \quad \text{e} \quad \mathbf{u}_2 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)^T$$

formam uma base ortonormal para R^2. Se $\mathbf{x} \in R^2$, então

$$\mathbf{x}^T\mathbf{u}_1 = \frac{x_1 + x_2}{\sqrt{2}} \quad \text{e} \quad \mathbf{x}^T\mathbf{u}_2 = \frac{x_1 - x_2}{\sqrt{2}}$$

Pelo Teorema 5.5.2, temos que

$$\mathbf{x} = \frac{x_1 + x_2}{\sqrt{2}}\mathbf{u}_1 + \frac{x_1 - x_2}{\sqrt{2}}\mathbf{u}_2$$

e, pelo Corolário 5.5.4, temos

$$\|\mathbf{x}\|^2 = \left(\frac{x_1 + x_2}{\sqrt{2}}\right)^2 + \left(\frac{x_1 - x_2}{\sqrt{2}}\right)^2 = x_1^2 + x_2^2$$

□

EXEMPLO 5. Dado que $\{1/\sqrt{2}, \cos 2x\}$ é um conjunto ortonormal em $C[-\pi, \pi]$ (em relação ao produto interno do Exemplo 3), determine o valor de $\int_{-\pi}^{\pi} \operatorname{sen}^4 x\, dx$ sem calcular derivadas.

SOLUÇÃO. Como

$$\operatorname{sen}^2 x = \frac{1 - \cos 2x}{2} = \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} + \left(-\frac{1}{2}\right)\cos 2x$$

obtemos, pela fórmula de Parseval,

$$\int_{-\pi}^{\pi} \operatorname{sen}^4 x\, dx = \pi\|\operatorname{sen}^2 x\|^2 = \pi\left(\frac{1}{2} + \frac{1}{4}\right) = \frac{3\pi}{4}$$

□

MATRIZES ORTOGONAIS

Particularmente importantes são as matrizes $n \times n$ cujas colunas formam um conjunto ortonormal em R^n.

Definição. Uma matriz Q $n \times n$ é dita **ortogonal** se seus vetores colunas formam um conjunto ortonormal em R^n.

Teorema 5.5.5. *Uma matriz Q $n \times n$ é ortogonal se e somente se $Q^TQ = I$.*

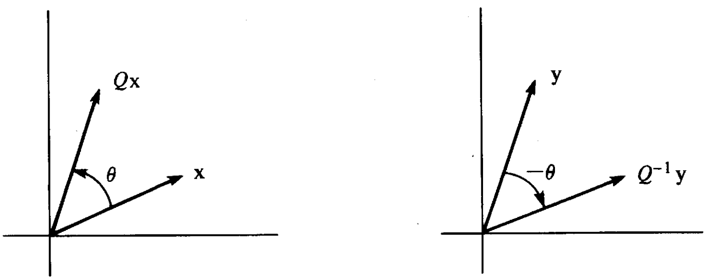


FIG. 5.5.1

Demonstração. Pela definição, uma matriz $n \times n$ Q é ortogonal se e somente se suas colunas satisfazem

$$\mathbf{q}_i^T\mathbf{q}_j = \delta_{ij}$$

No entanto, $\mathbf{q}_i^T\mathbf{q}_j$ é o elemento (i, j) de Q^TQ. Logo, Q é ortogonal se e somente se $Q^TQ = I$. □

Uma conseqüência do teorema é que, se Q é uma matriz ortogonal, então Q é invertível e $Q^{-1} = Q^T$.

EXEMPLO 6. Para qualquer θ fixo, a matriz

$$Q = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

é ortogonal e

$$Q^{-1} = Q^T = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix}$$ □

A matriz Q no Exemplo 6 pode ser considerada como uma transformação linear de R^2 sobre R^2 que roda cada vetor de um ângulo θ, mantendo fixo o comprimento do vetor (ver Exemplo 2 na Seção 2 do Cap. 4). Analogamente, Q^{-1} pode ser considerada como uma rotação por um ângulo $-\theta$ (veja Fig. 5.5.1).

Em geral, produtos internos são preservados sob multiplicação por um matriz ortogonal [isto é, $<\mathbf{x}, \mathbf{y}> = <Q\mathbf{x}, Q\mathbf{y}>$]. Isso é verdade porque

$$\langle Q\mathbf{x}, Q\mathbf{y}\rangle = (Q\mathbf{y})^TQ\mathbf{x} = \mathbf{y}^TQ^TQ\mathbf{x} = \mathbf{y}^T\mathbf{x} = \langle \mathbf{x}, \mathbf{y}\rangle$$

Em particular, se $\mathbf{x} = \mathbf{y}$, então $\|Q\mathbf{x}\|^2 = \|\mathbf{x}\|^2$, logo $\|Q\mathbf{x}\| = \|\mathbf{x}\|$. A multiplicação por uma matriz ortogonal preserva a norma dos vetores.

Propriedades de Matrizes Ortogonais

Se Q é uma matriz ortogonal $n \times n$, então:

1. As colunas de Q formam uma base ortonormal para R^n.
2. $Q^TQ = I$
3. $Q^T = Q^{-1}$
4. $\langle Q\mathbf{x}, Q\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$
5. $\|Q\mathbf{x}\|_2 = \|\mathbf{x}\|_2$

MATRIZES DE PERMUTAÇÃO

Uma *matriz de permutação* é uma matriz obtida permutando-se as colunas da matriz identidade. É claro, portanto, que matrizes de permutação são ortogonais. Se P é a matriz de permutação obtida colocando-se as colunas de I na ordem $(k_1, k_2, \ldots, k_n)$, então $P = (\mathbf{e}_{k1}, \mathbf{e}_{k2}, \ldots, \mathbf{e}_{kn})$. Se A é uma matriz $m \times n$, então

$$AP = (A\mathbf{e}_{k_1}, \ldots, A\mathbf{e}_{k_n}) = (\mathbf{a}_{k_1}, \ldots, \mathbf{a}_{k_n})$$

A multiplicação de A à direita por P coloca as colunas de A na ordem $(k_1, k_2, \ldots, k_n)$. Por exemplo, se

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \quad \text{e} \quad P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

então

$$AP = \begin{pmatrix} 3 & 1 & 2 \\ 3 & 1 & 2 \end{pmatrix}$$

Como $P = (\mathbf{e}_{k1}, \mathbf{e}_{k2}, \ldots, \mathbf{e}_{kn})$ é ortogonal, temos que

$$P^{-1} = P^T = \begin{pmatrix} \mathbf{e}_{k_1}^T \\ \vdots \\ \mathbf{e}_{k_n}^T \end{pmatrix}$$

A coluna k_1 de P^T vai ser $\mathbf{e}_1$, a coluna k_2 vai ser $\mathbf{e}_2$, e assim por diante. Logo, P^T também é uma matriz de permutação. A matriz P^T pode ser formada diretamente colocando-se as linhas de I na ordem $(k_1, k_2, \ldots, k_n)$. Em geral, uma matriz de permutação pode ser formada permutando-se as linhas ou as colunas de I.

Se Q é a matriz de permutação formada colocando-se as linhas de I na ordem $(k_1, k_2, \ldots, k_n)$ e se B é uma matriz $n \times r$, então

$$QB = \begin{pmatrix} \mathbf{e}_{k_1}^T \\ \vdots \\ \mathbf{e}_{k_n}^T \end{pmatrix} B = \begin{pmatrix} \mathbf{e}_{k_1}^T B \\ \vdots \\ \mathbf{e}_{k_n}^T B \end{pmatrix} = \begin{pmatrix} \mathbf{b}(k_1, :) \\ \vdots \\ \mathbf{b}(k_n, :) \end{pmatrix}$$

Logo a matriz QB é formada colocando-se as linhas de B na ordem $(k_1, k_2, \ldots, k_n)$. Por exemplo, se

$$Q = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 3 \end{pmatrix}$$

então

$$QB = \begin{pmatrix} 3 & 3 \\ 1 & 1 \\ 2 & 2 \end{pmatrix}$$

Em geral, se P é uma matriz de permutação $n \times n$, a multiplicação de uma matriz B $n \times r$ por P à esquerda permuta as linhas de B e a multiplicação de uma matriz A $m \times n$ por P à direita permuta as colunas de A.

CONJUNTOS ORTONORMAIS E MÍNIMOS QUADRÁTICOS

A ortogonalidade tem um papel importante na solução de problemas de mínimos quadráticos. Lembre-se de que se A é uma matriz $m \times n$ de posto n, então o problema de mínimos quadráticos para $A\mathbf{x} = \mathbf{b}$ tem uma única solução $\hat{\mathbf{x}}$ que é obtida resolvendo-se as equações normais $A^TA\mathbf{x} = A^T\mathbf{b}$. A projeção ortogonal $\mathbf{p} = A\hat{\mathbf{x}}$ é o vetor em $I(A)$ mais próximo de $\mathbf{b}$. O problema de mínimos quadráticos é particularmente fácil de resolver no caso em que os vetores colunas de A formam um conjunto ortonormal em R^m.

Teorema 5.5.6. *Se as colunas de* A *formam um conjunto ortonormal de vetores em R^m, então $A^TA = I$ e a solução do problema de mínimos quadráticos é*

$$\hat{\mathbf{x}} = A^T\mathbf{b}$$

Demonstração. O elemento (i, j) da matriz A^TA é formado pela i-ésima linha de A^T e a j-ésima coluna de A. Portanto, esse elemento é, na verdade, o produto escalar das i-ésima e j-ésima colunas A. Como os vetores coluna de A são ortonormais, segue que

$$A^TA = (\delta_{ij}) = I$$

Conseqüentemente, as equações normais tomam a forma

$$\mathbf{x} = A^T\mathbf{b} \qquad \square$$

O que acontece quando as colunas de A não são ortonormais? Na próxima seção, aprenderemos um método que permite encontrar uma base ortonormal para $I(A)$. A partir disso, obteremos uma fatoração

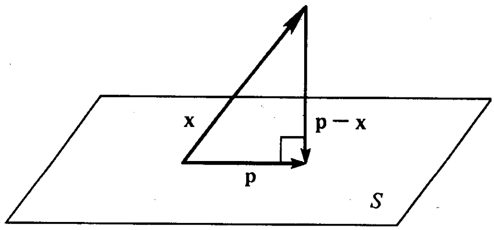


FIG. 5.5.2

de A em um produto QR, onde Q tem um conjunto ortonormal de colunas e R é triangular superior. Com essa fatoração, o problema dos mínimos quadráticos pode ser resolvido de maneira rápida e precisa.

Uma vez conhecida uma base ortonormal para $I(A)$, é possível escrever a projeção ortogonal $\mathbf{p} = A\hat{\mathbf{x}}$ em termos dos elementos da base. Na verdade, esse é um caso especial do problema de mínimos quadráticos mais geral que consiste em determinar o elemento $\mathbf{p}$ de um subespaço S de um espaço V munido de produto interno que se encontra à menor distância de dado elemento $\mathbf{x}$ de V. Esse problema é fácil de resolver se dispomos de uma base ortonormal para S. Em primeiro lugar, provaremos o seguinte teorema:

Teorema 5.5.7. *Seja S um subespaço do espaço V munido de produto interno e seja $\mathbf{x} \in V$. Seja $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ uma base ortonormal para S. Se*

(3)
$$\mathbf{p} = \sum_{i=1}^{n} c_i \mathbf{x}_i$$

onde

(4)
$$c_i = \langle \mathbf{x}, \mathbf{x}_i \rangle \qquad \textit{para todo } i$$

então $\mathbf{p} - \mathbf{x} \in S^{\perp}$ (ver Fig. 5.5.2).

Demonstração. Vamos provar primeiro que $(\mathbf{p} - \mathbf{x}) \perp \mathbf{x}_i$ para todo i.

$$\begin{aligned}\langle \mathbf{x}_i, \mathbf{p} - \mathbf{x} \rangle &= \langle \mathbf{x}_i, \mathbf{p} \rangle - \langle \mathbf{x}_i, \mathbf{x} \rangle \\ &= \left\langle \mathbf{x}_i, \sum_{j=1}^{n} c_j \mathbf{x}_j \right\rangle - c_i \\ &= \sum_{j=1}^{n} c_j \langle \mathbf{x}_i, \mathbf{x}_j \rangle - c_i \\ &= 0\end{aligned}$$

Portanto, $\mathbf{p} - \mathbf{x}$ é ortogonal a todos os $\mathbf{x}_i$. Se $\mathbf{y} \in S$, então

$$\mathbf{y} = \sum_{i=1}^{n} \alpha_i \mathbf{x}_i$$

de modo que

$$\langle \mathbf{p} - \mathbf{x}, \mathbf{y} \rangle = \left\langle \mathbf{p} - \mathbf{x}, \sum_{i=1}^{n} \alpha_i \mathbf{x}_i \right\rangle = \sum_{i=1}^{n} \alpha_i \langle \mathbf{p} - \mathbf{x}, \mathbf{x}_i \rangle = 0 \qquad \square$$

Se $\mathbf{x} \in S$, o resultado acima é trivial uma vez que, pelo Teorema 5.5.2, $\mathbf{p} - \mathbf{x} = \mathbf{0}$. Se $\mathbf{x} \notin S$, então $\mathbf{p}$ é o elemento de S mais próximo de $\mathbf{x}$.

Teorema 5.5.8. *Sob as hipóteses do Teorema 5.5.7,* $\mathbf{p}$ *é o elemento de S mais próximo de* $\mathbf{x}$, *isto é,*

$$\|\mathbf{y}-\mathbf{x}\| > \|\mathbf{p}-\mathbf{x}\|$$

para todo $\mathbf{y} \neq \mathbf{p}$ *em S.*

Demonstração. Se $\mathbf{y} \in S$ e $\mathbf{y} \neq \mathbf{p}$, então

$$\|\mathbf{y}-\mathbf{x}\|^2 = \|(\mathbf{y}-\mathbf{p}) + (\mathbf{p}-\mathbf{x})\|^2$$

Como $\mathbf{y} - \mathbf{p} \in S$, o Teorema 5.5.7 e o teorema de Pitágoras implicam que

$$\|\mathbf{y}-\mathbf{x}\|^2 = \|\mathbf{y}-\mathbf{p}\|^2 + \|\mathbf{p}-\mathbf{x}\|^2 > \|\mathbf{p}-\mathbf{x}\|^2$$

Portanto, $\|\mathbf{y} - \mathbf{x}\| > \|\mathbf{p} - \mathbf{x}\|$. □

O vetor $\mathbf{p}$ definido por (3) e (4) é chamado de *projeção ortogonal de* $\mathbf{x}$ *sobre S.*

Corolário 5.5.9. *Sejam S um subespaço não-nulo de* R^m *e* $\mathbf{b} \in R^m$. Se $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ *é uma base ortonormal para S e* $U = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k)$, *então a projeção ortogonal* $\mathbf{b}$ *de* $\mathbf{p}$ *sobre S é dada por*

$$\mathbf{p} = UU^T\mathbf{b}$$

Demonstração. Pelo Teorema 5.5.8, a projeção ortogonal $\mathbf{b}$ de $\mathbf{p}$ sobre S é dada por

$$\mathbf{p} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k = U\mathbf{a}$$

onde

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1^T\mathbf{b} \\ \mathbf{u}_2^T\mathbf{b} \\ \vdots \\ \mathbf{u}_k^T\mathbf{b} \end{pmatrix} = U^T\mathbf{b}$$

Portanto,

$$\mathbf{p} = UU^T\mathbf{b}$$ □

A matriz UU^T que aparece no Corolário 5.5.9 é a matriz de projeção correspondente ao subespaço S do R^m. Para projetar qualquer vetor $\mathbf{b} \in R^m$ sobre S precisamos conhecer apenas uma base ortonormal $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ de S, formar a matriz UU^T e multiplicá-la por $\mathbf{b}$.

Se P é a matriz de projeção correspondente a um subespaço S do R^m, então a projeção $\mathbf{p}$ de $\mathbf{b}$ sobre S é única. Se Q também é uma matriz de projeção correspondendo a S, então

$$Q\mathbf{b} = \mathbf{p} = P\mathbf{b}$$

Segue que

$$\mathbf{q}_j = Q\mathbf{e}_j = P\mathbf{e}_j = \mathbf{p}_j \qquad \text{para} \quad j = 1, \ldots, m$$

e, portanto, $Q = P$. Logo, a matriz de projeção sobre um subespaço S do R^m é única.

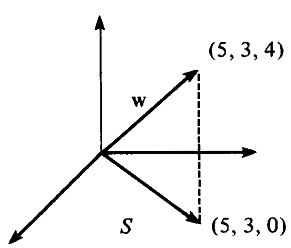


FIG. 5.5.3

EXEMPLO 7. Seja S o conjunto de todos os vetores em R^3 da forma $(x, y, 0)^T$. Encontre o vetor $\mathbf{p}$ em S mais próximo do vetor $\mathbf{w} = (5, 3, 4)^T$ (ver Fig. 5.5.3).

SOLUÇÃO. Sejam $\mathbf{u}_1 = (1, 0, 0)^T$ e $\mathbf{u}_2 = (0, 1, 0)^T$. É claro que $\mathbf{u}_1$ e $\mathbf{u}_2$ formam uma base ortonormal para S. Agora

$$c_1 = \mathbf{w}^T\mathbf{u}_1 = 5$$

$$c_2 = \mathbf{w}^T\mathbf{u}_2 = 3$$

O vetor $\mathbf{p}$ é exatamente o que era de se esperar que fosse:

$$\mathbf{p} = 5\mathbf{u}_1 + 3\mathbf{u}_2 = (5, 3, 0)^T$$

Uma maneira alternativa para a obtenção de $\mathbf{p}$ consiste em utilizar a matriz de projeção UU^T.

$$\mathbf{p} = UU^T\mathbf{w} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 5 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \\ 0 \end{pmatrix}$$ □

APROXIMAÇÃO DE FUNÇÕES

Em muitas aplicações é necessário aproximar uma função contínua utilizando os elementos de um conjunto de funções de um tipo especial. A aproximação mais comum consiste em utilizar polinômios de grau menor ou igual a um n fixo. Podemos usar o Teorema 5.5.8 para obter a melhor aproximação de mínimos quadráticos.

EXEMPLO 8. Encontre a melhor aproximação, por uma função linear, de mínimos quadráticos para a função e^x no intervalo $[0, 1]$.

SOLUÇÃO. Seja S o subespaço de todas as funções lineares pertencentes a $C[0, 1]$. Esse subespaço é gerado pelas funções 1 e x, que não são ortogonais entre si. Vamos procurar uma função da forma $x - a$ que seja ortogonal a 1.

$$\langle 1, x - a\rangle = \int_0^1 (x - a)\,dx = \tfrac{1}{2} - a$$

Logo $a = 1/2$. Como $\|x - 1/2\| = 1/\sqrt{12}$, segue que

$$u_1(x) = 1 \quad \text{e} \quad u_2(x) = \sqrt{12}\,(x - \tfrac{1}{2})$$

formam uma base ortogonal para S.

Sejam

$$a_1 = \int_0^1 u_1(x)\,e^x\,dx = e - 1$$

$$a_2 = \int_0^1 u_2(x)\,e^x\,dx = \sqrt{3}\,(3 - e)$$

A projeção

$$\begin{aligned} p(x) &= a_1u_1(x) + a_2u_2(x) \\ &= (e - 1)\cdot 1 + \sqrt{3}(3 - e)\left[\sqrt{12}(x - \tfrac{1}{2})\right] \\ &= (4e - 10) + 6(3 - e)x \end{aligned}$$

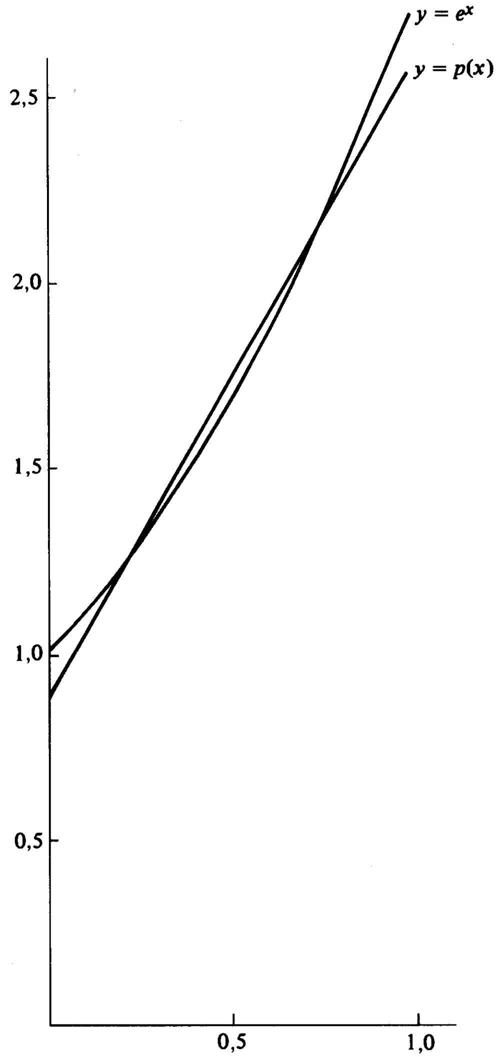


FIG. 5.5.4

é a melhor aproximação linear, no sentido dos mínimos quadráticos, de e^x no intervalo [0, 1] (ver Fig. 5.5.4). □

Aproximação por Polinômios Trigonométricos

Um *polinômio trigonométrico* de grau n é uma função da forma

$$t(x) = \frac{a_0}{2} + \sum_{k=1}^{n}(a_k \cos kx + b_k \operatorname{sen} kx)$$

Já sabemos que a coleção

$$\frac{1}{\sqrt{2}}, \cos x, \cos 2x, \ldots, \cos nx$$

forma um conjunto ortonormal em relação ao produto interno (2). Deixaremos a cargo do leitor a verificação do fato de que, ao adicionarmos as funções

$$\operatorname{sen} x, \operatorname{sen} 2x, \ldots, \operatorname{sen} nx$$

ao conjunto acima, obtemos outro conjunto ortonormal. Podemos, portanto, utilizar o Teorema 5.5.8 para encontrar a melhor aproximação de uma função contínua em termos de um polinômio trigonométrico de grau menor ou igual a um n dado, no sentido dos mínimos quadrados. Observe que

$$\left\langle f, \frac{1}{\sqrt{2}} \right\rangle \frac{1}{\sqrt{2}} = \langle f, 1 \rangle \frac{1}{2}$$

de modo que, se

$$a_0 = \langle f, 1 \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, dx$$

e

$$a_k = \langle f, \cos kx \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx\, dx$$

$$b_k = \langle f, \operatorname{sen} kx \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \operatorname{sen} kx\, dx$$

para $k = 1, 2, \ldots, n$, então esses coeficientes determinam a melhor aproximação de f por mínimos quadráticos. Os a_j e b_k são *coeficientes de Fourier*, bastante conhecidos, que aparecem em diversas aplicações envolvendo aproximação de funções por séries trigonométricas.

EXERCÍCIOS

1. Quais dos conjuntos de vetores a seguir formam uma base ortonormal para R^2?

(a) $\{(1, 0)^T, (0, 1)^T\}$

(b) $\{(\frac{3}{5}, \frac{4}{5})^T, (\frac{5}{13}, \frac{12}{13})^T\}$

(c) $\{(1, -1)^T, (1, 1)^T\}$

(d) $\left\{ \left(\frac{\sqrt{3}}{2}, \frac{1}{2} \right)^T, \left(-\frac{1}{2}, \frac{\sqrt{3}}{2} \right)^T \right\}$

2. Sejam

$$\mathbf{u}_1 = \left(\frac{1}{3\sqrt{2}}, \frac{1}{3\sqrt{2}}, -\frac{4}{3\sqrt{2}} \right)^T, \qquad \mathbf{u}_2 = \left(\frac{2}{3}, \frac{2}{3}, \frac{1}{3} \right)^T$$

$$\mathbf{u}_3 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0 \right)^T$$

(a) Mostre que $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ formam uma base ortonormal para R^3.

(b) Seja $\mathbf{x} = (1, 1, 1)^T$. Escreva $\mathbf{x}$ como uma combinação linear de $\mathbf{u}_1$, $\mathbf{u}_2$ e $\mathbf{u}_3$ usando o Teorema 5.5.2 e use a fórmula de Parseval para calcular $\|\mathbf{x}\|$.

3. Seja S o subespaço de R^3 gerado pelos vetores $\mathbf{u}_2$ e $\mathbf{u}_3$ do Exercício 2. Seja $\mathbf{x} = (1, 2, 2)^T$. Encontre a projeção ortogonal $\mathbf{p}$ de $\mathbf{x}$ sobre S. Mostre que $(\mathbf{p} - \mathbf{x}) \perp \mathbf{u}_2$ e $(\mathbf{p} - \mathbf{x}) \perp \mathbf{u}_3$.

4. Seja θ um número real fixo e sejam

$$\mathbf{x}_1 = \begin{pmatrix} \cos\theta \\ \operatorname{sen}\theta \end{pmatrix} \quad \text{e} \quad \mathbf{x}_2 = \begin{pmatrix} -\operatorname{sen}\theta \\ \cos\theta \end{pmatrix}$$

(a) Mostre que $\{\mathbf{x}_1, \mathbf{x}_2\}$ é uma base ortonormal para R^2.
(b) Escreva um vetor arbitrário $\mathbf{y}$ em R^2 como uma combinação linear $c_1\mathbf{x}_1 + c_2\mathbf{x}_2$.
(c) Verifique que

$$c_1^2 + c_2^2 = \|\mathbf{y}\|^2 = y_1^2 + y_2^2$$

5. Suponha que $\mathbf{u}_1$ e $\mathbf{u}_2$ formam uma base ortonormal para R^2 e seja $\mathbf{u}$ um vetor unitário em R^2. Se $\mathbf{u}^T\mathbf{u}_1 = 1/2$, determine o valor de $|\mathbf{u}^T\mathbf{u}_2|$.

6. Seja $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ uma base ortonormal para um espaço V munido de produto interno e sejam

$$\mathbf{u} = \mathbf{u}_1 + 2\mathbf{u}_2 + 2\mathbf{u}_3 \qquad \text{e} \qquad \mathbf{v} = \mathbf{u}_1 + 7\mathbf{u}_3$$

Determine o valor de:
(a) $\langle\mathbf{u}, \mathbf{v}\rangle$;
(b) $\|\mathbf{u}\|$ e $\|\mathbf{v}\|$;
(c) o ângulo θ entre $\mathbf{u}$ e $\mathbf{v}$.

7. As funções $\cos x$ e $\operatorname{sen} x$ formam um conjunto ortonormal em $C[-\pi, \pi]$. Se

$$f(x) = 3\cos x + 2\operatorname{sen} x \qquad \text{e} \qquad g(x) = \cos x - \operatorname{sen} x$$

use o Corolário 5.5.3 para determinar o valor de

$$\langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)\,dx$$

8. O conjunto

$$S = \left\{\frac{1}{\sqrt{2}}, \cos x, \cos 2x, \cos 3x, \cos 4x\right\}$$

é um conjunto ortonormal de vetores em $C[-\pi, \pi]$ em relação ao produto interno definido por (2).
(a) Use identidades trigonométricas para escrever a função $\operatorname{sen}^4 x$ como uma combinação linear de elementos de S.
(b) Use o item (a) e o Teorema 5.5.2 para encontrar os valores das integrais a seguir.

$$\text{(i)}\ \int_{-\pi}^{\pi} \operatorname{sen}^4 x \cos x\,dx \qquad \text{(ii)}\ \int_{-\pi}^{\pi} \operatorname{sen}^4 x \cos 2x\,dx$$
$$\text{(iii)}\ \int_{-\pi}^{\pi} \operatorname{sen}^4 x \cos 3x\,dx \qquad \text{(iv)}\ \int_{-\pi}^{\pi} \operatorname{sen}^4 x \cos 4x\,dx$$

9. Prove que a transposta de uma matriz ortogonal é uma matriz ortogonal.

10. Se Q é uma matriz ortogonal $n \times n$ e se $\mathbf{x}$ e $\mathbf{y}$ são vetores não-nulos em R^n, qual a relação entre o ângulo entre $Q\mathbf{x}$ e $Q\mathbf{y}$ e o ângulo entre $\mathbf{x}$ e $\mathbf{y}$? Prove.

11. Seja Q uma matriz ortogonal $n \times n$. Use indução matemática para provar cada uma das afirmações a seguir.
(a) $(Q^m)^{-1} = (Q^T)^m = (Q^m)^T$ para todo inteiro positivo m.
(b) $\|Q^m\mathbf{x}\| = \|\mathbf{x}\|$ para todo $\mathbf{x} \in R^n$.

12. Seja $\mathbf{u}$ um vetor unitário em R^n e seja $H = I - 2\mathbf{u}\mathbf{u}^T$. Mostre que H é ao mesmo tempo ortogonal e simétrica e, portanto, sua própria inversa.

13. Seja Q uma matriz ortogonal e seja $d = \det(Q)$. Mostre que $|d| = 1$.

14. Mostre que o produto de duas matrizes ortogonais é ortogonal. O produto de duas matrizes de permutação é uma matriz de permutação? Explique.

15. Mostre que, se U é uma matriz ortogonal $n \times n$, então

$$\mathbf{u}_1\mathbf{u}_1^T + \mathbf{u}_2\mathbf{u}_2^T + \cdots + \mathbf{u}_n\mathbf{u}_n^T = I$$

16. Use indução matemática para mostrar que, se $Q \in R^{n\times n}$ é, ao mesmo tempo, triangular superior e ortogonal, então $\mathbf{q}_j = \pm\,\mathbf{e}_j$, $j = 1, \ldots, n$.

17. Seja

$$A = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

(a) Mostre que as colunas de A formam um conjunto ortonormal em R^4.
(b) Resolva o problema de mínimos quadráticos para $A\mathbf{x} = \mathbf{b}$ para cada uma das escolhas de $\mathbf{b}$ a seguir.

(i) $\mathbf{b} = (4, 0, 0, 0)^T$
(ii) $\mathbf{b} = (1, 2, 3, 4)^T$
(iii) $\mathbf{b} = (1, 1, 2, 2)^T$

18. Seja A a matriz do Exercício 17.
(a) Encontre a matriz de projeção P que projeta ortogonalmente vetores em R^4 sobre $I(A)$.
(b) Para cada uma das soluções $\mathbf{x}$ encontradas no Exercício 17(b), calcule $A\mathbf{x}$ e compare com $P\mathbf{b}$.

19. Seja A a matriz do Exercício 17.
(a) Encontre uma base ortonormal para $N(A^T)$.
(b) Determine a matriz de projeção que projeta ortogonalmente vetores em R^4 sobre $N(A^T)$.

20. Sejam A uma matriz $m \times n$, P a matriz de projeção que projeta ortogonalmente vetores em R^m sobre $I(A)$ e Q a matriz de projeção que projeta ortogonalmente vetores em R^n sobre $I(A^T)$. Mostre que:
(a) $I - P$ é a matriz de projeção de R^m sobre $N(A^T)$;
(b) $I - Q$ é a matriz de projeção de R^n sobre $N(A)$.

21. Seja P a matriz de projeção correspondente a um subespaço S de R^m. Mostre que:

(a) $P^2 = P$ (b) $P^T = P$

22. Seja A uma matriz $m \times n$ cujos vetores colunas são ortogonais dois a dois e seja $\mathbf{b} \in R^m$. Mostre que, se $\hat{\mathbf{x}}$ é a solução de mínimos quadráticos para o sistema $A\mathbf{x} = \mathbf{b}$, então

$$\hat{x}_i = \frac{\mathbf{b}^T \mathbf{a}_i}{\mathbf{a}_i^T \mathbf{a}_i} \qquad i = 1, \ldots, n$$

23. Considere o espaço vetorial $C[-1, 1]$ com o produto interno

$$\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)\,dx$$

e a norma

$$\|f\| = (\langle f, f \rangle)^{1/2}$$

(a) Mostre que os vetores 1 e x são ortogonais.
(b) Calcule $\|1\|$ e $\|x\|$.
(c) Encontre a melhor aproximação de mínimos quadráticos, por uma função linear $l(x) = c_1 1 + c_2 x$, de $x^{1/3}$ em $[-1, 1]$.
(d) Esboce os gráficos de $x^{1/3}$ e de $l(x)$ em $[-1, 1]$.

24. Considere o espaço $C[0, 1]$ munido do produto interno

$$\langle f, g \rangle = \int_{0}^{1} f(x)g(x)\,dx$$

Seja S o subespaço gerado pelos vetores 1 e $2x - 1$.

(a) Mostre que 1 e $2x - 1$ são ortogonais.
(b) Determine $\|1\|$ e $\|2x - 1\|$.
(c) Encontre a melhor aproximação de mínimos quadráticos para $\sqrt{x}$ por uma função pertencente ao subespaço S.

25. Seja

$$S = \{1/\sqrt{2}, \cos x, \cos 2x, \ldots, \cos nx, \operatorname{sen} x, \operatorname{sen} 2x, \ldots, \operatorname{sen} nx\}$$

Mostre que S é um conjunto ortonormal em $C[-\pi, \pi]$ munido do produto interno definido por (2).

26. Encontre a melhor aproximação de mínimos quadráticos para $f(x) = |x|$ em $[-\pi, \pi]$ por um polinômio trigonométrico de grau menor ou igual a 2.

27. Seja $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k, \mathbf{x}_{k+1}, \ldots, \mathbf{x}_n\}$ uma base ortonormal para um espaço V munido de um produto interno. Seja S_1 o subespaço gerado por $\mathbf{x}_1, \ldots, \mathbf{x}_k$ e seja S_2 o subespaço gerado por $\mathbf{x}_{k+1}, \mathbf{x}_{k+2}, \ldots, \mathbf{x}_n$. Mostre que $S_1 \perp S_2$.

28. Seja $\mathbf{x}$ um elemento do espaço vetorial V do Exercício 27 e sejam $\mathbf{p}_1$ e $\mathbf{p}_2$ as projeções ortogonais de $\mathbf{x}$ sobre S_1 e S_2, respectivamente. Mostre que:
(a) $\mathbf{x} = \mathbf{p}_1 + \mathbf{p}_2$;
(b) se $\mathbf{x} \in S_1^{\perp}$, então $\mathbf{p}_1 = \mathbf{0}$ e, portanto, $S^{\perp} = S_2$.

29. Seja S um subespaço de um espaço V munido de um produto interno. Seja $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ uma base ortogonal para S e seja $\mathbf{x} \in V$. Mostre que a melhor aproximação de $\mathbf{x}$ de mínimos quadráticos por elementos de S é dada por

$$\mathbf{p} = \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{x}_i \rangle}{\langle \mathbf{x}_i, \mathbf{x}_i \rangle} \mathbf{x}_i$$

6 O PROCESSO DE ORTOGONALIZAÇÃO DE GRAM-SCHMIDT

Nesta seção, vamos aprender um processo para a construção de uma base ortonormal para um espaço V de dimensão n munido de um produto interno. Começando com um base $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$, o método envolve a utilização de projeções ortogonais para a construção de uma base ortonormal $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$.

Vamos construir os $\mathbf{u}_i$ de modo que $[\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}] = [\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}]$ para $k = 1, \ldots, n$. Para começar o processo, seja

$$\mathbf{u}_1 = \left(\frac{1}{\|\mathbf{x}_1\|}\right)\mathbf{x}_1 \tag{1}$$

$[\{\mathbf{u}_1\}] = [\{\mathbf{x}_1\}]$, já que $\mathbf{u}_1$ é um vetor unitário com mesma direção que $\mathbf{x}_1$. Seja $\mathbf{p}_1$ a projeção ortogonal de $\mathbf{x}_2$ sobre $[\{\mathbf{x}_1\}] = [\{\mathbf{u}_1\}]$.

$$\mathbf{p}_1 = \langle \mathbf{x}_2, \mathbf{u}_1 \rangle \mathbf{u}_1$$

Pelo Teorema 5.5.7,

$$(\mathbf{x}_2 - \mathbf{p}_1) \perp \mathbf{u}_1$$

Observe que $\mathbf{x}_2 - \mathbf{p}_1 \neq \mathbf{0}$, já que

$$\mathbf{x}_2 - \mathbf{p}_1 = \frac{-\langle \mathbf{x}_2, \mathbf{u}_1 \rangle}{\|\mathbf{x}_1\|}\mathbf{x}_1 + \mathbf{x}_2 \tag{2}$$

e $\mathbf{x}_1$ e $\mathbf{x}_2$ são linearmente independentes. Definindo

$$\mathbf{u}_2 = \frac{1}{\|\mathbf{x}_2 - \mathbf{p}_1\|}(\mathbf{x}_2 - \mathbf{p}_1) \tag{3}$$

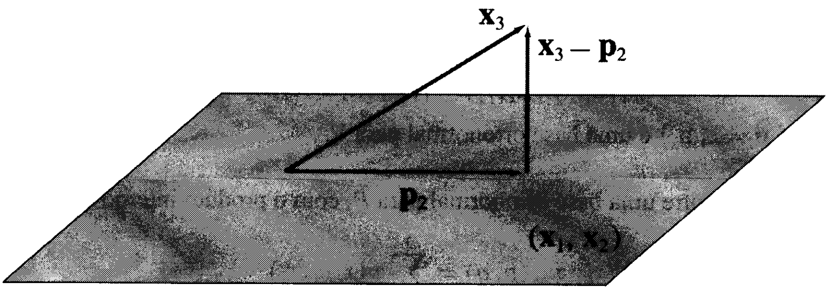


FIG. 5.6.1

temos que $\mathbf{u}_2$ é um vetor unitário ortogonal a $\mathbf{u}_1$. De (1), (2) e (3), temos que $[\{\mathbf{u}_1, \mathbf{u}_2\}] \subset [\{\mathbf{x}_1, \mathbf{x}_2\}]$. Como $\mathbf{u}_1$ e $\mathbf{u}_2$ são linearmente independentes, $\{\mathbf{u}_1, \mathbf{u}_2\}$ é uma base ortonormal para $[\{\mathbf{x}_1, \mathbf{x}_2\}]$, logo $[\{\mathbf{x}_1, \mathbf{x}_2\}] = [\mathbf{u}_1, \mathbf{u}_2\}]$.

Para construir $\mathbf{u}_3$, continue do mesmo modo. Seja $\mathbf{p}_2$ a projeção ortogonal de $\mathbf{x}_3$ sobre $[\{\mathbf{x}_1, \mathbf{x}_2\}] = [\{\mathbf{u}_1, \mathbf{u}_2\}]$,

$$\mathbf{p}_2 = \langle \mathbf{x}_3, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{x}_3, \mathbf{u}_2 \rangle \mathbf{u}_2$$

e defina

$$\mathbf{u}_3 = \frac{1}{\|\mathbf{x}_3 - \mathbf{p}_2\|}(\mathbf{x}_3 - \mathbf{p}_2)$$

e assim por diante (ver Fig. 5.6.1).

Teorema 5.6.1 (O Processo de Gram-Schmidt). *Seja* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *uma base para o espaço V munido de um produto interno. Seja*

$$\mathbf{u}_1 = \left(\frac{1}{\|\mathbf{x}_1\|}\right)\mathbf{x}_1$$

e defina $\mathbf{u}_2, \ldots, \mathbf{u}_n$ recursivamente por

$$\mathbf{u}_{k+1} = \frac{1}{\|\mathbf{x}_{k+1} - \mathbf{p}_k\|}(\mathbf{x}_{k+1} - \mathbf{p}_k) \text{ para } k = 1, \ldots, n-1$$

onde

$$\mathbf{p}_k = \langle \mathbf{x}_{k+1}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{x}_{k+1}, \mathbf{u}_2 \rangle \mathbf{u}_2 + \cdots + \langle \mathbf{x}_{k+1}, \mathbf{u}_k \rangle \mathbf{u}_k$$

é a projeção ortogonal de $\mathbf{x}_{k+1}$ *sobre* $[\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}]$. *O conjunto* $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ *é uma base ortonormal para V.*

Demonstração. Vamos argumentar indutivamente. É claro que $[\{\mathbf{u}_1\}] = [\{\mathbf{x}_1\}]$. Suponha que construímos $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ de modo que $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ seja um conjunto ortonormal e que

$$[\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}] = [\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}].$$

Como $\mathbf{p}_k$ é uma combinação linear dos $\mathbf{u}_i$ para $1 \leq i \leq k$, temos que $\mathbf{p}_k \in [\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}]$ e que $\mathbf{x}_{k+1} - \mathbf{p}_k \in \{\mathbf{x}_1, \ldots, \mathbf{x}_{k+1}\}$.

$$\mathbf{x}_{k+1} - \mathbf{p}_k = \mathbf{x}_{k+1} - \sum_{i=1}^{k} c_i \mathbf{x}_i$$

Como $\mathbf{x}_1, \ldots, \mathbf{x}_{k+1}$ são linearmente independentes, temos que $\mathbf{x}_{k+1} - \mathbf{p}_k$ é diferente de zero e, pelo Teorema 5.5.7, é ortogonal a cada um dos $\mathbf{u}_i$, $1 \leq i \leq k$. Portanto, $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{k+1}\}$ é um conjunto ortonormal de vetores em $[\{\mathbf{x}_1, \ldots, \mathbf{x}_{k+1}\}]$. Como $\mathbf{u}_1, \ldots, \mathbf{u}_{k+1}$ são linearmente independentes, eles formam uma base para $[\{\mathbf{x}_1, \ldots, \mathbf{x}_{k+1}\}]$, logo

$$[\{\mathbf{u}_1, \ldots, \mathbf{u}_{k+1}\}] = [\{\mathbf{x}_1, \ldots, \mathbf{x}_{k+1}\}].$$

Por indução, $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ é uma base ortonormal para V. □

EXEMPLO 1. Encontre uma base ortonormal para P_3 com o produto interno definido por

$$\langle p, q\rangle = \sum_{i=1}^{3} p(x_i)q(x_i)$$

onde $x_1 = -1$, $x_2 = 0$ e $x_3 = 1$.

SOLUÇÃO. Começando com a base $\{1, x, x^2\}$, podemos usar o processo de Gram-Schmidt para obter uma base ortonormal.

$$\|1\|^2 = \langle 1, 1\rangle = 3$$

de modo que

$$\mathbf{u}_1 = \left(\frac{1}{\|1\|}\right)1 = \frac{1}{\sqrt{3}}$$

Defina

$$p_1 = \left\langle x, \frac{1}{\sqrt{3}}\right\rangle \frac{1}{\sqrt{3}} = \left(-1\cdot\frac{1}{\sqrt{3}} + 0\cdot\frac{1}{\sqrt{3}} + 1\cdot\frac{1}{\sqrt{3}}\right)\frac{1}{\sqrt{3}} = 0$$

Então,

$$x - p_1 = x \qquad \text{e} \qquad \|x - p_1\|^2 = \langle x, x\rangle = 2$$

Logo,

$$\mathbf{u}_2 = \frac{1}{\sqrt{2}}x$$

Finalmente,

$$p_2 = \left\langle x^2, \frac{1}{\sqrt{3}}\right\rangle \frac{1}{\sqrt{3}} + \left\langle x^2, \frac{1}{\sqrt{2}}x\right\rangle \frac{1}{\sqrt{2}}x = \frac{2}{3}$$

$$\|x^2 - p_2\|^2 = \langle x^2 - \tfrac{2}{3}, x^2 - \tfrac{2}{3}\rangle = \tfrac{2}{3}$$

e, portanto,

$$\mathbf{u}_3 = \frac{\sqrt{6}}{2}\left(x^2 - \frac{2}{3}\right)$$ □

Polinômios ortogonais serão estudados com mais detalhes na Seção 7.

EXEMPLO 2. Seja

$$A = \begin{pmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ 1 & 4 & 2 \\ 1 & -1 & 0 \end{pmatrix}$$

Encontre uma base ortonormal para o espaço coluna de A.

SOLUÇÃO. As colunas de A são linearmente independentes e, portanto, formam uma base para um espaço tridimensional de R^4. O processo de Gram-Schmidt pode ser usado para se construir uma base ortonormal como se segue. Defina

$$
\begin{aligned}
r_{11} &= \|\mathbf{a}_1\| = 2 \\
\mathbf{q}_1 &= \frac{1}{r_{11}}\mathbf{a}_1 = \left(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}\right)^T \\
r_{12} &= \langle \mathbf{a}_2, \mathbf{q}_1\rangle = \mathbf{q}_1^T\mathbf{a}_2 = 3 \\
\mathbf{p}_1 &= r_{12}\mathbf{q}_1 = 3\mathbf{q}_1 \\
\mathbf{a}_2 - \mathbf{p}_1 &= \left(-\tfrac{5}{2}, \tfrac{5}{2}, \tfrac{5}{2}, -\tfrac{5}{2}\right)^T \\
r_{22} &= \|\mathbf{a}_2 - \mathbf{p}_1\| = 5 \\
\mathbf{q}_2 &= \frac{1}{r_{22}}(\mathbf{a}_2 - \mathbf{p}_1) = \left(-\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}\right)^T \\
r_{13} &= \langle \mathbf{a}_3, \mathbf{q}_1\rangle = \mathbf{q}_1^T\mathbf{a}_3 = 2, \quad r_{23} = \langle \mathbf{a}_3, \mathbf{q}_2\rangle = \mathbf{q}_2^T\mathbf{a}_3 = -2 \\
\mathbf{p}_2 &= r_{13}\mathbf{q}_1 + r_{23}\mathbf{q}_2 = (2, 0, 0, 2)^T \\
\mathbf{a}_3 - \mathbf{p}_2 &= (2, -2, 2, -2)^T \\
r_{33} &= \|\mathbf{a}_3 - \mathbf{p}_2\| = 4 \\
\mathbf{q}_3 &= \frac{1}{r_{33}}(\mathbf{a}_3 - \mathbf{p}_2) = \left(\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}\right)^T
\end{aligned}
$$

Os vetores $\mathbf{q}_1$, $\mathbf{q}_2$, $\mathbf{q}_3$ formam uma base ortonormal para $I(A)$. □

Podemos obter uma fatoração útil para a matriz A se guardarmos todos os produtos internos e normas calculados durante o processo de Gram-Schmidt. Para a matriz do Exemplo 2, se usarmos os r_{ij} para formar uma matriz,

$$
R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{pmatrix} = \begin{pmatrix} 2 & 3 & 2 \\ 0 & 5 & -2 \\ 0 & 0 & 4 \end{pmatrix}
$$

e se definirmos

$$
Q = (\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix}
$$

então é fácil verificar que $QR = A$. Esse resultado é demonstrado no próximo teorema.

Teorema 5.6.2 (Fatoração QR). *Se A é uma matriz $m \times n$ de posto r, então A pode ser fatorada em um produto QR, onde Q é uma matriz $m \times n$ com colunas ortonormais e R é uma matriz $n \times n$ triangular superior e invertível.*

Demonstração. Seja $\mathbf{p}_1, \ldots, \mathbf{p}_{n-1}$ definida como no Teorema 5.6.1 e seja $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ a base ortonormal de $I(A)$ obtida pelo processo de Gram-Schmidt. Defina

$$r_{11} = \|\mathbf{a}_1\|$$

$$r_{kk} = \|\mathbf{a}_k - \mathbf{p}_{k-1}\| \qquad \text{para} \qquad k = 2, \ldots, n$$

e

$$r_{ik} = \mathbf{q}_i^T \mathbf{a}_k \qquad \text{para} \quad i = 1, \ldots, k-1 \qquad \text{e} \qquad k = 2, \ldots, n$$

Pelo processo de Gram-Schmidt,

$$\begin{aligned} r_{11}\mathbf{q}_1 &= \mathbf{a}_1 \\ r_{kk}\mathbf{q}_k &= \mathbf{a}_k - r_{1k}\mathbf{q}_1 - r_{2k}\mathbf{q}_2 - \cdots - r_{k-1,k}\mathbf{q}_{k-1} \quad \text{para} \quad k = 2, \ldots, n \end{aligned} \tag{4}$$

O sistema (4) pode ser colocado na forma

$$\begin{aligned} \mathbf{a}_1 &= r_{11}\mathbf{q}_1 \\ \mathbf{a}_2 &= r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2 \\ &\vdots \\ \mathbf{a}_n &= r_{1n}\mathbf{q}_1 + \cdots + r_{nn}\mathbf{q}_n \end{aligned}$$

Definindo

$$Q = (\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n)$$

e R como a matriz triangular superior,

$$R = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & & & \\ 0 & 0 & \cdots & r_{nn} \end{pmatrix}$$

temos que a j-ésima coluna do produto QR é

$$Q\mathbf{r}_j = r_{1j}\mathbf{q}_1 + r_{2j}\mathbf{q}_2 + \cdots + r_{jj}\mathbf{q}_j = \mathbf{a}_j$$

para $j = 1, \ldots, n$. Portanto,

$$QR = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) = A$$

□

EXEMPLO 3. Calcule a fatoração QR de Gram-Schmidt para a matriz

$$A = \begin{pmatrix} 1 & -2 & -1 \\ 2 & 0 & 1 \\ 2 & -4 & 2 \\ 4 & 0 & 0 \end{pmatrix}$$

SOLUÇÃO

Etapa 1
Defina

$$r_{11} = \|\mathbf{a}_1\| = 5$$

$$\mathbf{q}_1 = \frac{1}{r_{11}}\mathbf{a}_1 = \left(\tfrac{1}{5}, \tfrac{2}{5}, \tfrac{2}{5}, \tfrac{4}{5}\right)^T$$

Etapa 2
Defina

$$
\begin{aligned}
r_{12} &= \mathbf{q}_1^T \mathbf{a}_2 = -2 \\
\mathbf{p}_1 &= r_{12}\mathbf{q}_1 = -2\mathbf{q}_1 \\
\mathbf{a}_2 - \mathbf{p}_1 &= \left(-\tfrac{8}{5}, \tfrac{4}{5}, -\tfrac{16}{5}, \tfrac{8}{5}\right)^T \\
r_{22} &= \|\mathbf{a}_2 - \mathbf{p}_1\| = 4 \\
\mathbf{q}_2 &= \frac{1}{r_{22}}(\mathbf{a}_2 - \mathbf{p}_1) = \left(-\tfrac{2}{5}, \tfrac{1}{5}, -\tfrac{4}{5}, \tfrac{2}{5}\right)^T
\end{aligned}
$$

Etapa 3
Defina

$$
\begin{aligned}
r_{13} &= \mathbf{q}_1^T \mathbf{a}_3 = 1, \qquad r_{23} = \mathbf{q}_2^T \mathbf{a}_3 = -1 \\
\mathbf{p}_2 &= r_{13}\mathbf{q}_1 + r_{23}\mathbf{q}_2 = \mathbf{q}_1 - \mathbf{q}_2 = \left(\tfrac{3}{5}, \tfrac{1}{5}, \tfrac{6}{5}, \tfrac{2}{5}\right)^T \\
\mathbf{a}_3 - \mathbf{p}_2 &= \left(-\tfrac{8}{5}, \tfrac{4}{5}, \tfrac{4}{5}, -\tfrac{2}{5}\right)^T \\
r_{33} &= \|\mathbf{a}_3 - \mathbf{p}_2\| = 2 \\
\mathbf{q}_3 &= \frac{1}{r_{33}}(\mathbf{a}_3 - \mathbf{p}_2) = \left(-\tfrac{4}{5}, \tfrac{2}{5}, \tfrac{2}{5}, -\tfrac{1}{5}\right)^T
\end{aligned}
$$

Em cada etapa, encontramos uma coluna de Q e uma coluna de R. A fatoração é dada por

$$
A = QR = \begin{pmatrix} \frac{1}{5} & -\frac{2}{5} & -\frac{4}{5} \\ \frac{2}{5} & \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & -\frac{4}{5} & \frac{2}{5} \\ \frac{4}{5} & \frac{2}{5} & -\frac{1}{5} \end{pmatrix} \begin{pmatrix} 5 & -2 & 1 \\ 0 & 4 & -1 \\ 0 & 0 & 2 \end{pmatrix}
$$

□

Vimos, na Seção 5, que, se as colunas de uma matriz A $m \times n$ formam um conjunto ortonormal, então a solução de mínimos quadráticos para $A\mathbf{x} = \mathbf{b}$ é, simplesmente, $\hat{\mathbf{x}} = A^T\mathbf{b}$. Se A tem posto n, mas suas colunas não formam um conjunto ortonormal em R^m, então a fatoração QR pode ser usada para se resolver o problema de mínimos quadráticos.

Teorema 5.6.3. *Se A é uma matriz $m \times n$ de posto n, então a solução para o problema de mínimos quadráticos para $A\mathbf{x} = \mathbf{b}$ é dada por $\hat{\mathbf{x}} = R^{-1}Q^T\mathbf{b}$, onde Q e R são as matrizes obtidas pela fatoração dada no Teorema 5.6.2. A solução $\hat{\mathbf{x}}$ pode ser obtida resolvendo-se o sistema $R\mathbf{x} = Q^T\mathbf{b}$ por substituição.*

Demonstração. Seja $\hat{\mathbf{x}}$ a solução do problema de mínimos quadráticos para $A\mathbf{x} = \mathbf{b}$ cuja existência é garantida pelo Teorema 5.4.2. Então $\hat{\mathbf{x}}$ satisfaz as equações normais

$$
A^T A\mathbf{x} = A^T\mathbf{b}
$$

Fatorando A em um produto QR, essas equações ficam

$$
(QR)^T QR\mathbf{x} = (QR)^T\mathbf{b}
$$

ou seja,

$$
R^T(Q^TQ)R\mathbf{x} = R^TQ^T\mathbf{b}
$$

Como Q tem colunas ortonormais, temos que $Q^TQ = I$, logo

$$R^TR\mathbf{x} = R^TQ^T\mathbf{b}$$

Como R^T é invertível, obtemos

$$R\mathbf{x} = Q^T\mathbf{b} \qquad \text{ou} \qquad \mathbf{x} = R^{-1}Q^T\mathbf{b}$$

□

EXEMPLO 4. Encontre a solução de mínimos quadráticos para

$$\begin{pmatrix} 1 & -2 & -1 \\ 2 & 0 & 1 \\ 2 & -4 & 2 \\ 4 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 1 \\ -2 \end{pmatrix}$$

SOLUÇÃO. A matriz de coeficientes desse sistema foi fatorada no Exemplo 3. Usando essa fatoração, obtemos

$$Q^T\mathbf{b} = \begin{pmatrix} \frac{1}{5} & \frac{2}{5} & \frac{2}{5} & \frac{4}{5} \\ -\frac{2}{5} & \frac{1}{5} & -\frac{4}{5} & \frac{2}{5} \\ -\frac{4}{5} & \frac{2}{5} & \frac{2}{5} & -\frac{1}{5} \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$$

O sistema $R\mathbf{x} = Q^T\mathbf{b}$ pode ser resolvido facilmente por substituição:

$$\left(\begin{array}{ccc|c} 5 & -2 & 1 & -1 \\ 0 & 4 & -1 & -1 \\ 0 & 0 & 2 & 2 \end{array}\right)$$

A solução é $\mathbf{x} = (-2/5, 0, 1)^T$. □

O PROCESSO DE GRAM-SCHMIDT MODIFICADO

No Cap. 7, vamos considerar métodos computacionais para resolver problemas de mínimos quadráticos. O método QR do Exemplo 4 não produz, em geral, resultados precisos quando efetuados com aritmética de precisão finita. Na prática, pode ocorrer uma perda de ortogonalidade devido a erros de aproximação no cálculo de $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$. Podemos obter melhor precisão numérica usando uma versão modificada do método de Gram-Schmidt. Nesta versão o vetor $\mathbf{q}_1$ é construído como anteriormente:

$$\mathbf{q}_1 = \frac{1}{\|\mathbf{a}_1\|}\mathbf{a}_1$$

Entretanto, os outros vetores $\mathbf{a}_2, \ldots, \mathbf{a}_n$ são modificados de modo a serem ortogonais a $\mathbf{q}_1$. Isso pode ser feito subtraindo-se de cada vetor $\mathbf{a}_k$ sua projeção ortogonal de $\mathbf{a}_k$ sobre $\mathbf{q}_1$.

$$\mathbf{a}_k^{(1)} = \mathbf{a}_k - (\mathbf{q}_1^T\mathbf{a}_k)\mathbf{q}_1 \qquad k = 2, \ldots, n$$

Em uma segunda etapa, definimos

$$\mathbf{q}_2 = \frac{1}{\|\mathbf{a}_2^{(1)}\|}\mathbf{a}_2^{(1)}$$

O vetor $\mathbf{q}_2$ já é ortogonal a $\mathbf{q}_1$. Modificamos os vetores restantes de modo a torná-los ortogonais a $\mathbf{q}_2$.

$$\mathbf{a}_k^{(2)} = \mathbf{a}_k^{(1)} - (\mathbf{q}_2^T\mathbf{a}_k^{(1)})\mathbf{q}_2 \qquad k = 3, \ldots, n$$

De maneira análoga, determinamos sucessivamente $\mathbf{q}_3, \mathbf{q}_4, \ldots, \mathbf{q}_n$. Na última etapa, precisamos apenas definir

$$\mathbf{q}_n = \frac{1}{\|\mathbf{a}_n^{(n-1)}\|}\mathbf{a}_n^{(n-1)}$$

para obter o conjunto ortonormal $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$. O próximo algoritmo resume esse processo.

Algoritmo 5.6.4 (O Processo de Gram-Schmidt Modificado)

Para $k = 1, 2, \ldots, n$ *defina*

$$r_{kk} = \|\mathbf{a}_k\|$$

$$\mathbf{q}_k = \frac{1}{r_{kk}}\mathbf{a}_k$$

Para $j = k+1, k+2, \ldots, n$, *defina*

$$r_{kj} = \mathbf{q}_k^T\mathbf{a}_j$$

$$\mathbf{a}_j = \mathbf{a}_j - r_{kj}\mathbf{q}_k$$

Fim do ciclo

Fim do ciclo

Se o processo de Gram-Schmidt modificado é aplicado aos vetores colunas de uma matriz A $m \times n$ de posto n, então, como anteriormente, podemos obter uma fatoração QR de A. Essa fatoração pode ser usada computacionalmente para se determinar a solução de mínimos quadráticos para $A\mathbf{x} = \mathbf{b}$.

EXERCÍCIOS

1. Para cada matriz a seguir, use o processo de Gram-Schmidt para encontrar uma base ortonormal para $I(A)$.

(a) $A = \begin{pmatrix} -1 & 3 \\ 1 & 5 \end{pmatrix}$ (b) $A = \begin{pmatrix} 2 & 5 \\ 1 & 10 \end{pmatrix}$

2. Fatore cada uma das matrizes no Exercício 1 em um produto QR, onde Q é uma matriz ortogonal e R é triangular superior.

3. Dada a base $\{(1, 2, -2)^T, (4, 3, 2)^T, (1, 2, 1)^T\}$ para R^3, use o processo de Gram-Schmidt para encontrar uma base ortonormal.

4. Considere o espaço vetorial $C[-1, 1]$ munido do produto interno

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,dx$$

Encontre uma base ortonormal para o subespaço gerado por 1, x e x^2.

5. Sejam

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \quad \text{e} \quad \mathbf{b} = \begin{pmatrix} 12 \\ 6 \\ 18 \end{pmatrix}$$

(a) Use o processo de Gram-Schmidt para encontrar uma base ortonormal para o espaço coluna de A.
(b) Fatore A em um produto QR, onde as colunas de Q formam um conjunto ortonormal de vetores e R é triangular superior.
(c) Resolva o problema de mínimos quadráticos para

$$A\mathbf{x} = \mathbf{b}$$

6. Repita o Exercício 5 para

$$A = \begin{pmatrix} 3 & -1 \\ 4 & 2 \\ 0 & 2 \end{pmatrix} \quad \text{e} \quad \mathbf{b} = \begin{pmatrix} 0 \\ 20 \\ 10 \end{pmatrix}$$

17. Os vetores $\mathbf{x}_1 = (1/2)(1, 1, 1, -1)^T$ e $\mathbf{x}_2 = (1/6)(1, 1, 3, 5)^T$ formam um conjunto ortonormal em R^4. Estenda esse conjunto a uma base ortonormal para R^4, encontrando uma base ortonormal para o núcleo de

$$\begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & 3 & 5 \end{pmatrix}$$

[***Sugestão:*** Encontre primeiro uma base para o núcleo e depois use o processo de Gram- Schmidt.]

8. Use o processo de Gram-Schmidt para encontrar uma base ortonormal para o subespaço de R^4 gerado por $\mathbf{x}_1 = (4, 2, 2, 1)^T$, $\mathbf{x}_2 = (2, 0, 0, 2)^T$ e $\mathbf{x}_3 = (1, 1, -1, 1)^T$.

9. Repita o Exercício 8 usando o processo de Gram-Schmidt modificado e compare as respostas.

10. Mostre que, quando se usa aritmética exata, o processo de Gram-Schmidt modificado produz o mesmo conjunto ortonormal que o processo de Gram-Schmidt clássico.

11. O que vai acontecer se o processo de Gram-Schmidt for aplicado a um conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ com $\mathbf{v}_1$ e $\mathbf{v}_2$ linearmente independentes, mas com $\mathbf{v}_3 \in [\{\mathbf{v}_1, \mathbf{v}_2\}]$? O processo não funciona? Se for esse o caso, como? Explique.

12. Seja A uma matriz $m \times n$ de posto n e seja $\mathbf{b} \in R^m$. Se Q e R são as matrizes obtidas pela aplicação do processo de Gram-Schmidt às colunas de A e se

$$\mathbf{p} = c_1\mathbf{q}_1 + c_2\mathbf{q}_2 + \cdots + c_n\mathbf{q}_n$$

é a projeção ortogonal de $\mathbf{b}$ sobre $I(A)$, mostre que:

(a) $\mathbf{c} = Q^T\mathbf{b}$
(b) $\mathbf{p} = QQ^T\mathbf{b}$
(c) $QQ^T = A(A^TA)^{-1}A^T$

7 POLINÔMIOS ORTOGONAIS

Já vimos como usar polinômios para ajustar curvas a dados e para aproximar funções contínuas. Como ambos são problemas de mínimos quadráticos, eles podem ser simplificados escolhendo-se uma base ortogonal para a classe de polinômios usados para a aproximação. Isso nos leva ao conceito de polinômios ortogonais.

Nesta seção, vamos estudar famílias de polinômios ortogonais associados a diversos produtos internos em $C[a, b]$. Veremos que os polinômios em cada uma dessas classes satisfazem uma relação de recorrência contendo três termos. Essa relação de recorrência é particularmente útil em aplicações computacionais. Certas famílias de polinômios ortogonais têm aplicações importantes em diversas áreas da matemática. Vamos nos referir a esses polinômios como polinômios clássicos e examiná-los com mais detalhe. Em particular, os polinômios clássicos satisfazem certas classes de equações diferenciais lineares de segunda ordem que aparecem na solução de muitas equações diferenciais parciais da física matemática.

SEQÜÊNCIAS ORTOGONAIS

Como a demonstração do Teorema 5.6.1 foi por indução, o processo de Gram-Schmidt é válido para um conjunto enumerável. Logo, se $\mathbf{x}_1, \mathbf{x}_2, \ldots$ é uma seqüência em um espaço vetorial V munido de um produto interno e se $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ são linearmente independentes para cada n, então o processo de Gram-Schmidt pode ser usado para se formar uma seqüência $\mathbf{u}_1, \mathbf{u}_2, \ldots$ onde, $\{\mathbf{u}_1, \mathbf{u}_2, \ldots\}$ é um conjunto ortonormal e $[\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}] = [\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}]$ para cada n. Em particular, a partir da seqüência $1, x, x^2, \ldots$, é possível construir uma "seqüência ortonormal" $p_0(x), p_1(x), \ldots$.

Seja P o espaço vetorial de todos os polinômios e defina o produto interno $<\cdot,\cdot>$ em P por

$$(1) \qquad \langle p, q\rangle = \int_a^b p(x)q(x)w(x)\,dx$$

onde $w(x)$ é uma função contínua positiva. O intervalo pode ser aberto ou fechado, finito ou infinito. Se, no entanto,

$$\int_a^b p(x)w(x)\,dx$$

é imprópria, exigimos que a integral convirja para todo $p \in P$.

Definição. Seja $p_0(x), p_1(x), \ldots$ uma seqüência de polinômios com grau $p_i(x) = i$ para cada i. Se $\langle p_i(x), p_j(x)\rangle = 0$ sempre que $i \neq j$, então $\{p_n(x)\}$ é uma **seqüência de polinômios ortogonais**. Se $\langle p_i, p_j\rangle = \delta_{ij}$, então $\{p_n(x)\}$ é uma **seqüência de polinômios ortonormais**.

Teorema 5.7.1. *Se $p_0, p_1, \ldots$ é uma seqüência de polinômios ortogonais, então*

(i) $p_0, \ldots, p_{n-1}$ *forma uma base para* P_n;
(ii) $p_n \in P_n^{\perp}$ *(isto é, p_n é ortogonal a todos os polinômios de grau menor do que n).*

Demonstração. Pelo Teorema 5.5.1, $p_0, p_1, \ldots, p_{n-1}$ são linearmente independentes em P_n. Como dim $P_n = n$, esses n vetores formam uma base para P_n. Seja $p(x)$ um polinômio arbitrário de grau menor do que n. Então

$$p(x) = \sum_{i=0}^{n-1} c_i p_i(x)$$

e, portanto,

$$\langle p_n, p\rangle = \left\langle p_n, \sum_{i=0}^{n-1} c_i p_i \right\rangle = \sum_{i=0}^{n-1} c_i \langle p_n, p_i\rangle = 0$$

Logo, $p_n \in P_n^{\perp}$. □

Se $\{p_0, p_1, \ldots, p_{n-1}\}$é um conjunto ortogonal em P_n e se

$$u_i = \left(\frac{1}{\|p_i\|}\right) p_i \qquad \text{para} \qquad i = 0, \ldots, n-1$$

então $\{u_0, \ldots, u_{n-1}\}$ é uma base ortonormal para P_n. Logo, se $p \in P_n$, temos

$$\begin{aligned} p &= \sum_{i=0}^{n-1} \langle p, u_i\rangle\, u_i \\ &= \sum_{i=0}^{n-1} \left\langle p, \left(\frac{1}{\|p_i\|}\right) p_i \right\rangle \left(\frac{1}{\|p_i\|}\right) p_i \\ &= \sum_{i=0}^{n-1} \frac{\langle p, p_i\rangle}{\langle p_i, p_i\rangle} p_i \end{aligned}$$

Analogamente, se $f \in C[a, b]$, então a melhor aproximação de mínimos quadráticos de f por elementos em P_n é dada por

$$p = \sum_{i=0}^{n-1} \frac{\langle f, p_i\rangle}{\langle p_i, p_i\rangle} p_i$$

onde $p_0, p_1, \ldots, p_{n-1}$ são polinômios ortogonais.

Uma outra propriedade interessante de polinômios ortogonais é que eles satisfazem uma relação de recorrência com três termos.

Teorema 5.7.2. *Seja $p_0, p_1, \ldots$ uma seqüência de polinômios ortogonais. Seja a_i o coeficiente do termo de maior grau em p_i para cada i e denote o polinômio nulo por $p_{-1}(x)$. Então*

$$\alpha_{n+1}p_{n+1}(x) = (x - \beta_{n+1})p_n(x) - \alpha_n\gamma_n p_{n-1}(x) \qquad (n \geq 0)$$

onde $\alpha_0 = \gamma_0 = 1$ e

$$\alpha_n = \frac{a_{n-1}}{a_n}, \qquad \beta_n = \frac{\langle p_{n-1}, xp_{n-1}\rangle}{\langle p_{n-1}, p_{n-1}\rangle}, \qquad \gamma_n = \frac{\langle p_n, p_n\rangle}{\langle p_{n-1}, p_{n-1}\rangle} \qquad (n \geq 1)$$

Demonstração. Como $p_0, p_1, \ldots, p_{n+1}$ formam uma base para P_{n+2}, temos

$$xp_n(x) = \sum_{k=0}^{n+1} c_{nk}p_k(x) \tag{2}$$

onde

$$c_{nk} = \frac{\langle xp_n, p_k\rangle}{\langle p_k, p_k\rangle} \tag{3}$$

Para qualquer produto interno definido por (1), temos

$$\langle xf, g\rangle = \langle f, xg\rangle$$

Em particular,

$$\langle xp_n, p_k\rangle = \langle p_n, xp_k\rangle$$

Pelo Teorema 5.7.1, se $k < n - 1$, então

$$c_{nk} = \frac{\langle xp_n, p_k\rangle}{\langle p_k, p_k\rangle} = \frac{\langle p_n, xp_k\rangle}{\langle p_k, p_k\rangle} = 0$$

Logo, (2) fica

$$xp_n(x) = c_{n,n-1}p_{n-1}(x) + c_{n,n}p_n(x) + c_{n,n+1}p_{n+1}(x)$$

Essa expressão pode ser colocada na forma

$$c_{n,n+1}p_{n+1}(x) = (x - c_{n,n})p_n(x) - c_{n,n-1}p_{n-1}(x) \tag{4}$$

Comparando os coeficientes dos termos de maior grau em cada lado da Equação (4), vemos que

$$c_{n,n+1}a_{n+1} = a_n$$

ou

$$c_{n,n+1} = \frac{a_n}{a_{n+1}} = \alpha_{n+1} \tag{5}$$

Segue de (4) que

$$c_{n,n+1}\langle p_n, p_{n+1}\rangle = \big\langle p_n, (x - c_{n,n})p_n\big\rangle - c_{n,n-1}\langle p_n, p_{n-1}\rangle$$
$$0 = \langle p_n, xp_n\rangle - c_{nn}\langle p_n, p_n\rangle$$

e, portanto,

$$c_{nn} = \frac{\langle p_n, xp_n\rangle}{\langle p_n, p_n\rangle} = \beta_{n+1}$$

De (3), temos que

$$\begin{aligned}\langle p_{n-1}, p_{n-1}\rangle c_{n,n-1} &= \langle xp_n, p_{n-1}\rangle \\ &= \langle p_n, xp_{n-1}\rangle \\ &= \langle p_n, p_n\rangle c_{n-1,n}\end{aligned}$$

e, portanto, de (5), obtemos

$$c_{n,n-1} = \frac{\langle p_n, p_n \rangle}{\langle p_{n-1}, p_{n-1} \rangle}\alpha_n = \gamma_n \alpha_n \qquad \square$$

Ao gerar uma seqüência de polinômios ortogonais pela relação de recorrência no Teorema 5.7.2, temos a liberdade de escolher o coeficiente a_{n+1} do termo de maior grau que queremos em cada etapa. Isso é razoável, já que qualquer múltiplo não-nulo de um p_{n+1} particular também vai ser ortogonal a $p_0, \ldots, p_n$. Se escolhermos todos os a_i como sendo iguais a 1, por exemplo, a relação de recorrência fica

$$p_{n+1}(x) = (x - \beta_n)p_n(x) - \gamma_n p_{n-1}(x)$$

POLINÔMIOS ORTOGONAIS CLÁSSICOS

Vamos considerar, agora, alguns exemplos. Devido à sua importância, vamos considerar os polinômios clássicos, começando com o exemplo mais simples, os polinômios de Legendre.

Polinômios de Legendre. Os polinômios de Legendre são ortogonais em relação ao produto interno

$$\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\,dx$$

Seja $P_n(x)$ o polinômio de Legendre de grau n. Escolhendo o coeficiente do termo de maior grau, de modo que $P_n(1) = 1$ para todo n, a relação de recorrência para os polinômios de Legendre é

$$(n + 1)P_{n+1}(x) = (2n + 1)xP_n(x) - nP_{n-1}(x)$$

Usando essa fórmula, é fácil gerar a seqüência de polinômios de Legendre. Os cinco primeiros polinômios da seqüência são

$$P_0(x) = 1$$
$$P_1(x) = x$$
$$P_2(x) = \tfrac{1}{2}(3x^2 - 1)$$
$$P_3(x) = \tfrac{1}{2}(5x^3 - 3x)$$
$$P_4(x) = \tfrac{1}{8}(35x^4 - 30x^2 + 3)$$

Polinômios de Tchebycheff. Esses polinômios são ortogonais em relação ao produto interno

$$\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)(1 - x^2)^{-1/2}\,dx$$

É costume normalizar os coeficientes dos termos de maior grau de modo que $a_0 = 1$ e $a_k = 2^{k-1}$ para $k = 1, 2, \ldots$. Os polinômios de Tchebycheff são denotados por $T_n(x)$ e têm a propriedade interessante de que

$$T_n(\cos\theta) = \cos n\theta$$

Essa propriedade, junto com a identidade trigonométrica

$$\cos(n + 1)\theta = 2\cos\theta\cos n\theta - \cos(n - 1)\theta$$

pode ser usada para se obter a relação de recorrência

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

para $n \geq 1$.

TABELA 1

Tchebycheff	Hermite	Laguerre ($\lambda = 0$)
$T_{n+1} = 2x\,T_n - T_{n+1},\ n \geq 1$	$H_{n+1} = 2x\,H_n - 2n\,H_{n-1}$	$(n+1)\,L_{n+1}^{(0)} = (2n+1-x)\,L_0^{(0)} - n\,L_{n-1}^{(0)}$
$T_0 = 1$	$H_0 = 1$	$L_0^{(0)} = 1$
$T_1 = x$	$H_1 = 2x$	$L_1^{(0)} = 1 - x$
$T_2 = 2x^2 - 1$	$H_2 = 4x^2 - 2$	$L_2^{(0)} = \frac{1}{2}(x^2 - 4x + 2)$
$T_3 = 4x^3 - 3x$	$H_3 = 8x^3 = 12x$	$L_3^{(0)} = \frac{1}{6}(-x^3 + 9x^2 - 18x + 6)$

Polinômios de Jacobi. Os polinômios de Legendre e de Tchebycheff são casos especiais de polinômios de Jacobi. Os polinômios de Jacobi são ortogonais em relação ao produto interno

$$\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)(1-x)^{\lambda}(1+x)^{\mu}\,dx$$

onde λ, $\mu > -1$.

Polinômios de Hermite. Os polinômios de Hermite são definidos no intervalo $(-\infty, \infty)$. Eles são ortogonais em relação ao produto interno

$$\langle p, q \rangle = \int_{-\infty}^{\infty} p(x)q(x)e^{-x^2}\,dx$$

A relação de recorrência para os polinômios de Hermite é dada por

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x)$$

Polinômios de Laguerre. Os polinômios de Laguerre são definidos no intervalo $(0, \infty)$ e são ortogonais em relação ao produto interno

$$\langle p, q \rangle = \int_{0}^{\infty} p(x)q(x)x^{\lambda}e^{-x}\,dx$$

onde $\lambda > -1$. A relação de recorrência para os polinômios de Laguerre é dada por

$$(n+1)L_{n+1}^{(\lambda)}(x) = (2n+\lambda+1-x)L_n^{(\lambda)}(x) - (n+\lambda)L_{n-1}^{(\lambda)}(x)$$

A Tabela 1 compara os polinômios de Tchebycheff, Hermite e Laguerre.

APLICAÇÃO: INTEGRAÇÃO NUMÉRICA

Uma aplicação importante de polinômios ortogonais ocorre em integração numérica. Para aproximar

$$\int_a^b f(x)w(x)\,dx \tag{6}$$

aproximamos, primeiro, $f(x)$ por um polinômio interpolador. Podemos determinar um polinômio $P(x)$ que coincide com $f(x)$ em n pontos $x_1, \ldots, x_n$, no intervalo $[a, b]$, usando a *fórmula de interpolação de Lagrange,*

$$P(x) = \sum_{i=1}^{n} f(x_i)\delta_i(x)$$

onde

$$\delta_i(x) = \frac{\prod_{\substack{j=1\\ j\neq i}}^{n}(x - x_j)}{\prod_{\substack{j=1\\ j\neq i}}^{n}(x_i - x_j)}$$

A integral (6) é aproximada, então, por

$$\int_a^b P(x)w(x)\,dx = \sum_{i=1}^{n} A_i f(x_i) \tag{7}$$

onde

$$A_i = \int_a^b \delta_i(x)w(x)\,dx \qquad i = 1, \ldots, n$$

Pode-se mostrar que (7) fornece o valor exato da integral sempre que $f(x)$ for um polinômio de grau menor do que n. Se os pontos $x_1, \ldots, x_n$ forem escolhidos adequadamente, a fórmula (7) vai ser exata para polinômios de grau maior. De fato, pode-se mostrar que, se $p_0, p_1, p_2, \ldots$ é uma seqüência de polinômios ortogonais em relação ao produto interno (1) e se $x_1, \ldots, x_n$ são os zeros de $p_n(x)$, então a fórmula (7) é exata para todos os polinômios de grau menor do que $2n$. O próximo teorema garante que as raízes de p_n são todas reais e pertencentes ao intervalo aberto (a, b).

Teorema 5.7.3. *Se $p_0, p_1, p_2, \ldots$ é uma seqüência de polinômios ortogonais em relação ao produto interno* (1), *então os zeros de $p_n(x)$ são todos reais e distintos e pertencem ao intervalo (a, b).*

Demonstração. Sejam $x_1, \ldots, x_m$ os zeros de $p_n(x)$ que pertencem a (a, b) e nos quais $p_n(x)$ muda de sinal. Então $p_n(x)$ tem que ter um fator $(x - x_i)^{k_i}$, onde k_i é ímpar, para $i = 1, \ldots, m$. Podemos escrever

$$p_n(x) = (x - x_1)^{k_1}(x - x_2)^{k_2} \cdots (x - x_m)^{k_m} q(x)$$

onde $q(x)$ não muda de sinal em (a, b) e $q(x_i) \neq 0$ para $i = 1, \ldots, m$. É claro que $m \leq n$. Vamos mostrar que $m = n$. Seja

$$r(x) = (x - x_1)(x - x_2) \cdots (x - x_m)$$

O produto

$$p_n(x)r(x) = (x - x_1)^{k_1+1}(x - x_2)^{k_2+1} \cdots (x - x_m)^{k_m+1} q(x)$$

envolve apenas potências pares de $(x - x_i)$ para cada i e, portanto, não pode mudar de sinal em (a, b). Logo,

$$\langle p_n, r\rangle = \int_a^b p_n(x)r(x)w(x)\,dx \neq 0$$

Como p_n é ortogonal a todos os polinômios de grau menor do que n, temos que grau$(r(x)) = m \geq n$. □

Fórmulas de integração numérica do tipo (7), onde os x_i são raízes de polinômios ortogonais, são chamadas de *fórmulas de quadratura gaussianas*. A demonstração de que a fórmula é exata para polinômios de grau menor do que $2n$ pode ser encontrada na maioria dos livros-textos de análise numérica para a graduação.

EXERCÍCIOS

1. Use as fórmulas de recorrência para calcular (a) T_4, T_5 e (b) H_4, H_5.

2. Sejam $p_0(x)$, $p_1(x)$ e $p_2(x)$ polinômios ortogonais em relação ao produto interno

$$\langle p(x), q(x)\rangle = \int_{-1}^{1} \frac{p(x)q(x)}{1+x^2}\,dx$$

Use o Teorema 5.7.2 para calcular $p_1(x)$ e $p_2(x)$ se todos os polinômios têm seus coeficientes do termo de maior grau iguais a 1.

3. Mostre que os polinômios de Tchebycheff têm as seguintes propriedades:

(a) $2T_m(x)T_n(x) = T_{m+n}(x) + T_{m-n}(x)$, para $m > n$
(b) $T_m(T_n(x)) = T_{mn}(x)$

4. Encontre a melhor aproximação de mínimos quadráticos de e^x em $[-1, 1]$ em relação ao produto interno

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,dx$$

5. Seja $p_0, p_1, \ldots$ uma seqüência de polinômios ortogonais e denote por a_n o coeficiente do termo de maior grau de p_n. Prove que

$$\|p_n\|^2 = a_n\,\langle x^n, p_n\rangle$$

6. Seja $T_n(x)$ o polinômio de Tchebycheff de grau n e defina

$$U_{n-1}(x) = \frac{1}{n}T_n'(x)$$

para $n = 1, 2, \ldots$.
(a) Calcule $U_0(x)$, $U_1(x)$ e $U_2(x)$.
(b) Se $x = \cos\theta$, mostre que

$$U_{n-1}(x) = \frac{\operatorname{sen} n\theta}{\operatorname{sen}\theta}$$

7. Seja $U_{n-1}(x)$ como no Exercício 6 para $n \geq 1$ e defina $U_{-1}(x) = 0$. Mostre que

(a) $T_n(x) = U_n(x) - xU_{n-1}(x)$, para $n \geq 0$
(b) $U_n(x) = 2xU_{n-1}(x) - U_{n-2}(x)$, para $n \geq 1$

8. Mostre que os U_i definidos no Exercício 6 são ortogonais em relação ao produto interno

$$\langle p, q\rangle = \int_{-1}^{1} p(x)q(x)(1-x^2)^{1/2}\,dx$$

Os U_i são chamados de polinômios de Tchebycheff de segunda espécie.

9. Para $n = 0, 1, 2$, mostre que o polinômio de Legendre $P_n(x)$ satisfaz a equação de segunda ordem

$$(1-x^2)y'' - 2xy' + n(n+1)y = 0$$

10. Prove cada uma das fórmulas a seguir.

(a) $H_n'(x) = 2nH_{n-1}(x), \quad n = 0, 1, \ldots$
(b) $H_n''(x) - 2xH_n'(x) + 2nH_n(x) = 0, \quad n = 0, 1, \ldots$

11. Dada uma função $f(x)$ cujo gráfico contém os pontos $(1, 2)$, $(2, -1)$, $(3, 4)$, use a fórmula de interpolação de Lagrange para construir um polinômio de segundo grau que coincide com f nos pontos dados.

12. Mostre que, se $f(x)$ é um polinômio de grau menor do que n, então $f(x)$ tem que ser igual ao polinômio interpolador $P(x)$ em (7) e, portanto, a soma em (7) fornece o valor exato de $\int_a^b f(x)\,w(x)\,dx$.

13. Use os zeros do polinômio de Legendre $P_2(x)$ para obter uma fórmula de quadratura com dois pontos

$$\int_{-1}^{1} f(x)\,dx \approx A_1 f(x_1) + A_2 f(x_2)$$

14. (a) Qual o grau dos polinômios para os quais a fórmula de quadratura do Exercício 13 é exata?
(b) Use a fórmula do Exercício 13 para aproximar

$$\int_{-1}^{1} (x^3 + 3x^2 + 1)\,dx \qquad \text{e} \qquad \int_{-1}^{1} \frac{1}{1+x^2}\,dx$$

Compare as aproximações com os valores exatos.

EXERCÍCIOS COM O MATLAB PARA O CAPÍTULO 5

1. Defina

$$\mathbf{x} = [\,0:4, 4, -4, 1, 1]' \qquad \text{e} \qquad \mathbf{y} = \texttt{ones}(9, 1)$$

(a) Use a função `norm` do MATLAB para calcular $\|\mathbf{x}\|$, $\|\mathbf{y}\|$, $\|\mathbf{x}+\mathbf{y}\|$ e verifique que vale a desigualdade triangular. Use também o MATLAB para verificar que a lei do paralelogramo

$$\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)$$

é satisfeita.

(b) Se

$$t = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$$

por que sabemos que $|t|$ tem que ser menor ou igual a 1? Use o MATLAB para calcular o valor de t e use a função `acos` para calcular o ângulo entre $\mathbf{x}$ e $\mathbf{y}$. Converta o ângulo para graus multiplicando-o por $180/\pi$. (Observe que o número π é denotado por `pi` no MATLAB.)

(c) Use o MATLAB para calcular o vetor projeção de $\mathbf{x}$ sobre $\mathbf{y}$. Faça $\mathbf{z} = \mathbf{x} - \mathbf{p}$ e verifique que $\mathbf{z}$ é ortogonal a $\mathbf{p}$ calculando o produto interno dos dois vetores. Calcule $\|x\|^2$ e $\|\mathbf{z}\|^2 + \|\mathbf{p}\|^2$ para verificar que o teorema de Pitágoras é satisfeito.

2. (Ajuste de Mínimos Quadrados, por uma Função Linear, para um Conjunto de Dados) A tabela a seguir, com valores de x e y, foi dada na Seção 4 deste capítulo (ver Fig. 5.4.2).

x	−1,0	0,0	2,1	2,3	2,4	5,3	6,0	6,5	8,0
y	−1,02	−0,52	0,55	0,70	0,70	2,13	2,52	2,82	3,54

Os nove pontos estão praticamente em linha reta; logo, os dados podem ser aproximados por uma função linear $z = c_1x + c_2$. Coloque as coordenadas x e y dos pontos em vetores colunas $\mathbf{x}$ e $\mathbf{y}$, respectivamente. Defina $V = [\mathbf{x}, \texttt{ones}(\mathbf{x})]$ e use o operador "\" do MATLAB para calcular os coeficientes c_1 e c_2 como soluções de mínimos quadráticos para o sistema linear 9×2 $V\mathbf{c} = \mathbf{y}$. Para ver o resultado graficamente, defina

$$\mathbf{w} = -1 : 0.1 : 8 \qquad \text{e} \qquad \mathbf{z} = c(1) * \mathbf{w} + c(2) * \texttt{ones}(\mathbf{w})$$

e coloque em um gráfico os pontos dados e o ajuste por mínimos quadráticos usando o comando

$$\texttt{plot}(\mathbf{x}, \mathbf{y}, \text{`}x\text{'}, \mathbf{w}, \mathbf{z})$$

3. (Construção de Curvas de Temperatura por Polinômios de Mínimos Quadráticos) Entre os dados importantes para modelos de previsão do tempo estão conjuntos de dados contendo valores da temperatura em diversos pontos da atmosfera. Esses valores são medidos diretamente, através de balões atmosféricos, ou inferidos por sensores remotos em satélites específicos. A tabela a seguir mostra um conjunto típico de dados obtidos por balões atmosféricos. A temperatura T em graus Kelvin pode ser como uma função da pressão atmosférica p, medida em decibars. A pressão entre 1 e 3 decibars corresponde ao topo da atmosfera e as entre 9 e 10 decibars correspondem à parte mais baixa da atmosfera.

p	1	2	3	4	5	6	7	8	9	10
T	222	227	223	233	244	253	260	266	270	266

(a) Coloque os valores da pressão em um vetor coluna **p** fazendo $\mathbf{p} = [1 : 10]'$ e coloque os valores da temperatura em um vetor coluna **T**. Para encontrar o melhor ajuste de mínimos quadráticos por uma função linear $c_1x + c_2$ para esses dados, escreva um sistema com mais equações do que incógnitas $V\mathbf{c} = \mathbf{T}$. A matriz de coeficientes V pode ser gerada no MATLAB por

$$V = [\mathbf{p}, \texttt{ones}(10, 1)]$$

ou, de modo alternativo, definindo

$$A = \texttt{vander}(\mathbf{p}); \qquad V = A(:, 9 : 10)$$

Observação. Dado um vetor qualquer $\mathbf{x} = (x_1, x_2, \ldots, x_{n+1})^T$, o comando `vander(x)` do MATLAB gera uma matriz de Vandermonde da forma

$$\begin{pmatrix} x_1^n & x_1^{n-1} & \cdots & x_1 & 1 \\ x_2^n & x_2^{n-1} & \cdots & x_2 & 1 \\ \vdots & & & & \\ x_{n+1}^n & x_{n+1}^{n-1} & \cdots & x_{n+1} & 1 \end{pmatrix}$$

Para um ajuste linear, apenas as duas últimas colunas da matriz de Vandermonde são usadas. Maiores informações sobre o comando `vander` podem ser obtidas digitando-se `help vander`. Uma vez construída a matriz V, a solução de mínimos quadráticos **c** pode ser obtida usando-se o operador "\" do MATLAB.

(b) Para ver como a função linear se ajusta aos dados, defina um intervalo para os valores da pressão digitando

$$\mathbf{q} = 1 : 0.1 : 10;$$

Os valores correspondentes da função podem ser determinados digitando-se

$$\mathbf{z} = \texttt{polyval}(\mathbf{c}, \mathbf{q});$$

Podemos fazer o gráfico da função e dos pontos dados usando o comando

$$\texttt{plot}(\mathbf{q}, \mathbf{z}, \mathbf{p}, \mathbf{T}, \text{'}x\text{'})$$

(c) Vamos, agora, tentar obter um melhor ajuste usando uma aproximação por um polinômio de grau 3. De novo, vamos calcular os coeficientes do polinômio cúbico

$$c_1x^3 + c_2x^2 + c_3x + c_4$$

que vai nos dar o melhor ajuste de mínimos quadráticos encontrando a solução de mínimos quadráticos para um sistema com mais equações do que incógnitas $V\mathbf{c} = \mathbf{T}$. A matriz dos coeficientes é determinada pelas quatro últimas colunas da matriz $A = \texttt{vander}(\mathbf{p})$. Para ver o resultado graficamente, faça, novamente,

$$\mathbf{z} = \texttt{polyval}(\mathbf{c}, \mathbf{q})$$

e faça o gráfico do polinômio de grau três e dos pontos dados usando o mesmo comando que anteriormente. Onde você obteve o melhor ajuste, na parte mais alta ou mais baixa da atmosfera?

(d) Para obter um bom ajuste, tanto na parte mais alta quanto na parte mais baixa da atmosfera, tente usar um polinômio de grau seis. Determine os coeficientes como anteriormente usando as sete últimas colunas de A. Faça $\mathbf{z}$ = `polyval`($\mathbf{c}$, $\mathbf{q}$) e faça um gráfico do resultado.

4. (Subespaços Fundamentais: Bases Ortonormais) Os espaços vetoriais $N(A)$, $I(A)$, $N(A^T)$ e $I(A^T)$ são os quatro espaços fundamentais associados a uma matriz A. Podemos usar o MATLAB para construir bases ortonormais para cada um dos espaços fundamentais associados a uma determinada matriz. Podemos, depois, construir as matrizes de projeção correspondentes a cada um desses subespaços.

(a) Defina

$$A = \texttt{rand}(5, 2) * \texttt{rand}(2, 5)$$

Quanto você espera que sejam o posto e a nulidade de A? Explique. Use o MATLAB para verificar suas respostas digitando `rank`(A) e Z = `null`(A). As colunas de Z formam uma base ortonormal para $N(A)$.

(b) A seguir, defina

$$Q = \texttt{orth}(A), \qquad W = \texttt{null}(A'), \qquad S = [Q \quad W]$$

A matriz S deveria ser ortogonal. Por quê? Explique. Calcule $S*S'$ e compare seu resultado com `eye`(5). Em teoria, ambas as matrizes A^TW e W^TA deveriam ter apenas elementos nulos. Por quê? Explique. Use o MATLAB para calcular A^T*W e W^T*A.

(c) Prove que, se Q e W tivessem sido calculadas em aritmética exata, então teríamos

$$I - WW^T = QQ^T \qquad \text{e} \qquad QQ^TA = A$$

(***Sugestão:*** Escreva SS^T em termos de Q e W.)

Use o MATLAB para verificar essas identidades.

(d) Prove que, se Q tivesse sido calculada em aritmética exata, então teríamos $QQ^T\mathbf{b} = \mathbf{b}$ para todo $\mathbf{b} \in I(A)$. Use o MATLAB para verificar isso fazendo $\mathbf{b} = A*$`rand`(5, 1) e depois calculando $Q*Q^T*\mathbf{b}$ e comparando com $\mathbf{b}$.

(e) Como as colunas de Q formam uma base ortonormal para $I(A)$, temos que QQ^T é a matriz de projeção correspondente a $I(A)$. Logo, qualquer que seja $\mathbf{c} \in R^5$, o vetor $\mathbf{q} = QQ^T\mathbf{c}$ é a projeção ortogonal de $\mathbf{c}$ sobre $I(A)$. Faça $\mathbf{c}$ = `rand`(5, 1) e calcule o vetor projeção $\mathbf{q}$. O vetor $\mathbf{r} = \mathbf{c} - \mathbf{q}$ deveria estar em $N(A^T)$. Por quê? Explique. Use o MATLAB para calcular $A'*\mathbf{r}$.

(f) A matriz WW^T é a matriz de projeção correspondente a $N(A^T)$. Use o MATLAB para calcular a projeção ortogonal $\mathbf{w} = WW^T\mathbf{c}$ de $\mathbf{c}$ sobre $N(A^T)$ e compare o resultado com $\mathbf{r}$.

(g) Faça Y = `orth`(A') e use-a para calcular a matriz de projeção U correspondente a $I(A^T)$. Faça $\mathbf{b}$ = `rand`(5, 1) e calcule o vetor projeção $\mathbf{y} = U*\mathbf{b}$ de $\mathbf{b}$ sobre $I(A^T)$. Calcule, também, $U*\mathbf{y}$ e compare com $\mathbf{y}$. O vetor $\mathbf{s} = \mathbf{b} - \mathbf{y}$ deveria estar em $N(A)$. Por quê? Explique. Use o MATLAB para calcular $A*\mathbf{s}$.

(h) Use a matriz Z = `null`(A) para calcular a matriz de projeção V correspondente a $N(A)$. Calcule $V*\mathbf{b}$ e compare com $\mathbf{s}$.

CAPÍTULO 6

AUTOVALORES

Na Seção 1, vamos discutir a equação $A\mathbf{x} = \lambda\mathbf{x}$. Essa equação ocorre em muitas aplicações de álgebra linear. Se a equação tiver uma solução não-trivial $\mathbf{x}$, dizemos que λ é um *autovalor* de A e que $\mathbf{x}$ é um *autovetor* associado a λ. Uma das principais aplicações de autovalores é a solução de sistemas de equações diferenciais lineares. Essa aplicação é apresentada na Seção 2.

Se A é uma matriz $n \times n$, então A representa uma transformação linear de R^n em si mesmo. Autovalores e autovetores nos proporcionam uma chave para entender como o operador funciona. Por exemplo, se $\lambda > 0$, o efeito do operador em qualquer autovetor associado a λ é simplesmente o de esticar ou encolher por um fator constante. De fato, o efeito do operador é determinado facilmente em qualquer combinação linear de autovetores. Em particular, se for possível encontrar uma base de autovetores para R^n, o operador pode ser representado por uma matriz diagonal D em relação a essa base e a matriz A pode ser fatorada em um produto XDX^{-1}. Na Seção 3, vamos ver como fazer isso e discutir uma série de aplicações.

Na Seção 4, vamos considerar matrizes com elementos complexos. Nesse contexto, estaremos interessados em matrizes cujos autovetores formam uma base para C^n (o espaço vetorial de todas as n-uplas de números complexos). A Seção 5 trata de aplicações de autovalores a equações do segundo grau em diversas variáveis e também de aplicações envolvendo máximos e mínimos de funções de várias variáveis. Vamos considerar matrizes simétricas positivas definidas na Seção 6. Os autovalores de tais matrizes são sempre reais e positivos. Essas matrizes aparecem em uma grande variedade de aplicações. Por fim, na Seção 7, vamos estudar matrizes com elementos não-negativos e algumas aplicações à economia.

1 AUTOVALORES E AUTOVETORES

Nesta seção, vamos nos preocupar com o problema de encontrar um escalar λ tal que o sistema $n \times n$

$$A\mathbf{x} = \lambda\mathbf{x}$$

tem uma solução não-trivial.

Definição. Seja A uma matriz $n \times n$. Um escalar λ é um **autovalor** ou **valor característico** de A se existe um vetor não-nulo $\mathbf{x}$ tal que $A\mathbf{x} = \lambda\mathbf{x}$. O vetor $\mathbf{x}$ é um **autovetor** ou **vetor característico** associado a λ.

EXEMPLO 1. Sejam

$$A = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} \quad \text{e} \quad \mathbf{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Como

$$Ax = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix} = 3 \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 3\mathbf{x}$$

temos que $\lambda = 3$ é um autovalor de A e $\mathbf{x} = (2, 1)^T$ é um autovetor associado a λ. De fato, qualquer múltiplo não-nulo de $\mathbf{x}$ é um autovetor, já que

$$A(\alpha\mathbf{x}) = \alpha A\mathbf{x} = \alpha\lambda\mathbf{x} = \lambda(\alpha\mathbf{x})$$

Então, por exemplo, $(4, 2)^T$ também é um autovalor associado a $\lambda = 3$.

$$\begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 12 \\ 6 \end{pmatrix} = 3 \begin{pmatrix} 4 \\ 2 \end{pmatrix}$$

□

A equação $A\mathbf{x} = \lambda\mathbf{x}$ pode ser colocada na forma

(1) $$(A - \lambda I)\mathbf{x} = \mathbf{0}$$

Logo, λ é um autovalor de A se e somente se (1) tem uma solução não-trivial. O conjunto de soluções de (1) é $N(A - \lambda I)$, que é um subespaço de R^n. Logo, se λ é um autovalor de A, então $N(A - \lambda I) \neq \{\mathbf{0}\}$ e qualquer vetor não-nulo em $N(A - \lambda I)$ é um autovetor associado a λ. O subespaço $N(A - \lambda I)$ é chamado de *auto-espaço* associado a λ.

A Equação (1) terá uma solução não-trivial se e somente se $A - \lambda I$ é singular ou, equivalentemente,

(2) $$\det(A - \lambda I) = 0$$

Expandindo o determinante em (2), obtemos um polinômio de grau n na variável λ,

$$p(\lambda) = \det(A - \lambda I)$$

Esse polinômio é chamado de *polinômio característico*, e a Equação (2) é a *equação característica* para a matriz A. As raízes do polinômio característico são os autovalores de A. Se contarmos as raízes de acordo com sua multiplicidade, o polinômio característico tem exatamente n raízes. Então, A tem n autovalores, alguns dos quais podem estar repetidos e alguns podem ser números complexos. Neste último caso, vai ser necessário aumentar nosso corpo de escalares para os números complexos e permitir que nossos vetores e matrizes tenham coeficientes complexos.

Já estabelecemos algumas condições equivalentes para que λ seja um autovalor de A.

> Seja A uma matriz $n \times n$ e seja λ um escalar. As seguintes afirmações são equivalentes:
>
> (a) λ é um autovalor de A;
> (b) $(A - \lambda I)\mathbf{x} = \mathbf{0}$ tem uma solução não-trivial;
> (c) $N(A - \lambda I) \neq \{\mathbf{0}\}$;
> (d) $A - \lambda I$ é singular;
> (e) $\det(A - \lambda I) = 0$.

Vamos agora usar a afirmação (e) para determinar os autovalores em alguns exemplos.

EXEMPLO 2. Encontre os autovalores e os autovetores associados da matriz

$$A = \begin{pmatrix} 3 & 2 \\ 3 & -2 \end{pmatrix}$$

SOLUÇÃO. A equação característica é

$$\begin{vmatrix} 3-\lambda & 2 \\ 3 & -2-\lambda \end{vmatrix} = 0$$

ou

$$\lambda^2 - \lambda - 12 = 0$$

Logo, os autovalores de A são $\lambda_1 = 4$ e $\lambda_2 = -3$. Para encontrar os autovetores associados a $\lambda_1 = 4$, precisamos encontrar o núcleo de $A - 4I$.

$$A - 4I = \begin{pmatrix} -1 & 2 \\ 3 & -6 \end{pmatrix}$$

Resolvendo a equação $(A - 4I)\mathbf{x} = \mathbf{0}$, obtemos

$$\mathbf{x} = (2x_2, x_2)^T$$

Logo, qualquer múltiplo não-nulo de $(2, 1)^T$ é um autovetor associado a λ_1 e $\{(2, 1)^T\}$ é uma base para o auto-espaço associado a λ_1. Analogamente, para encontrar os autovetores associados a λ_2, precisamos resolver $(A + 3I)\mathbf{x} = \mathbf{0}$. Nesse caso $\{(-1, 3)^T\}$ é uma base para $N(A + 3I)$ e qualquer múltiplo não-nulo de $(-1, 3)^T$ é um autovetor associado a λ_2. □

EXEMPLO 3. Seja

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

Encontre os autovalores e os auto-espaços associados.

SOLUÇÃO.

$$\begin{vmatrix} 2-\lambda & -3 & 1 \\ 1 & -2-\lambda & 1 \\ 1 & -3 & 2-\lambda \end{vmatrix} = -\lambda(\lambda - 1)^2$$

Logo, o polinômio característico tem raízes $\lambda_1 = 0$, $\lambda_2 = \lambda_3 = 1$. O auto-espaço associado a $\lambda_1 = 0$ é $N(A)$, que determinamos da maneira usual.

$$\left(\begin{array}{ccc|c} 2 & -3 & 1 & 0 \\ 1 & -2 & 1 & 0 \\ 1 & -3 & 2 & 0 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Fazendo $x_3 = \alpha$, vemos que $x_1 = x_2 = x_3 = \alpha$. Logo, o auto-espaço associado a $\lambda_1 = 0$ consiste em todos os vetores da forma $\alpha(1, 1, 1)^T$. Para encontrar o auto-espaço associado a $\lambda = 1$, precisamos resolver o sistema $(A - I)\mathbf{x} = \mathbf{0}$.

$$\left(\begin{array}{ccc|c} 1 & -3 & 1 & 0 \\ 1 & -3 & 1 & 0 \\ 1 & -3 & 1 & 0 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & -3 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Fazendo $x_3 = \alpha$ e $x_3 = \beta$, obtemos $x_1 = 3\alpha - \beta$. Logo, o auto-espaço associado a $\lambda = 1$ consiste em todos os vetores da forma

$$\begin{pmatrix} 3\alpha - \beta \\ \alpha \\ \beta \end{pmatrix} = \alpha \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

□

EXEMPLO 4. Encontre os autovalores e os auto-espaços associados da matriz

$$A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$$

SOLUÇÃO.

$$\begin{vmatrix} 1-\lambda & 2 \\ -2 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 + 4$$

As raízes do polinômio característico são $\lambda_1 = 1 + 2i$, $\lambda_2 = 1 - 2i$.

$$A - \lambda_1 I = \begin{pmatrix} -2i & 2 \\ -2 & -2i \end{pmatrix} = -2\begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix}$$

Logo, $\{(1, i)^T\}$ é uma base para o auto-espaço associado a $\lambda_1 = 1 + 2i$. Analogamente,

$$A - \lambda_2 I = \begin{pmatrix} 2i & 2 \\ -2 & 2i \end{pmatrix} = 2\begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix}$$

e $\{(1, -i)^T\}$ é uma base para $N(A - \lambda_2 I)$. □

AUTOVALORES COMPLEXOS

Se A é uma matriz $n \times n$ com coeficientes reais, então o polinômio característico tem coeficientes reais e, portanto, todas as suas raízes complexas aparecem em pares conjugados. Então, se $\lambda = a + bi$ ($b \neq 0$) é um autovalor de A, $\bar{\lambda} = a - bi$ é, também, um autovalor de A. O símbolo $\bar{\lambda}$ (que se lê *lambda barra*) é usado para denotar o complexo conjugado de λ. Uma notação análoga pode ser usada para matrizes. Se $A = (a_{ij})$ é uma matriz com coeficientes complexos, então $\bar{A} = (\bar{a}_{ij})$ é a matriz formada tomando-se o complexo conjugado de cada elemento de A. Definimos uma matriz *real* como aquela com a propriedade de que $\bar{A} = A$. Em geral, se A e B são matrizes tais que a multiplicação AB é possível, então $\overline{AB} = \bar{A}\,\bar{B}$ (ver Exercício 17).

Não são apenas os autovalores de A que aparecem em pares conjugados, mas também os autovetores. De fato, se λ é um autovalor complexo de uma matriz real A $n \times n$ e se $\mathbf{z}$ é um autovetor associado a λ, então

$$A\bar{\mathbf{z}} = \bar{A}\bar{\mathbf{z}} = \overline{A\mathbf{z}} = \overline{\lambda\mathbf{z}} = \bar{\lambda}\bar{\mathbf{z}}$$

Logo, $\bar{\mathbf{z}}$ é um autovetor associado a $\bar{\lambda}$. No Exemplo 4, calculamos o autovetor $\mathbf{z} = (1, i)^T$, associado ao autovalor $\lambda = 1 + 2i$ e o autovetor associado a $\bar{\lambda}$, $\bar{\mathbf{z}} = (1, -i)^T$.

O PRODUTO E A SOMA DE AUTOVALORES

É fácil determinar a soma e o produto dos autovalores de uma matriz A $n \times n$. Se $p(\lambda)$ é o polinômio característico de A, então

$$(3) \qquad p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & & a_{2n} \\ \vdots & & & . \\ a_{n1} & a_{n2} & & a_{nn}-\lambda \end{vmatrix}$$

Expandindo em relação à primeira coluna, obtemos

$$\det(A - \lambda I) = (a_{11} - \lambda)\det(M_{11}) + \sum_{i=2}^{n} a_{i1}(-1)^{i+1}\det(M_{i1})$$

onde as matrizes M_{i1}, $i = 2, ..., n$ não contêm os dois elementos diagonais $(a_{11} - \lambda)$ e $(a_{ii} - \lambda)$. Expandindo $\det(M_{11})$ da mesma maneira, concluímos que

$$(a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) \tag{4}$$

é o único termo na expansão de $\det(A - \lambda I)$ envolvendo um produto de mais de $n - 2$ elementos diagonais. Ao expandir (4), o coeficiente de λ^n vai ser $(-1)^n$. Logo, o coeficiente do termo de maior grau de $p(\lambda)$ é $(-1)^n$ e, se $\lambda_1, ..., \lambda_n$ são os autovalores de A, temos

$$\begin{aligned} p(\lambda) &= (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n) \\ &= (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda) \end{aligned} \tag{5}$$

De (3) e (5), obtemos

$$\boxed{\lambda_1 \cdot \lambda_2 \cdots \lambda_n = p(0) = \det(A)}$$

De (4), vemos também que o coeficiente de $(-\lambda)^{n-1}$ é $\sum_{i=1}^{n} a_{ij}$. Se determinarmos o mesmo coeficiente usando (5), obteremos $\sum_{i=1}^{n} \lambda_i$. Portanto,

$$\boxed{\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii}}$$

A soma dos elementos diagonais de A é chamada de *traço* de A e é denotada por $\mathrm{tr}(A)$.

EXEMPLO 5. Se

$$A = \begin{pmatrix} 5 & -18 \\ 1 & -1 \end{pmatrix}$$

então

$$\det(A) = -5 + 18 = 13 \qquad \text{e} \qquad \mathrm{tr}(A) = 5 - 1 = 4$$

O polinômio característico de A é dado por

$$\begin{vmatrix} 5 - \lambda & -18 \\ 1 & -1 - \lambda \end{vmatrix} = \lambda^2 - 4\lambda + 13$$

e, portanto, os autovalores de A são $\lambda_1 = 2 + 3i$ e $\lambda_2 = 2 - 3i$. Observe que

$$\lambda_1 + \lambda_2 = 4 = \mathrm{tr}(A)$$

e

$$\lambda_1\lambda_2 = 13 = \det(A)$$ □

Nos exemplos que vimos até agora n era sempre menor do que 4. Para n maior, é muito mais difícil encontrar as raízes do polinômio caraterístico. No Cap. 7, vamos aprender métodos numéricos para calcular autovalores. (Esses métodos não vão envolver o polinômio caraterístico.) Se os autovalores de A forem calculados através de algum método numérico, é preciso verificar a precisão comparando a soma dos autovalores ao traço de A.

MATRIZES SEMELHANTES

Vamos terminar esta seção com um resultado importante sobre os autovalores de matrizes semelhantes. Lembre-se de que uma matriz B é dita *semelhante* a uma matriz A se existe uma matriz invertível S tal que $B = S^{-1}AS$.

Teorema 6.1.1. *Sejam A e B matrizes $n \times n$. Se B é semelhante a A, então as duas matrizes têm o mesmo polinômio característico e, portanto, os mesmos autovalores.*

Demonstração. Vamos denotar por $p_A(x)$ e $p_B(x)$ os polinômios característicos de A e B, respectivamente. Se B é semelhante a A, existe uma matriz invertível S tal que $B = S^{-1}AS$. Temos, então,

$$\begin{aligned} p_B(\lambda) &= \det(B - \lambda I) \\ &= \det(S^{-1}AS - \lambda I) \\ &= \det(S^{-1}(A - \lambda I)S) \\ &= \det(S^{-1})\det(A - \lambda I)\det(S) \\ &= p_A(\lambda) \end{aligned}$$

Os autovalores da matriz são as raízes do polinômio característico. Como as duas matrizes têm o mesmo polinômio característico, elas têm que ter os mesmos autovalores. □

EXEMPLO 6. Sejam

$$T = \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix} \quad \text{e} \quad S = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}$$

É fácil ver que os autovalores de T são $\lambda_1 = 2$ e $\lambda_2 = 3$. Definindo $A = S^{-1}TS$, os autovalores de A devem ser iguais aos de T.

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

Deixamos a cargo do leitor a verificação de que os autovalores dessa matriz são $\lambda_1 = 2$ e $\lambda_2 = 3$. □

EXERCÍCIOS

1. Encontre os autovalores e os auto-espaços correspondentes de cada uma das matrizes a seguir.

(a) $\begin{pmatrix} 3 & 2 \\ 4 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 6 & -4 \\ 3 & -1 \end{pmatrix}$

(c) $\begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}$ (d) $\begin{pmatrix} 3 & -8 \\ 2 & 3 \end{pmatrix}$

(e) $\begin{pmatrix} 1 & 1 \\ -2 & 3 \end{pmatrix}$ (f) $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

(g) $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ (h) $\begin{pmatrix} 1 & 2 & 1 \\ 0 & 3 & 1 \\ 0 & 5 & -1 \end{pmatrix}$

(i) $\begin{pmatrix} 4 & -5 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$ (j) $\begin{pmatrix} -2 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$

(k) $\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$ (l) $\begin{pmatrix} 3 & 0 & 0 & 0 \\ 4 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$

2. Mostre que os autovalores de uma matriz triangular são os elementos diagonais da matriz.
3. Seja A uma matriz $n \times n$. Mostre que A é singular se e somente se $\lambda = 0$ é um autovalor de A.
4. Seja A uma matriz invertível e seja λ um autovalor de A. Mostre que $1/\lambda$ é um autovalor de A^{-1}.
5. Seja λ um autovalor de A e seja $\mathbf{x}$ uma autovetor associado. Use indução matemática para mostrar que λ^m é um autovalor de A^m e que $\mathbf{x}$ é um autovetor de A^m associado a λ^m para $m = 1, 2, \ldots$.
6. Uma matriz A $n \times n$ é dita idempotente se $A^2 = A$. Mostre que, se λ é um autovalor de uma matriz idempotente, então λ tem que ser igual a 0 ou a 1.
7. Uma matriz A $n \times n$ é dita nilpotente se $A^k = O$ para algum inteiro positivo k. Mostre que todos os autovalores de uma matriz nilpotente são nulos.
8. Seja A uma matriz $n \times n$ e seja $B = A - \alpha I$, onde α é um escalar. Qual a relação entre os autovalores de A e de B? Explique.
9. Mostre que A e A^T têm os mesmos autovalores. Elas têm necessariamente os mesmos autovetores? Explique.
10. Mostre que a matriz

$$A = \begin{pmatrix} \cos\theta & -\text{sen}\theta \\ \text{sen}\theta & \cos\theta \end{pmatrix}$$

vai ter autovalores complexos se θ não for um múltiplo de π. Interprete esse resultado geometricamente.
11. Seja A uma matriz 2×2. Se $\text{tr}(A) = 8$ e $\det(A) = 12$, quais são os autovalores de A?
12. Seja $A = (a_{ij})$ uma matriz $n \times n$ com autovalores $\lambda_1, \ldots, \lambda_n$. Mostre que

$$\lambda_j = a_{jj} + \sum_{i \neq j}(a_{ii} - \lambda_i) \qquad \text{para} \quad j = 1, \ldots, n$$

13. Seja A uma matriz 2×2 e seja $p(\lambda) = \lambda^2 + b\lambda + c$ o polinômio característico de A. Mostre que $b = -\text{tr}(A)$ e que $c = \det(A)$.
14. Seja λ um autovalor não-nulo de A e seja $\mathbf{x}$ um autovetor associado. Mostre que $A^m\mathbf{x}$ também é um autovetor associado a λ para $m = 1, 2, \ldots$.
15. Seja A uma matriz $n \times n$ e seja λ um autovalor de A. Se $A - \lambda I$ tem posto k, qual é a dimensão do auto-espaço associado a λ? Explique.
16. Seja A uma matriz $n \times n$. Mostre que um vetor $\mathbf{x}$ em R^n é um autovetor de A se e somente se o subespaço S de R^n gerado por $\mathbf{x}$ e $A\mathbf{x}$ tem dimensão 1.
17. Sejam $\alpha = a + bi$ e $\beta = c + di$ escalares complexos e sejam A e B matrizes com coeficientes reais.
(a) Mostre que

$$\overline{\alpha + \beta} = \overline{\alpha} + \overline{\beta} \qquad \text{e} \qquad \overline{\alpha\beta} = \overline{\alpha}\,\overline{\beta}$$

(b) Mostre que os elementos (i, j) de $\overline{AB}$ e de $\overline{A}\ \overline{B}$ são iguais e, portanto, que

$$\overline{AB} = \overline{A}\,\overline{B}$$

18. Sejam $\mathbf{x}_1, \ldots, \mathbf{x}_r$ autovetores de uma matriz A $n \times n$ e seja S o subespaço de R^n gerado por $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r$. Mostre que S é invariante sob A (isto é, mostre que $A\mathbf{x} \in S$ sempre que $\mathbf{x} \in S$).
19. Seja $B = S^{-1}AS$ e seja $\mathbf{x}$ um autovetor de B associado a um autovalor λ. Mostre que $S\mathbf{x}$ é um autovetor de A associado a λ.
20. Mostre que, se duas matrizes A e B têm um autovetor em $\mathbf{x}$ em comum (mas não necessariamente um autovalor em comum), então $\mathbf{x}$ é também um autovetor de qualquer matriz da forma $C = \alpha A + \beta B$.
21. Seja A uma matriz $n \times n$ e seja λ um autovalor de A. Mostre que, se $\mathbf{x}$ é um autovetor associado a λ, então $\mathbf{x}$ pertence ao espaço coluna de A.
22. Seja $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ uma base ortonormal para R^n e sejam $\lambda_1, \ldots, \lambda_n$ escalares. Defina

$$A = \lambda_1\mathbf{u}_1\mathbf{u}_1^T + \lambda_2\mathbf{u}_2\mathbf{u}_2^T + \cdots + \lambda_n\mathbf{u}_n\mathbf{u}_n^T$$

Mostre que A é uma matriz simétrica com autovalores $\lambda_1, \lambda_2, \ldots, \lambda_n$ e que $\mathbf{u}_i$ é um autovetor associado a λ_i para cada i.

23. Seja A uma matriz tal que a soma de cada uma de suas colunas é igual a uma constante fixa δ. Mostre que δ é um autovalor de A.

24. Sejam λ_1 e λ_2 autovalores distintos de A. Seja $\mathbf{x}$ um autovetor de A associado a λ_1 e seja $\mathbf{y}$ um autovetor de A^T associado a λ_2. Mostre que $\mathbf{x}$ e $\mathbf{y}$ são ortogonais.

25. Sejam A e B matrizes $n \times n$. Mostre que:

(a) Se λ é um autovalor não-nulo de AB, então também é um autovalor de BA;

(b) Se $\lambda = 0$ é um autovalor de AB, então $\lambda = 0$ também é um autovalor de BA.

26. Prove que não existem matrizes A e B $n \times n$ tais que $AB - BA = I$.

[***Sugestão:*** veja os Exercícios 8 e 25.]

27. Seja $p(\lambda) = (-1)^n(\lambda^n - a_{n-1}\lambda^{n-1} - \dots - a_1\lambda - a_0)$ um polinômio de grau $n \geq 1$ e seja

$$C = \begin{pmatrix} a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & & \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

(a) Mostre que, se λ_i é uma raiz de $p(\lambda) = 0$, então λ_i é um autovalor de C com autovetor associado $\mathbf{x} = (\lambda_i^{n-1}, \lambda_i^{n-2}, \dots, \lambda_i, 1)^T$.

(b) Use o item (a) para mostrar que, se $p(\lambda)$ tem raízes distintas $\lambda_1, \lambda_2, \dots, \lambda_n$, então $p(\lambda)$ é o polinômio característico de C.

A matriz C é chamada de *matriz companheira* de $p(\lambda)$.

28. O resultado dado no Exercício 27(b) é válido mesmo que as raízes de $p(\lambda)$ não sejam distintas. Prove isso da seguinte maneira:

(a) Seja

$$D_m(\lambda) = \begin{pmatrix} a_m & a_{m-1} & \cdots & a_1 & a_0 \\ 1 & -\lambda & \cdots & 0 & 0 \\ \vdots & & & & \\ 0 & 0 & \cdots & 1 & -\lambda \end{pmatrix}$$

e use indução matemática para mostrar que

$$\det(D_m(\lambda)) = (-1)^m(a_m\lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_1\lambda + a_0)$$

(b) Mostre que

$$\det(C - \lambda I) = (a_{n-1} - \lambda)(-\lambda)^{n-1} - \det(D_{n-2}) = p(\lambda)$$

2 SISTEMAS DE EQUAÇÕES DIFERENCIAIS LINEARES

Autovalores têm um papel importante na solução de sistemas de equações diferenciais lineares. Nesta seção, vamos ver como usá-los na solução de sistemas de equações diferenciais lineares com coeficientes constantes. Vamos começar considerando sistemas de equações de primeira ordem da forma

$$\begin{aligned} y_1' &= a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n \\ y_2' &= a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n \\ &\vdots \\ y_n' &= a_{n1}y_1 + a_{n2}y_2 + \cdots + a_{nn}y_n \end{aligned}$$

onde $y_i = f_i(t)$ é uma função em $C^1[a, b]$ para cada i. Fazendo $\mathbf{Y} = (y_1, y_2, ..., y_n)^T$ e $\mathbf{Y}' = (y_1', y_2', ..., y_n')^T$, o sistema pode ser escrito na forma

$$\mathbf{Y}' = A\mathbf{Y}$$

$\mathbf{Y}$ e $\mathbf{Y}'$ são ambas funções vetoriais de t. Vamos considerar o caso mais simples primeiro. Quando $n = 1$, o sistema é simplesmente

$$y' = ay$$

É claro que qualquer função da forma

$$y(t) = ce^{at} \qquad (c \text{ uma constante arbitrária})$$

satisfaz essa equação. Uma generalização natural dessa solução para o caso $n > 1$ é

$$\mathbf{Y} = \begin{pmatrix} x_1 e^{\lambda t} \\ x_2 e^{\lambda t} \\ \vdots \\ x_n e^{\lambda t} \end{pmatrix} = e^{\lambda t}\mathbf{x}$$

onde $\mathbf{x} = (x_1, x_2, ..., x_n)^T$. Para verificar que uma função vetorial dessa forma satisfaz o sistema, calculamos sua derivada:

$$\mathbf{Y}' = \lambda e^{\lambda t}\mathbf{x} = \lambda\mathbf{Y}$$

Então, escolhendo um autovalor de A, λ, e um autovetor associado, $\mathbf{x}$, temos

$$A\mathbf{Y} = e^{\lambda t}A\mathbf{x} = \lambda e^{\lambda t}\mathbf{x} = \lambda\mathbf{Y} = \mathbf{Y}'$$

Logo $\mathbf{Y}$ é uma solução do sistema. Portanto, se λ é um autovalor de A e $\mathbf{x}$ é um autovetor associado, $e^{\lambda t}\mathbf{x}$ é uma solução do sistema $\mathbf{Y}' = A\mathbf{Y}$. Isso é válido independentemente de λ ser real ou complexo. Observe que, se $\mathbf{Y}_1$ e $\mathbf{Y}_2$ são soluções do sistema, então $\alpha\mathbf{Y}_1 + \beta\mathbf{Y}_2$ também é solução, pois

$$\begin{aligned}(\alpha\mathbf{Y}_1 + \beta\mathbf{Y}_2)' &= \alpha\mathbf{Y}_1' + \beta\mathbf{Y}_2' \\ &= \alpha A\mathbf{Y}_1 + \beta A\mathbf{Y}_2 \\ &= A(\alpha\mathbf{Y}_1 + \beta\mathbf{Y}_2)\end{aligned}$$

Por indução temos que, se $\mathbf{Y}_1, ..., \mathbf{Y}_n$ são soluções de $\mathbf{Y}' = A\mathbf{Y}$, então qualquer combinação linear $c_1\mathbf{Y}_1 + ... + c_n\mathbf{Y}_n$ também é solução.

Em geral, as soluções de um sistema $n \times n$ de primeira ordem da forma

$$\mathbf{Y}' = A\mathbf{Y}$$

vai formar um subespaço de dimensão n do espaço vetorial de todas as funções contínuas com valores vetoriais. Se, além disso, impusermos que $\mathbf{Y}(t)$ assume um determinado valor $\mathbf{Y}_0$ quando $t = 0$, o problema vai ter uma única solução. Um problema da forma

$$\mathbf{Y}' = A\mathbf{Y}, \qquad \mathbf{Y}(0) = \mathbf{Y}_0$$

é chamado um *problema de valor inicial.*

EXEMPLO 1. Resolva o sistema

$$\begin{aligned} y_1' &= 3y_1 + 4y_2 \\ y_2' &= 3y_1 + 2y_2 \end{aligned}$$

SOLUÇÃO

$$A = \begin{pmatrix} 3 & 4 \\ 3 & 2 \end{pmatrix}$$

Os autovalores de A são $\lambda_1 = 6$ e $\lambda_2 = -1$. Resolvendo $(A - \lambda I)\mathbf{x} = \mathbf{0}$ com $\lambda = \lambda_1$ e $\lambda = \lambda_2$, vemos que $\mathbf{x}_1 = (4, 3)^T$ é um autovetor associado a λ_1 e que $\mathbf{x}_2 = (1, -1)^T$ é um autovetor associado a λ_2. Logo, qualquer vetor da forma

$$\begin{aligned}\mathbf{Y} &= c_1 e^{\lambda_1 t}\mathbf{x}_1 + c_2 e^{\lambda_2 t}\mathbf{x}_2 \\ &= \begin{pmatrix} 4c_1 e^{6t} + c_2 e^{-t} \\ 3c_1 e^{6t} - c_2 e^{-t} \end{pmatrix}\end{aligned}$$

é uma solução do sistema. □

Suponha que, no Exemplo 1, tivéssemos admitido que $y_1 = 6$ e $y_2 = 1$ quando $t = 0$. Então

$$\mathbf{Y}(0) = \begin{pmatrix} 4c_1 + c_2 \\ 3c_1 - c_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 1 \end{pmatrix}$$

e, portanto, $c_1 = 1$ e $c_2 = 2$. Logo, a solução do problema de valor inicial é dada por

$$\begin{aligned}\mathbf{Y} &= e^{6t}\mathbf{x}_1 + 2e^{-t}\mathbf{x}_2 \\ &= \begin{pmatrix} 4e^{6t} + 2e^{-t} \\ 3e^{6t} - 2e^{-t} \end{pmatrix}\end{aligned}$$

APLICAÇÃO 1: MISTURAS

Dois tanques estão conectados como ilustrado na Fig. 6.2.1. Inicialmente, o tanque A contém 200 litros de água, onde foram dissolvidos 60 gramas de sal, e o tanque B contém 200 litros de água pura. Bombeia-se líquido para dentro e para fora dos dois tanques a taxas mostradas no diagrama. Determine a quantidade de sal em cada tanque no instante t.

SOLUÇÃO. Seja $y_1(t)$ e $y_2(t)$ a quantidade de gramas de sal nos tanques A e B, respectivamente, no instante t. Inicialmente,

$$\mathbf{Y}(0) = \begin{pmatrix} y_1(0) \\ y_2(0) \end{pmatrix} = \begin{pmatrix} 60 \\ 0 \end{pmatrix}$$

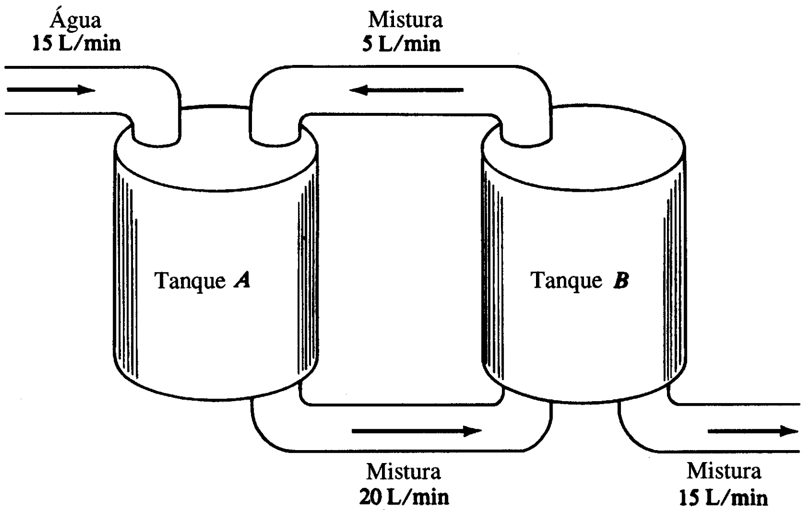


FIG. 6.2.1

A quantidade total de líquido em cada tanque permanece 200 litros, já que a quantidade de líquido bombeada para dentro é igual à quantidade bombeada para fora em cada tanque. A taxa de variação da quantidade de sal em cada tanque é igual à taxa em que está sendo adicionado sal menos a taxa em que está sendo bombeado para fora. Para o tanque A, a taxa em que se está adicionando sal é dada por

$$(5\text{ L/min})\cdot\left(\frac{y_2(t)}{200}\text{g/L}\right)=\frac{y_2(t)}{40}\text{g/min}$$

e a taxa de sal que está sendo bombeado para fora é

$$(20\text{ L/min})\cdot\left(\frac{y_1(t)}{200}\text{g/L}\right)=\frac{y_1(t)}{10}\text{g/min}$$

Portanto, a taxa de variação para o tanque A é dada por

$$y_1'(t)=\frac{y_2(t)}{40}-\frac{y_1(t)}{10}$$

Analogamente, para o tanque B, a taxa de variação é dada por

$$y_2'(t)=\frac{20y_1(t)}{200}-\frac{20y_2(t)}{200}=\frac{y_1(t)}{10}-\frac{y_2(t)}{10}$$

Para encontrar $y_1(t)$ e $y_2(t)$, precisamos resolver o problema de valor inicial

$$\mathbf{Y}'=A\mathbf{Y},\quad \mathbf{Y}(0)=\mathbf{Y}_0$$

onde

$$A=\begin{pmatrix}-\frac{1}{10} & \frac{1}{40}\\ \frac{1}{10} & -\frac{1}{10}\end{pmatrix},\qquad \mathbf{Y}_0=\begin{pmatrix}60\\0\end{pmatrix}$$

Os autovalores de A são $\lambda_1=-3/20$, $\lambda_2=-1/20$ com autovetores associados

$$\mathbf{x}_1=\begin{pmatrix}1\\-2\end{pmatrix}\quad\text{e}\quad \mathbf{x}_2=\begin{pmatrix}1\\2\end{pmatrix}$$

A solução, então, tem que ser da forma

$$\mathbf{Y}=c_1e^{-3t/20}\mathbf{x}_1+c_2e^{-t/20}\mathbf{x}_2$$

Quando $t=0$, $\mathbf{Y}=\mathbf{Y}_0$, logo

$$c_1\mathbf{x}_1+c_2\mathbf{x}_2=\mathbf{Y}_0$$

e podemos encontrar c_1 e c_2 resolvendo

$$\begin{pmatrix}1 & 1\\-2 & 2\end{pmatrix}\begin{pmatrix}c_1\\c_2\end{pmatrix}=\begin{pmatrix}60\\0\end{pmatrix}$$

A solução desse sistema é $c_1=c_2=30$. Portanto, a solução do problema de valor inicial é

$$\mathbf{Y}(t)=\begin{pmatrix}y_1(t)\\y_2(t)\end{pmatrix}=\begin{pmatrix}30e^{-3t/20}+30e^{-t/20}\\-60e^{-3t/20}+60e^{-t/20}\end{pmatrix}$$ □

AUTOVALORES COMPLEXOS

Seja A uma matriz real $n\times n$ com um autovalor complexo $\lambda=a+bi$ e seja $\mathbf{x}$ um autovetor associado. O vetor $\mathbf{x}$ pode ser separado em suas partes real e imaginária.

$$\mathbf{x} = \begin{pmatrix} \operatorname{Re} x_1 + i \operatorname{Im} x_1 \\ \operatorname{Re} x_2 + i \operatorname{Im} x_2 \\ \vdots \\ \operatorname{Re} x_n + i \operatorname{Im} x_n \end{pmatrix} = \begin{pmatrix} \operatorname{Re} x_1 \\ \operatorname{Re} x_2 \\ \vdots \\ \operatorname{Re} x_n \end{pmatrix} + i \begin{pmatrix} \operatorname{Im} x_1 \\ \operatorname{Im} x_2 \\ \vdots \\ \operatorname{Im} x_n \end{pmatrix} = \operatorname{Re} \mathbf{x} + i \operatorname{Im} \mathbf{x}$$

Como os elementos de A são todos reais, temos que $\bar{\lambda} = a - bi$ é também um autovalor de A com autovetor

$$\bar{\mathbf{x}} = \begin{pmatrix} \operatorname{Re} x_1 - i \operatorname{Im} x_1 \\ \operatorname{Re} x_2 - i \operatorname{Im} x_2 \\ \vdots \\ \operatorname{Re} x_n - i \operatorname{Im} x_n \end{pmatrix} = \operatorname{Re} \mathbf{x} - i \operatorname{Im} \mathbf{x}$$

e, portanto, $e^{\lambda t}\mathbf{x}$ e $e^{\bar{\lambda} t}\ \bar{\mathbf{x}}$ são, ambos, soluções do sistema de primeira ordem $\mathbf{Y}' = A\mathbf{Y}$. Qualquer combinação linear dessas duas soluções vai ser, também, uma solução. Logo, se definirmos

$$\mathbf{Y}_1 = \tfrac{1}{2}(e^{\lambda t}\mathbf{x} + e^{\bar{\lambda} t}\bar{\mathbf{x}}) = \operatorname{Re}(e^{\lambda t}\mathbf{x})$$

e

$$\mathbf{Y}_2 = \frac{1}{2i}(e^{\lambda t}\mathbf{x} - e^{\bar{\lambda} t}\bar{\mathbf{x}}) = \operatorname{Im}(e^{\lambda t}\mathbf{x})$$

então as funções vetoriais $\mathbf{Y}_1$ e $\mathbf{Y}_2$ são soluções de $\mathbf{Y}' = A\mathbf{Y}$ com componentes reais. Separando as partes real e imaginária de

$$\begin{aligned} e^{\lambda t}\mathbf{x} &= e^{(a+ib)t}\mathbf{x} \\ &= e^{at}(\cos bt + i \operatorname{sen} bt)(\operatorname{Re} \mathbf{x} + i \operatorname{Im} \mathbf{x}) \end{aligned}$$

vemos que

$$\begin{aligned} \mathbf{Y}_1 &= e^{at}\ [(\cos bt) \operatorname{Re} \mathbf{x} - (\operatorname{sen} bt) \operatorname{Im} \mathbf{x}] \\ \mathbf{Y}_2 &= e^{at}\ [(\cos bt) \operatorname{Im} \mathbf{x} + (\operatorname{sen} bt) \operatorname{Re} \mathbf{x}] \end{aligned}$$

EXEMPLO 2. Resolva o sistema

$$\begin{aligned} y_1' &= y_1 + y_2 \\ y_2' &= -2y_1 + 3y_2 \end{aligned}$$

SOLUÇÃO. Seja

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

Os autovalores de A são $\lambda = 2 + i$ e $\bar{\lambda} = 2 - i$ com autovetores associados $\mathbf{x} = (1, 1 + i)^T$ e $\bar{\mathbf{x}} = (1, 1 - i)^T$, respectivamente.

$$\begin{aligned} e^{\lambda t}\mathbf{x} &= \begin{pmatrix} e^{2t}(\cos t + i \operatorname{sen} t) \\ e^{2t}(\cos t + i \operatorname{sen} t)(1 + i) \end{pmatrix} \\ &= \begin{pmatrix} e^{2t}\cos t + i e^{2t} \operatorname{sen} t \\ e^{2t}(\cos t - \operatorname{sen} t) + i e^{2t}(\cos t + \operatorname{sen} t) \end{pmatrix} \end{aligned}$$

Sejam

$$\mathbf{Y}_1 = \operatorname{Re}(e^{\lambda t}\mathbf{x}) = \begin{pmatrix} e^{2t}\cos t \\ e^{2t}(\cos t - \operatorname{sen} t) \end{pmatrix}$$

e

$$\mathbf{Y}_2 = \operatorname{Im}(e^{\lambda t}\mathbf{x}) = \begin{pmatrix} e^{2t}\operatorname{sen} t \\ e^{2t}(\cos t + \operatorname{sen} t) \end{pmatrix}$$

Qualquer combinação linear

$$\mathbf{Y} = c_1\mathbf{Y}_1 + c_2\mathbf{Y}_2$$

vai ser uma solução do sistema. □

Se a matriz $n \times n$ A de coeficientes do sistema $\mathbf{Y}' = A\mathbf{Y}$ tiver n autovetores linearmente independentes, a solução geral pode ser obtida pelos métodos que foram apresentados. O caso em que A tem menos do que n autovetores linearmente independentes é mais complicado e, por isso, não será apresentado neste livro.

SISTEMAS DE ORDEM MAIOR

Dado um sistema de segunda ordem da forma

$$\mathbf{Y}'' = A_1\mathbf{Y} + A_2\mathbf{Y}'$$

podemos escrevê-lo como um sistema de primeira ordem definindo

$$\begin{aligned} y_{n+1}(t) &= y_1'(t) \\ y_{n+2}(t) &= y_2'(t) \\ &\vdots \\ y_{2n}(t) &= y_n'(t) \end{aligned}$$

Fazendo

$$\mathbf{Y}_1 = \mathbf{Y} = (y_1, y_2, \ldots, y_n)^T$$

e

$$\mathbf{Y}_2 = \mathbf{Y}' = (y_{n+1}, \ldots, y_{2n})^T$$

temos

$$\mathbf{Y}_1' = O\mathbf{Y}_1 + I\mathbf{Y}_2$$

e

$$\mathbf{Y}_2' = A_1\mathbf{Y}_1 + A_2\mathbf{Y}_2$$

Essas equações podem ser combinadas para se obter o sistema de ordem $2n \times 2n$

$$\begin{pmatrix} \mathbf{Y}_1' \\ \mathbf{Y}_2' \end{pmatrix} = \begin{pmatrix} O & I \\ A_1 & A_2 \end{pmatrix}\begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{pmatrix}$$

Se os valores de $\mathbf{Y}_1 = \mathbf{Y}$ e de $\mathbf{Y}_2 = \mathbf{Y}'$ forem especificados quando $t = 0$, o problema de valor inicial vai ter uma única solução.

EXEMPLO 3. Resolva o problema de valor inicial

$$y_1'' = 2y_1 + y_2 + y_1' + y_2'$$
$$y_2'' = -5y_1 + 2y_2 + 5y_1' - y_2'$$
$$y_1(0) = y_2(0) = y_1'(0) = 4, \qquad y_2'(0) = -4$$

SOLUÇÃO. Fazendo $y_3 = y_1'$ e $y_4 = y_2'$, obtemos o sistema de primeira ordem

$$\begin{aligned}
y_1' &= y_3 \\
y_2' &= y_4 \\
y_3' &= 2y_1 + y_2 + y_3 + y_4 \\
y_4' &= -5y_1 + 2y_2 + 5y_3 - y_4
\end{aligned}$$

A matriz de coeficientes para esse sistema,

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

tem autovalores

$$\lambda_1 = 1, \qquad \lambda_2 = -1, \qquad \lambda_3 = 3, \qquad \lambda_4 = -3$$

Os autovetores correspondentes são

$$\mathbf{x}_1 = (1, -1, 1, -1)^T, \qquad \mathbf{x}_2 = (1, 5, -1, -5)^T,$$
$$\mathbf{x}_3 = (1, 1, 3, 3)^T, \qquad \mathbf{x}_4 = (1, -5, -3, 15)^T$$

Logo, a solução é da forma

$$c_1\mathbf{x}_1e^t + c_2\mathbf{x}_2e^{-t} + c_3\mathbf{x}_3e^{3t} + c_4\mathbf{x}_4e^{-3t}$$

Podemos usar a condição inicial para encontrar c_1, c_2, c_3 e c_4. Para $t = 0$, temos

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + c_3\mathbf{x}_3 + c_4\mathbf{x}_4 = (4, 4, 4, -4)^T$$

ou, equivalentemente,

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 5 & 1 & -5 \\ 1 & -1 & 3 & -3 \\ -1 & -5 & 3 & 15 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ 4 \\ -4 \end{pmatrix}$$

A solução desse sistema é $\mathbf{c} = (2, 1, 1, 0)^T$ e, portanto, a solução do problema de valor inicial é

$$\mathbf{Y} = 2\mathbf{x}_1e^t + \mathbf{x}_2e^{-t} + \mathbf{x}_3e^{3t}$$

Logo,

$$\begin{pmatrix} y_1 \\ y_2 \\ y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 2e^t + e^{-t} + e^{3t} \\ -2e^t + 5e^{-t} + e^{3t} \\ 2e^t - e^{-t} + 3e^{3t} \\ -2e^t - 5e^{-t} + 3e^{3t} \end{pmatrix}$$

□

Em geral, se tivermos um sistema de ordem m da forma

$$\mathbf{Y}^{(m)} = A_1\mathbf{Y} + A_2\mathbf{Y}' + \cdots + A_m\mathbf{Y}^{(m-1)}$$

onde cada A_i é uma matriz $n \times n$, podemos transformá-lo em um sistema de primeira ordem definindo

$$\mathbf{Y}_1 = \mathbf{Y}, \mathbf{Y}_2 = \mathbf{Y}_1', \ldots, \mathbf{Y}_m = \mathbf{Y}_{m-1}'$$

Terminamos com um sistema da forma

$$\begin{pmatrix} \mathbf{Y}_1' \\ \mathbf{Y}_2' \\ \vdots \\ \mathbf{Y}_{m-1}' \\ \mathbf{Y}_m' \end{pmatrix} = \begin{pmatrix} O & I & O & \cdots & O \\ O & O & I & \cdots & O \\ \vdots & & & & \\ O & O & O & \cdots & I \\ A_1 & A_2 & A_3 & \cdots & A_m \end{pmatrix} \begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_{m-1} \\ \mathbf{Y}_m \end{pmatrix}$$

Se, além disso, impusermos que $\mathbf{Y}, \mathbf{Y}', \ldots, \mathbf{Y}^{(m-1)}$ assumam valores específicos quando $t = 0$, o problema terá exatamente uma solução.

Se o sistema for, simplesmente, da forma $\mathbf{Y}^{(m)} = A\mathbf{Y}$, não é necessário, em geral, introduzir novas variáveis. Nesse caso precisamos, apenas, calcular as m-ésimas raízes dos autovalores de A. Se λ é um autovalor de A, $\mathbf{x}$ um autovetor associado, σ uma m-ésima raiz de λ e $\mathbf{Y} = e^{\sigma t}\mathbf{x}$, então

$$\mathbf{Y}^{(m)} = \sigma^m e^{\sigma t}\mathbf{x} = \lambda\mathbf{Y}$$

e

$$A\mathbf{Y} = e^{\sigma t}A\mathbf{x} = \lambda e^{\sigma t}\mathbf{x} = \lambda\mathbf{Y}$$

Portanto, $\mathbf{Y} = e^{\sigma t}\mathbf{x}$ é uma solução do sistema.

APLICAÇÃO 2: MOVIMENTO HARMÔNICO

A Fig. 6.2.2 mostra duas massas ligadas por molas com extremidades A e B fixas. As massas estão livres para se mover horizontalmente. Vamos supor que as três molas são uniformes e que, inicialmente, o sistema está na posição de equilíbrio. Exerce-se uma força no sistema para mover as massas. Os deslocamentos horizontais das massas no instante t vão ser denotados por $x_1(t)$ e $x_2(t)$, respectivamente. Vamos supor que não existem forças de atrito, como a fricção. As únicas forças agindo na massa m_1 no instante t são as causadas pelas molas 1 e 2. A força correspondente à mola 1 vai ser igual a $-kx_1$ e a força correspondente à mola 2 vai ser $k(x_2 - x_1)$. Pela segunda lei de Newton,

$$m_1x_1''(t) = -kx_1 + k(x_2 - x_1)$$

Analogamente, as únicas forças agindo na segunda massa são devidas às molas 2 e 3. Usando a segunda lei de Newton novamente, obtemos

$$m_2x_2''(t) = -k(x_2 - x_1) - kx_2$$

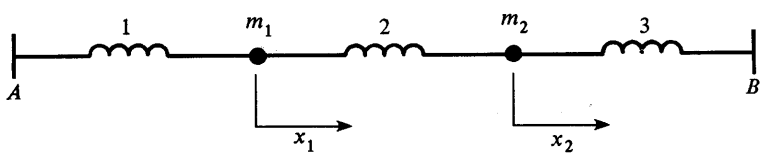


FIG. 6.2.2

Chegamos, então, a um sistema de segunda ordem

$$x_1'' = -\frac{k}{m_1}(2x_1 - x_2)$$
$$x_2'' = -\frac{k}{m_2}(-x_1 + 2x_2)$$

Suponha agora que $m_1 = m_2 = 1$, $k = 1$ e a velocidade inicial de ambas as massas é de $+2$ unidades por segundo. Para determinar os deslocamentos x_1 e x_2 em função de t, escrevemos o sistema na forma

(1) $$\mathbf{X}'' = A\mathbf{X}$$

A matriz de coeficientes

$$A = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$$

tem autovalores $\lambda_1 = -1$ e $\lambda_2 = -3$. Associados a λ_1 temos o autovetor $\mathbf{v}_1 = (1, 1)^T$ e $\sigma_1 = \pm i$. Logo, $e^{it}\mathbf{v}_1$ e $e^{-it}\mathbf{v}_1$ são, ambos, soluções de (1). Temos que

$$\tfrac{1}{2}(e^{it} + e^{-it})\mathbf{v}_1 = (\operatorname{Re} e^{it})\mathbf{v}_1 = (\cos t)\mathbf{v}_1$$

e

$$\frac{1}{2i}(e^{it} - e^{-it})\mathbf{v}_1 = (\operatorname{Im} e^{it})\mathbf{v}_1 = (\operatorname{sen} t)\mathbf{v}_1$$

são soluções de (1). Analogamente, para $\lambda_2 = -3$, temos o autovetor $\mathbf{v}_2 = (1, -1)^T$ e $\sigma_2 = \pm\sqrt{3}i$. Segue que

$$(\operatorname{Re} e^{\sqrt{3}it})\mathbf{v}_2 = (\cos\sqrt{3}t)\mathbf{v}_2$$

e

$$(\operatorname{Im} e^{\sqrt{3}it})\mathbf{v}_2 = (\operatorname{sen}\sqrt{3}t)\mathbf{v}_2$$

também são soluções de (1). Logo, a solução geral vai ser da forma

$$\begin{aligned} \mathbf{X}(t) &= c_1(\cos t)\mathbf{v}_1 + c_2(\operatorname{sen} t)\mathbf{v}_1 + c_3(\cos\sqrt{3}t)\mathbf{v}_2 + c_4(\operatorname{sen}\sqrt{3}t)\mathbf{v}_2 \\ &= \begin{pmatrix} c_1\cos t + c_2\operatorname{sen} t + c_3\cos\sqrt{3}t + c_4\operatorname{sen}\sqrt{3}t \\ c_1\cos t + c_2\operatorname{sen} t - c_3\cos\sqrt{3}t - c_4\operatorname{sen}\sqrt{3}t \end{pmatrix} \end{aligned}$$

No instante $t = 0$, temos

$$x_1(0) = x_2(0) = 0 \quad \text{e} \quad x_1'(0) = x_2'(0) = 2$$

Segue que

$$\begin{aligned} c_1 + c_3 &= 0 \\ c_1 - c_3 &= 0 \end{aligned} \quad \text{e} \quad \begin{aligned} c_2 + \sqrt{3}c_4 &= 2 \\ c_2 - \sqrt{3}c_4 &= 2 \end{aligned}$$

e, portanto,

$$c_1 = c_3 = c_4 = 0 \qquad \text{e} \qquad c_2 = 2$$

Logo, a solução do problema de valor inicial é, simplesmente,

$$\mathbf{X}(t) = \begin{pmatrix} 2\operatorname{sen} t \\ 2\operatorname{sen} t \end{pmatrix}$$

As massas vão oscilar com freqüência 1 e amplitude 2.

EXERCÍCIOS

1. Encontre a solução geral para cada um dos sistemas a seguir.

(a) $\begin{aligned} y_1' &= y_1 + y_2 \\ y_2' &= -2y_1 + 4y_2 \end{aligned}$ (b) $\begin{aligned} y_1' &= 2y_1 + 4y_2 \\ y_2' &= -y_1 - 3y_2 \end{aligned}$

(c) $\begin{aligned} y_1' &= y_1 - 2y_2 \\ y_2' &= -2y_1 + 4y_2 \end{aligned}$ (d) $\begin{aligned} y_1' &= y_1 - y_2 \\ y_2' &= y_1 + y_2 \end{aligned}$

(e) $\begin{aligned} y_1' &= 3y_1 - 2y_2 \\ y_2' &= 2y_1 + 3y_2 \end{aligned}$ (f) $\begin{aligned} y_1' &= y_1 + y_3 \\ y_2' &= 2y_2 + 6y_3 \\ y_3' &= y_2 + 3y_3 \end{aligned}$

2. Resolva cada um dos problemas de valor inicial a seguir.

(a) $\begin{aligned} y_1' &= -y_1 + 2y_2 \\ y_2' &= 2y_1 - y_2 \end{aligned}$ $\quad y_1(0) = 3,\ y_2(0) = 1$

(b) $\begin{aligned} y_1' &= y_1 - 2y_2 \\ y_2' &= 2y_1 + y_2 \end{aligned}$ $\quad y_1(0) = 1,\ y_2(0) = -2$

(c) $\begin{aligned} y_1' &= 2y_1 - 6y_3 \\ y_2' &= y_1 - 3y_3 \\ y_3' &= y_2 - 2y_3 \end{aligned}$ $\quad y_1(0) = y_2(0) = y_3(0) = 2$

(d) $\begin{aligned} y_1' &= y_1 + 2y_3 \\ y_2' &= y_2 - y_3 \\ y_3' &= y_1 + y_2 + y_3 \end{aligned}$ $\quad y_1(0) = y_2(0) = 1,\ y_3(0) = 4$

3. Suponha que

$$\mathbf{Y} = c_1 e^{\lambda_1 t}\mathbf{x}_1 + c_2 e^{\lambda_2 t}\mathbf{x}_2 + \cdots + c_n e^{\lambda_n t}\mathbf{x}_n$$

é a solução do problema de valor inicial

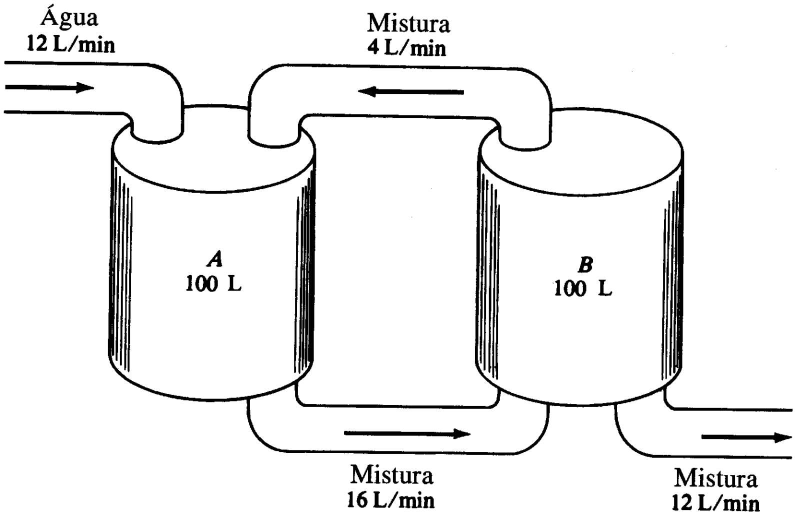
$$\mathbf{Y}' = A\mathbf{Y}, \qquad \mathbf{Y}(0) = \mathbf{Y}_0$$

(a) Mostre que

$$\mathbf{Y}_0 = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n$$

(b) Sejam $X = (\mathbf{x}_1, ..., \mathbf{x}_n)$ e $\mathbf{c} = (c_1, ..., c_n)^T$. Supondo que $\mathbf{x}_1, ..., \mathbf{x}_n$ são linearmente independentes, mostre que $\mathbf{c} = X^{-1}\mathbf{Y}_0$.

4. Dois tanques contêm, cada um, 100 litros de uma mistura. A mistura no tanque A contém 40 gramas de sal, enquanto a mistura no tanque B contém 20 gramas de sal. Bombeia-se líquido para dentro e para fora dos tanques de acordo com a figura a seguir. Determine a quantidade de sal em cada tanque no instante t.



5. Encontre a solução geral de cada sistema a seguir.

(a) $y_1'' = -2y_2$
$y_2'' = y_1 + 3y_2$

(b) $y_1'' = 2y_1 + y_2'$
$y_2'' = 2y_2 + y_1'$

6. Resolva o problema de valor inicial

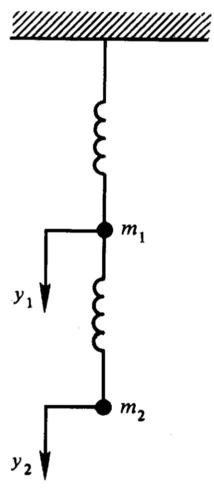
$$y_1'' = -2y_2 + y_1' + 2y_2'$$
$$y_2'' = 2y_1 + 2y_1' - y_2'$$
$$y_1(0) = 1, \quad y_2(0) = 0, \quad y_1'(0) = -3, \quad y_2'(0) = 2$$

7. No problema descrito na segunda aplicação, suponha que as soluções são da forma $x_1 = a_1$ sen σt, $x_2 = a_2$ sen σt. Substitua essas expressões no sistema e resolva para a freqüência σ e as amplitudes a_1 e a_2.

8. Resolva o problema descrito na segunda aplicação com as condições iniciais

$$x_1(0) = x_2(0) = 1, \quad x_1'(0) = 4, \quad \text{e} \quad x_2'(0) = 2$$

9. Duas massas são conectadas por molas, conforme ilustrado no diagrama a seguir. Ambas as molas têm a mesma constante e uma das extremidades da primeira mola está fixa. Se y_1 e y_2 representam os deslocamentos da posição de equilíbrio, encontre um sistema de equações diferenciais de segunda ordem que descreva o movimento do sistema de massas e molas.

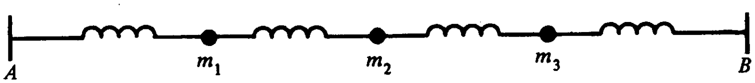


10. Três massas estão conectadas por uma série de molas entre dois pontos fixos conforme mostra a figura a seguir. Suponha que todas as molas têm a mesma constante e represente por $x_1(t)$, $x_2(t)$ e $x_3(t)$ os deslocamentos das massas respectivas no instante t.

(a) Encontre um sistema de equações diferenciais de segunda ordem que descreva o movimento desse sistema de massas e molas.

(b) Resolva o sistema se $m_1 = m_3 = 1/3$, $m_2 = 1/4$, $k = 1$ e

$$x_1(0) = x_2(0) = x_3(0) = 1$$
$$x_1'(0) = x_2'(0) = x_3'(0) = 0$$



11. Transforme a equação de ordem n

$$y^{(n)} = a_0 y + a_1 y' + \cdots + a_{n-1} y^{(n-1)}$$

em um sistema de primeira ordem definindo $y_1 = y$ e $y_j = y_{j-1}'$ para $j = 2, ..., n$. Encontre o polinômio característico da matriz de coeficientes desse sistema.

3 DIAGONALIZAÇÃO

Nesta seção, vamos considerar o problema de fatorar uma matriz A $n \times n$ em um produto da forma XDX^{-1}, onde D é diagonal. Vamos dar uma condição necessária e suficiente para a existência de uma tal fatoração e olhar uma série de exemplos. Começamos mostrando que autovetores associados a autovalores distintos são linearmente independentes.

Teorema 6.3.1. *Se $\lambda_1, \lambda_2, ..., \lambda_k$ são autovalores distintos de uma matriz A $n \times n$ com autovetores associados $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$, então $\mathbf{x}_1, ..., \mathbf{x}_k$ são linearmente independentes.*

Demonstração. Seja r a dimensão do subespaço de R^n gerado por $\mathbf{x}_1, ..., \mathbf{x}_k$ e suponha que $r < k$. Podemos supor (reordenando os $\mathbf{x}_i$ e λ_i se necessário) que $\mathbf{x}_1, ..., \mathbf{x}_r$ são linearmente independentes. Como $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_r, \mathbf{x}_{r+1}$ são linearmente dependentes, existem escalares $c_1, ..., c_r, c_{r+1}$, nem todos nulos, tais que

(1) $$c_1\mathbf{x}_1 + \cdots + c_r\mathbf{x}_r + c_{r+1}\mathbf{x}_{r+1} = \mathbf{0}$$

Observe que c_{r+1} tem que ser diferente de zero, pois, caso contrário, $\mathbf{x}_1, ..., \mathbf{x}_r$ seriam linearmente dependentes. Logo, $c_{r+1}\mathbf{x}_{r+1} \neq \mathbf{0}$ e, portanto, $c_1, ..., c_r$ não podem ser todos nulos. Multiplicando (1) por A, obtemos

$$c_1 A\mathbf{x}_1 + \cdots + c_r A\mathbf{x}_r + c_{r+1} A\mathbf{x}_{r+1} = \mathbf{0}$$

ou

(2) $$c_1\lambda_1\mathbf{x}_1 + \cdots + c_r\lambda_r\mathbf{x}_r + c_{r+1}\lambda_{r+1}\mathbf{x}_{r+1} = \mathbf{0}$$

Subtraindo λ_{r+1} vezes (1) de (2), obtemos

$$c_1(\lambda_1 - \lambda_{r+1})\mathbf{x}_1 + \cdots + c_r(\lambda_r - \lambda_{r+1})\mathbf{x}_r = \mathbf{0}$$

Isso contradiz a independência linear de $\mathbf{x}_1, ..., \mathbf{x}_r$. Logo, r tem que ser igual a k. □

Definição. Uma matriz A $n \times n$ é dita **diagonalizável** se existirem uma matriz invertível X e uma matriz diagonal D satisfazendo

$$X^{-1}AX = D$$

Nesse caso, dizemos que X **diagonaliza** A.

Teorema 6.3.2. *Uma matriz A $n \times n$ é diagonalizável se e somente se A tem n autovetores linearmente independentes.*

Demonstração. Suponha que A tem n autovetores linearmente independentes $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$. Seja λ_i o autovalor associado a $\mathbf{x}_i$ para cada i. (Alguns dos λ_i podem ser iguais.) Seja X a matriz cujo j-ésimo vetor coluna é o vetor $\mathbf{x}_j$ para $j = 1, ..., n$. Então $A\mathbf{x}_j = \lambda_j\mathbf{x}_j$ é o j-ésimo vetor coluna de AX. Logo,

$$\begin{aligned} AX &= (A\mathbf{x}_1, A\mathbf{x}_2, \ldots, A\mathbf{x}_n) \\ &= (\lambda_1\mathbf{x}_1, \lambda_2\mathbf{x}_2, \ldots, \lambda_n\mathbf{x}_n) \\ &= (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \\ &= XD \end{aligned}$$

Como X tem n colunas linearmente independentes, X é invertível e, portanto,

$$D = X^{-1}XD = X^{-1}AX$$

Reciprocamente, suponha que A é diagonalizável. Então, existe uma matriz invertível X tal que $AX = XD$. Se $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$ são os vetores colunas de X, temos

$$A\mathbf{x}_j = \lambda_j\mathbf{x}_j \qquad (\lambda_j = d_{jj})$$

para cada j. Logo, para cada j, λ_j é um autovalor de A com autovetor associado $\mathbf{x}_j$. Como as colunas de X são linearmente independentes, A tem n autovetores linearmente independentes. □

Observações

1. Se A é diagonalizável, então os vetores colunas da matriz X que diagonaliza A são autovetores de A e os elementos diagonais de D são os autovalores associados.
2. A matriz X não é única. Trocando-se a ordem das colunas de uma matriz diagonalizante X, ou multiplicando-as por escalares não-nulos, obteremos outra matriz diagonalizante.
3. Se A é $n \times n$ e tem n autovalores distintos, então A é diagonalizável. Se os autovalores não forem distintos, A pode ser ou não diagonalizável, dependendo se tem ou não n autovetores linearmente independentes.
4. Se A é diagonalizável, então A pode ser fatorada em um produto XDX^{-1}.

Como conseqüência da observação 4, temos que

$$A^2 = (XDX^{-1})(XDX^{-1}) = XD^2X^{-1}$$

e, em geral,

$$\begin{aligned} A^k &= XD^kX^{-1} \\ &= X\begin{pmatrix} (\lambda_1)^k & & & \\ & (\lambda_2)^k & & \\ & & \ddots & \\ & & & (\lambda_n)^k \end{pmatrix}X^{-1} \end{aligned}$$

Uma vez obtida uma fatoração $A = XDX^{-1}$, é fácil calcular as potências de A.

EXEMPLO 1. Seja

$$A = \begin{pmatrix} 2 & -3 \\ 2 & -5 \end{pmatrix}$$

Os autovalores de A são $\lambda_1 = 1$ e $\lambda_2 = -4$. Os autovetores associados a λ_1 e λ_2 são $\mathbf{x}_1 = (3, 1)^T$ e $\mathbf{x}_2 = (1, 2)^T$. Seja

$$X = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$$

Temos

$$\begin{aligned} X^{-1}AX &= \frac{1}{5}\begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 2 & -3 \\ 2 & -5 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & -4 \end{pmatrix} \end{aligned}$$

e

$$XDX^{-1} = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -4 \end{pmatrix}\begin{pmatrix} \frac{2}{5} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{3}{5} \end{pmatrix} = \begin{pmatrix} 2 & -3 \\ 2 & -5 \end{pmatrix} = A$$

□

EXEMPLO 2. Seja

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

É fácil ver que os autovalores de A são $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 1$. Temos o autovetor $(1, 1, 1)^T$ associado a $\lambda_1 = 0$ e os autovetores $(1, 2, 0)^T$ e $(0, -2, 1)^T$ associados a $\lambda = 1$. Seja

$$X = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & -2 \\ 1 & 0 & 1 \end{pmatrix}$$

Então,

$$\begin{aligned} X^{-1}AX &= \begin{pmatrix} -2 & 1 & 2 \\ 3 & -1 & -2 \\ 2 & -1 & -1 \end{pmatrix}\begin{pmatrix} 3 & -1 & -2 \\ 2 & 0 & -2 \\ 2 & -1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & -2 \\ 1 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

Embora $\lambda = 1$ seja um autovalor múltiplo, a matriz ainda pode ser diagonalizada já que tem três autovetores linearmente independentes. Observe, também, que

$$A^k = XD^kX^{-1} = XDX^{-1} = A$$

para todo $k \geq 1$.

□

Se uma matriz A $n \times n$ tem menos que n autovetores linearmente independentes, dizemos que A é *defectiva*. Segue-se do Teorema 6.3.2 que uma matriz defectiva não é diagonalizável.

EXEMPLO 3. Seja

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Ambos os autovalores de A são iguais a 1. Qualquer autovetor associado a $\lambda = 1$ tem que ser um múltiplo de $\mathbf{x}_1 = (1, 0)^T$. Logo A não pode ser diagonalizada. □

EXEMPLO 4. Sejam

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 1 & 0 & 2 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 4 & 0 \\ -3 & 6 & 2 \end{pmatrix}$$

A e B têm os mesmos autovalores

$$\lambda_1 = 4, \qquad \lambda_2 = \lambda_3 = 2$$

O auto-espaço de A correspondente a $\lambda_1 = 4$ é gerado por $\mathbf{e}_2$ e o auto-espaço correspondente a $\lambda = 2$ é gerado por $\mathbf{e}_3$. Como A tem apenas dois autovetores linearmente independentes, não é diagonalizável. Por outro lado, a matriz B tem autovetor $\mathbf{x}_1 = (0, 1, 3)^T$ associado a $\lambda_1 = 4$ e autovetores $\mathbf{x}_2 = (2, 1, 0)^T$ e $\mathbf{e}_3$ associados a $\lambda = 2$. Portanto, B tem três autovetores linearmente independentes e é diagonalizável. Embora $\lambda = 2$ seja um autovalor de multiplicidade 2, a matriz B é diagonalizável já que o auto-espaço correspondente tem dimensão 2. Geometricamente, a matriz B estica dois vetores linearmente independentes por um fator 2. Podemos pensar no autovalor $\lambda = 2$ como tendo *multiplicidade geométrica* 2. Por outro lado, a matriz A apenas estica vetores ao longo do eixo dos z por um fator 2. Nesse caso, o autovalor $\lambda = 2$ tem multiplicidade algébrica 2, mas sua multiplicidade geométrica é 1 (ver Fig. 6.3.1). □

APLICAÇÃO 1

Lembre-se da Aplicação 3 na Seção 3 do Cap. 1. Em uma determinada cidade, 30% das mulheres casadas se divorciam a cada ano e 20% das mulheres solteiras ou divorciadas se casam a cada ano. Existem 8.000 mulheres casadas, 2.000 mulheres solteiras ou divorciadas e a população permanece constante. Encontre o número de mulheres casadas e o número de mulheres solteiras ou divorciadas após 5 anos. Quais as perspectivas a longo prazo se esses percentuais de casamentos e divórcios continuarem indefinidamente no futuro?

SOLUÇÃO. Para encontrar o número de mulheres casadas e o de solteiras ou divorciadas depois de 1 ano, multiplicamos o vetor $\mathbf{x} = (8000, 2000)^T$ por

$$A = \begin{pmatrix} 0{,}7 & 0{,}2 \\ 0{,}3 & 0{,}8 \end{pmatrix}$$

O número de mulheres casadas e o de solteiras ou divorciadas depois de 5 anos serão dados pelas componentes do vetor $A^5\mathbf{x}$. Para calcular $A^5\mathbf{x}$, vamos fatorar A em um produto XDX^{-1}. Os autovalores de A são $\lambda_1 = 1$ e $\lambda_2 = 1/2$, com autovetores associados $\mathbf{x}_1 = (2, 3)^T$ e $\mathbf{x}_2 = (1, -1)^T$, respectivamente.

Seja

$$X = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}$$

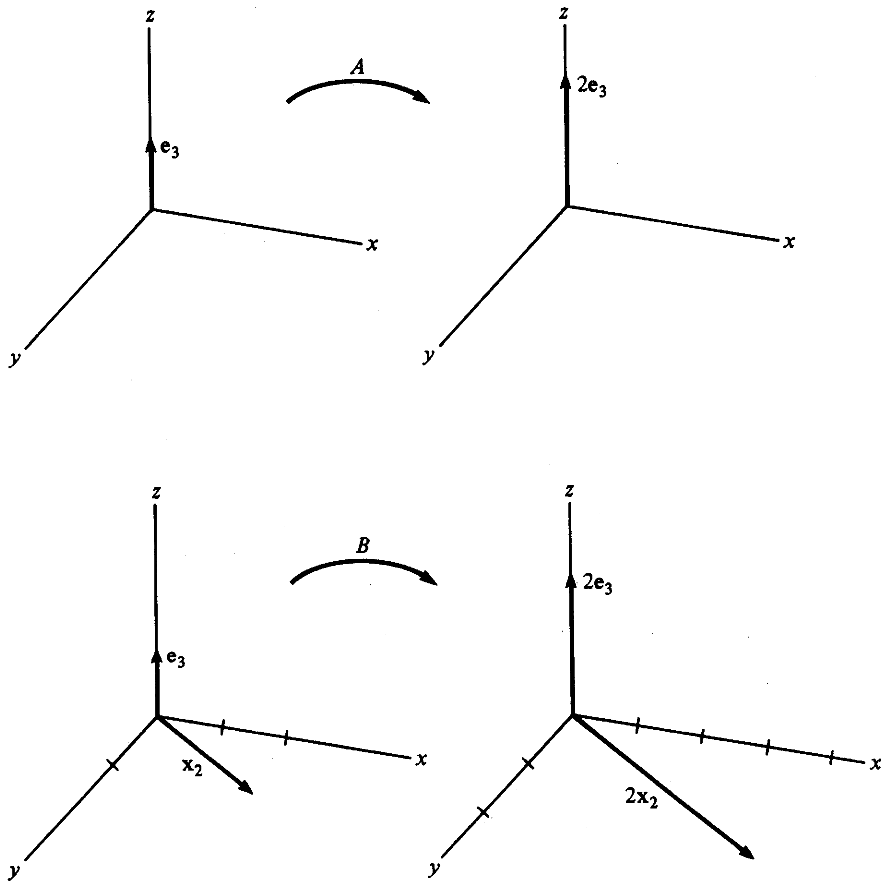


FIG. 6.3.1

Temos que

$$A = XDX^{-1} = \frac{1}{5}\begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 3 & -2 \end{pmatrix}$$

Logo,

$$\begin{aligned} A^5\mathbf{x} &= \frac{1}{5}\begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}^5\begin{pmatrix} 1 & 1 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} 8000 \\ 2000 \end{pmatrix} \\ &= \begin{pmatrix} 4125 \\ 5875 \end{pmatrix} \end{aligned}$$

Após 5 anos, 4125 estarão casadas e 5875 estarão solteiras ou divorciadas. Após n anos, o número de mulheres casadas e o de solteiras ou divorciadas serão dados por $A^n\mathbf{x}$. Para encontrar a tendência a longo prazo, fazemos o limite quando n tende a infinito.

$$\lim_{n\to\infty} A^n\mathbf{x} = \lim_{n\to\infty} XD^nX^{-1}\mathbf{x}$$

Quando $n \to \infty$, $D^n \to \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Portanto,

$$\lim_{n\to\infty} A^n\mathbf{x} = \frac{1}{5}\begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} 8000 \\ 2000 \end{pmatrix} = \begin{pmatrix} 4000 \\ 6000 \end{pmatrix}$$

A longo prazo, 40% das mulheres estarão casadas, enquanto 60% estarão solteiras ou divorciadas. □

APLICAÇÃO 2: GENES LIGADOS AO SEXO

Genes ligados ao sexo são genes localizados no cromossoma X. Por exemplo, o gene para o daltonismo para azul e verde é um gene recessivo relacionado com o sexo. Para encontrar um modelo matemático que descreva o daltonismo em uma determinada população, é necessário dividir a população em duas classes, homens e mulheres. Seja $x_1^{(0)}$ a proporção de genes para o daltonismo na população masculina e seja $x_2^{(0)}$ a proporção na população feminina. (Como o daltonismo é recessivo, a proporção de mulheres daltônicas é, de fato, menor do que $x_2^{(0)}$.) Como os homens recebem um cromossoma X da mãe e nenhum do pai, a proporção $x_1^{(1)}$ de homens daltônicos na próxima geração será a mesma que a proporção de genes recessivos na geração atual de mulheres. Como as mulheres recebem um cromossoma X da mãe e outro do pai, a proporção $x_2^{(1)}$ de genes recessivos na próxima geração de mulheres será a média entre $x_1^{(0)}$ e $x_2^{(0)}$. Então:

$$x_2^{(0)} = x_1^{(1)}$$
$$\tfrac{1}{2}x_1^{(0)} + \tfrac{1}{2}x_2^{(0)} = x_2^{(1)}$$

Se $x_1^{(0)} = x_2^{(0)}$, a proporção não vai mudar na próxima geração. Vamos supor que $x_1^{(0)} \neq x_2^{(0)}$ e escrever o sistema como uma equação matricial.

$$\begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \end{pmatrix} = \begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \end{pmatrix}$$

Vamos denotar por A a matriz dos coeficientes e por $\mathbf{x}^{(n)} = (x_1^{(n)}, x_2^{(n)})^T$ a proporção de genes para o daltonismo nas populações masculina e feminina da $(n + 1)$-ésima geração. Então,

$$\mathbf{x}^{(n)} = A^n\mathbf{x}^{(0)}$$

Para calcular A^n, observamos que A tem autovalores 1 e $-1/2$ e, portanto, pode ser fatorada em um produto

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

Logo,

$$\begin{aligned}\mathbf{x}^{(n)} &= \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}^n\begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{pmatrix}\begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \end{pmatrix} \\ &= \frac{1}{3}\begin{pmatrix} 1-(-\frac{1}{2})^{n-1} & 2+(-\frac{1}{2})^{n-1} \\ 1-(-\frac{1}{2})^{n} & 2+(-\frac{1}{2})^{n} \end{pmatrix}\begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \end{pmatrix}\end{aligned}$$

e, portanto,

$$\lim_{n\to\infty} \mathbf{x}^{(n)} = \frac{1}{3}\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \end{pmatrix}$$

$$= \begin{pmatrix} \dfrac{x_1^{(0)} + 2x_2^{(0)}}{3} \\ \dfrac{x_1^{(0)} + 2x_2^{(0)}}{3} \end{pmatrix}$$

As proporções de genes para o daltonismo nas populações masculina e feminina vão tender a um mesmo valor quando o número de gerações cresce. Se a proporção de homens daltônicos for p e se, durante um certo número de gerações, nenhuma pessoa de fora entrou na população, justifica-se supor que a proporção de daltonismo na população feminina também é p. Como o daltonismo é recessivo, esperaríamos que a proporção de mulheres daltônicas fosse da ordem de p^2. Então, se 1% da população masculina for daltônica, esperamos que em torno de 0,01% da população feminina seja daltônica.

A EXPONENCIAL DE UMA MATRIZ

Dado um escalar a, a exponencial e^a pode ser expressa como uma série de potências

$$e^a = 1 + a + \frac{1}{2!}a^2 + \frac{1}{3!}a^3 + \cdots$$

Analogamente, para uma matriz A $n \times n$, podemos definir sua exponencial e^A através de uma série de potências convergente

$$e^A = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots \tag{3}$$

Exponenciais de matrizes como em (3) aparecem em uma grande variedade de aplicações. Para uma matriz diagonal

$$D = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

é fácil calcular sua exponencial:

$$e^D = \lim_{m\to\infty}\left(I + D + \frac{1}{2!}D^2 + \cdots + \frac{1}{m!}D^m\right)$$

$$= \lim_{m\to\infty}\begin{pmatrix} \sum_{k=1}^{m} \frac{1}{k!}\lambda_1^k & & \\ & \ddots & \\ & & \sum_{k=1}^{m} \frac{1}{k!}\lambda_n^k \end{pmatrix}$$

$$= \begin{pmatrix} e^{\lambda_1} & & & \\ & e^{\lambda_2} & & \\ & & \ddots & \\ & & & e^{\lambda_n} \end{pmatrix}$$

É mais difícil calcular a exponencial de uma matriz geral A $n \times n$. Se, no entanto, A for diagonalizável, então

$$A^k = XD^kX^{-1} \quad \text{para} \quad k = 1, 2, \ldots$$

e, portanto,

$$\begin{aligned} e^A &= X\left(I + D + \frac{1}{2!}D^2 + \frac{1}{3!}D^3 + \cdots\right)X^{-1} \\ &= Xe^DX^{-1} \end{aligned}$$

EXEMPLO 5. Calcule e^A para

$$A = \begin{pmatrix} -2 & -6 \\ 1 & 3 \end{pmatrix}$$

SOLUÇÃO. Os autovalores de A são $\lambda_1 = 1$ e $\lambda_2 = 0$ com autovetores associados $\mathbf{x}_1 = (-2, 1)^T$ e $\mathbf{x}_2 = (-3, 1)^T$. Logo,

$$A = XDX^{-1} = \begin{pmatrix} -2 & -3 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ -1 & -2 \end{pmatrix}$$

e

$$\begin{aligned} e^A = Xe^DX^{-1} &= \begin{pmatrix} -2 & -3 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} e^1 & 0 \\ 0 & e^0 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ -1 & -2 \end{pmatrix} \\ &= \begin{pmatrix} 3-2e & 6-6e \\ e-1 & 3e-2 \end{pmatrix} \end{aligned}$$

□

A exponencial de uma matriz pode ser usada no problema de valor inicial

$$\text{(4)} \qquad \mathbf{Y}' = A\mathbf{Y}, \qquad \mathbf{Y}(0) = \mathbf{Y}_0$$

estudado na Seção 2. No caso de uma equação e uma incógnita,

$$y' = ay, \quad y(0) = y_0$$

a solução é

$$\text{(5)} \qquad y = e^{at}y_0$$

Podemos generalizar isso e expressar a solução de (4) em termos da exponencial de uma matriz, e^{At}, onde $At = tA$ (isto é, t vezes a matriz A). Em geral, uma série de potências pode ser derivada termo a termo dentro de seu raio de convergência. Como a expansão de e^{At} tem raio de convergência infinito, temos

$$\begin{aligned} \frac{d}{dt}e^{At} &= \frac{d}{dt}\left(I + tA + \frac{1}{2!}t^2A^2 + \frac{1}{3!}t^3A^3 + \cdots\right) \\ &= \left(A + tA^2 + \frac{1}{2!}t^2A^3 + \cdots\right) \\ &= A\left(I + tA + \frac{1}{2!}t^2A^2 + \cdots\right) \\ &= Ae^{At} \end{aligned}$$

Se, como em (5), definirmos

$$\mathbf{Y}(t) = e^{At}\mathbf{Y}_0$$

então

$$\mathbf{Y}' = Ae^{At}\mathbf{Y}_0 = A\mathbf{Y}$$

e

$$\mathbf{Y}(0) = \mathbf{Y}_0$$

Logo, a solução de

$$\mathbf{Y}' = A\mathbf{Y}, \qquad \mathbf{Y}(0) = \mathbf{Y}_0$$

é, simplesmente,

(6) $$\mathbf{Y} = e^{At}\mathbf{Y}_0$$

Embora a forma dessa solução pareça diferente das soluções na Seção 2, não há, de fato, nenhuma diferença. Na Seção 2, colocamos a solução na forma

$$c_1e^{\lambda_1 t}\mathbf{x}_1 + c_2e^{\lambda_2 t}\mathbf{x}_2 + \cdots + c_ne^{\lambda_n t}\mathbf{x}_n$$

onde $\mathbf{x}_i$ era um autovetor associado ao autovalor λ_i para $i = 1, ..., n$. Os c_i satisfazendo a condição inicial foram determinados resolvendo-se o sistema

$$X\mathbf{c} = \mathbf{Y}_0$$

com matriz de coeficientes $X = (\mathbf{x}_1, ..., \mathbf{x}_n)$.

Se A é diagonalizável, podemos escrever (6) na forma

$$\mathbf{Y} = Xe^{Dt}X^{-1}\mathbf{Y}_0$$

Logo,

$$\begin{aligned}\mathbf{Y} &= Xe^{Dt}\mathbf{c} \\ &= (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)\begin{pmatrix} c_1e^{\lambda_1 t} \\ c_2e^{\lambda_2 t} \\ \vdots \\ c_ne^{\lambda_n t}\end{pmatrix} \\ &= c_1e^{\lambda_1 t}\mathbf{x}_1 + \cdots + c_ne^{\lambda_n t}\mathbf{x}_n\end{aligned}$$

Resumindo, a solução do problema de valor inicial (4) é dada por

$$\mathbf{Y} = e^{At}\mathbf{Y}_0$$

Se A for diagonalizável, essa solução pode ser escrita na forma

$$\begin{aligned}\mathbf{Y} &= Xe^{Dt}X^{-1}\mathbf{Y}_0 \\ &= c_1e^{\lambda_1 t}\mathbf{x}_1 + c_2e^{\lambda_2 t}\mathbf{x}_2 + \cdots + c_ne^{\lambda_n t}\mathbf{x}_n \quad (\mathbf{c} = X^{-1}\mathbf{Y}_0)\end{aligned}$$

EXEMPLO 6. Use a exponencial de uma matriz para resolver o problema de valor inicial

$$\mathbf{Y}' = A\mathbf{Y}, \qquad \mathbf{Y}(0) = \mathbf{Y}_0$$

onde

$$A = \begin{pmatrix} 3 & 4 \\ 3 & 2 \end{pmatrix}, \qquad \mathbf{Y}_0 = \begin{pmatrix} 6 \\ 1 \end{pmatrix}$$

(Esse problema foi resolvido no Exemplo 1 da Seção 2.)

SOLUÇÃO. Os autovalores de A são $\lambda_1 = 6$ e $\lambda_2 = -1$ com autovetores associados $\mathbf{x}_1 = (4, 3)^T$ e $\mathbf{x}_2 = (1, -1)^T$. Logo,

$$A = XDX^{-1} = \begin{pmatrix} 4 & 1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 6 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{7} & \frac{1}{7} \\ \frac{3}{7} & -\frac{4}{7} \end{pmatrix}$$

e a solução é dada por

$$\begin{aligned} \mathbf{Y} &= e^{At}\mathbf{Y}_0 \\ &= Xe^{Dt}X^{-1}\mathbf{Y}_0 \\ &= \begin{pmatrix} 4 & 1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} e^{6t} & 0 \\ 0 & e^{-t} \end{pmatrix} \begin{pmatrix} \frac{1}{7} & \frac{1}{7} \\ \frac{3}{7} & -\frac{4}{7} \end{pmatrix} \begin{pmatrix} 6 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 4e^{6t} + 2e^{-t} \\ 3e^{6t} - 2e^{-t} \end{pmatrix} \end{aligned}$$

Compare com o Exemplo 1 da Seção 2. □

EXEMPLO 7. Use a exponencial de uma matriz para resolver o problema de valor inicial

$$\mathbf{Y}' = A\mathbf{Y}, \qquad \mathbf{Y}(0) = \mathbf{Y}_0$$

onde

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \mathbf{Y}_0 = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}$$

SOLUÇÃO. Como a matriz A não é diagonalizável, vamos calcular e^{At} pela definição. Observe que $A^3 = O$, de modo que

$$\begin{aligned} e^{At} &= I + tA + \frac{1}{2!}t^2A^2 \\ &= \begin{pmatrix} 1 & t & t^2/2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

A solução do problema de valor inicial é dada por

$$\begin{aligned} \mathbf{Y} &= e^{At}\mathbf{Y}_0 \\ &= \begin{pmatrix} 1 & t & t^2/2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix} \\ &= \begin{pmatrix} 2 + t + 2t^2 \\ 1 + 4t \\ 4 \end{pmatrix} \end{aligned}$$

□

EXERCÍCIOS

1. Fatore cada uma das matrizes A a seguir em um produto XDX^{-1}, onde D é diagonal.

(a) $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ (b) $A = \begin{pmatrix} 5 & 6 \\ -2 & -2 \end{pmatrix}$

(c) $A = \begin{pmatrix} 2 & -8 \\ 1 & -4 \end{pmatrix}$ (d) $A = \begin{pmatrix} 2 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{pmatrix}$

(e) $A = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 3 \\ 1 & 1 & -1 \end{pmatrix}$ (f) $A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ 3 & 6 & -3 \end{pmatrix}$

2. Para cada uma das matrizes no Exercício 1, use a fatoração XDX^{-1} para calcular A^6.

3. Para cada uma das matrizes invertíveis no Exercício 1, use a fatoração XDX^{-1} para calcular A^{-1}.

4. Para cada uma das matrizes a seguir, encontre uma matriz B tal que $B^2 = A$.

(a) $A = \begin{pmatrix} 2 & 1 \\ -2 & -1 \end{pmatrix}$ (b) $A = \begin{pmatrix} 9 & -5 & 3 \\ 0 & 4 & 3 \\ 0 & 0 & 1 \end{pmatrix}$

5. Seja A uma matriz diagonalizável $n \times n$ com matriz diagonalizante X. Mostre que a matriz $Y = (X^{-1})^T$ diagonaliza A^T.

6. Seja A uma matriz diagonalizável cujos autovalores são iguais a 1 ou a -1. Mostre que $A^{-1} = A$.

7. Mostre que qualquer matriz 3×3 da forma

$$\begin{pmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & b \end{pmatrix}$$

não é diagonalizável.

8. Para cada uma das matrizes a seguir, encontre todos os valores possíveis do escalar α que faz com que a matriz não seja diagonalizável ou mostre que não existe tal valor.

(a) $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \alpha \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & \alpha \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 2 & -1 & \alpha \end{pmatrix}$ (d) $\begin{pmatrix} 4 & 6 & -2 \\ -1 & -1 & 1 \\ 0 & 0 & \alpha \end{pmatrix}$

9. Seja A uma matriz 4×4 e seja λ um autovalor de multiplicidade 3. Se $A - \lambda I$ tem posto 1, A é diagonalizável? Explique.

10. Seja A uma matriz $n \times n$ com autovalores reais positivos $\lambda_1 > \lambda_2 > \ldots > \lambda_n$. Seja $\mathbf{x}_i$ o autovetor associado a λ_i para cada i e seja $\mathbf{x} = \alpha_1\mathbf{x}_1 + \ldots + \alpha_n\mathbf{x}_n$.

(a) Mostre que $A^m\mathbf{x} = \sum_{i=1}^{n} \alpha_i \lambda_i^m \mathbf{x}_i$.

(b) Se $\lambda_1 = 1$, mostre que $\lim_{m\to\infty} A^m\mathbf{x} = \alpha_1\mathbf{x}_1$.

11. Seja A uma matriz $n \times n$ com um autovalor λ de multiplicidade n. Mostre que A é diagonalizável se e somente se $A = \lambda I$.

12. Mostre que uma matriz nilpotente não-nula não é diagonalizável.

13. Seja A uma matriz diagonalizável com matriz diagonalizante X. Mostre que os vetores colunas de X associados aos autovalores não-nulos de A formam uma base para $R(A)$.

14. Segue do Exercício 13 que, para uma matriz diagonalizável, o número de autovalores não-nulos (contados de acordo com suas multiplicidades) é igual ao posto da matriz. Dê um exemplo de uma matriz não diagonalizável cujo posto não é igual ao número de autovalores não-nulos.

15. Seja A uma matriz $n \times n$ e seja λ um autovalor de A cujo auto-espaço tem dimensão k, onde $1 < k < n$. Qualquer base $\{\mathbf{x}_1, ..., \mathbf{x}_k\}$ para o auto-espaço pode ser estendida a uma base $\{\mathbf{x}_1, ..., \mathbf{x}_k, \mathbf{x}_{k+1}, ..., \mathbf{x}_n\}$ para R^n. Sejam $X = (\mathbf{x}_1, ..., \mathbf{x}_n)$ e $B = X^{-1}AX$.

(a) Mostre que B é da forma

$$\begin{pmatrix} \lambda I & B_{12} \\ O & B_{22} \end{pmatrix}$$

onde I é a matriz identidade $k \times k$.

(b) Use o Teorema 6.1.1 para mostrar que λ é um autovalor de A com multiplicidade pelo menos k.

16. Sejam $\mathbf{x}$ e $\mathbf{y}$ vetores não-nulos em R^n, $n \geq 2$, e seja $A = \mathbf{x}\mathbf{y}^T$. Mostre que:

(a) Zero é um autovalor de A com $n - 1$ autovetores linearmente independentes e, portanto, tem multiplicidade pelo menos $n - 1$ (ver Exercício 15);

(b) O outro autovalor de A é

$$\lambda_n = \operatorname{tr} A = \mathbf{x}^T\mathbf{y}$$

e $\mathbf{x}$ é um autovetor associado a λ_n;

(c) Se $\lambda_n = \mathbf{x}^T\mathbf{y} \neq 0$, então A é diagonalizável.

17. Seja A uma matriz diagonalizável $n \times n$. Prove que, se B é uma matriz semelhante a A, então B é diagonalizável.

18. Mostre que, se A e B são duas matrizes $n \times n$ com a mesma matriz diagonalizante X, então $AB = BA$.

19. A cidade de Mawtookit mantém uma população constante de 300.000 pessoas. Uma pesquisa em ciência política estimou que existem na cidade 150.000 independentes, 90.000 democratas e 60.000 republicanos.* Estimou-se, também, que, a cada ano, 20% dos independentes tornam-se democratas e 10% tornam-se republicanos. Analogamente, 20% dos democratas tornam-se independentes e 10% tornam-se republicanos, enquanto 10% dos republicanos viram democratas e 10% tornam-se independentes a cada ano. Seja

$$\mathbf{x} = \begin{pmatrix} 150.000 \\ 90.000 \\ 60.000 \end{pmatrix}$$

e seja $\mathbf{x}^{(1)}$ um vetor representando o número de pessoas em cada grupo após 1 ano.

(a) Encontre uma matriz A tal que $A\mathbf{x} = \mathbf{x}^{(1)}$.

(b) Mostre que $\lambda_1 = 1,0$, $\lambda_2 = 0,5$ e $\lambda_3 = 0,7$ são os autovalores de A e fatore A em um produto XDX^{-1}, onde D é diagonal.

(c) Qual grupo deverá dominar a longo prazo? Justifique sua resposta calculando $\lim_{n\to\infty} A^n\mathbf{x}$.

20. Na Aplicação 1, suponha que existem inicialmente p mulheres casadas e $10.000 - p$ mulheres solteiras ou divorciadas, onde $0 \leq p \leq 10.000$. Determine quantas mulheres casadas e quantas solteiras ou divorciadas deverão existir a longo prazo. Sua resposta depende de p? Explique.

21. Use a definição de exponencial de uma matriz para calcular e^A para cada uma das matrizes a seguir.

(a) $A = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$ (b) $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

(c) $A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

*Nos Estados Unidos, existem apenas dois grandes partidos, o Partido Democrata e o Partido Republicano; eleitores de outros partidos, bem pequenos, são chamados de independentes. (*N. T.*)

22. Calcule e^A para cada uma das matrizes a seguir.

(a) $A = \begin{pmatrix} -2 & -1 \\ 6 & 3 \end{pmatrix}$ (b) $A = \begin{pmatrix} 3 & 4 \\ -2 & -3 \end{pmatrix}$

(c) $A = \begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$

23. Resolva o problema de valor inicial $\mathbf{Y}' = A\mathbf{Y}$, $\mathbf{Y}(0) = \mathbf{Y}_0$ em cada um dos itens seguintes calculando $e^{At}\mathbf{Y}_0$.

(a) $A = \begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix}$, $\mathbf{Y}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

(b) $A = \begin{pmatrix} 2 & 3 \\ -1 & -2 \end{pmatrix}$, $\mathbf{Y}_0 = \begin{pmatrix} -4 \\ 2 \end{pmatrix}$

(c) $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{pmatrix}$, $\mathbf{Y}_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

(d) $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ -1 & -1 & -1 \end{pmatrix}$, $\mathbf{Y}_0 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$

24. Seja λ um autovalor de uma matriz A $n \times n$ e seja $\mathbf{x}$ um autovetor associado a λ. Mostre que e^λ é um autovalor de e^A e que $\mathbf{x}$ é um autovetor associado a e^λ.

25. Mostre que e^A é invertível para toda matriz diagonalizável A.

26. Seja A uma matriz diagonalizável com polinômio característico

$$p(\lambda) = a_1\lambda^n + a_2\lambda^{n-1} + \cdots + a_{n+1}$$

(a) Se D é uma matriz diagonal cujos elementos diagonais são os autovalores de A, mostre que

$$p(D) = a_1 D^n + a_2 D^{n-1} + \cdots + a_{n+1} I = O$$

(b) Mostre que $p(A) = O$.

(c) Mostre que, se $a_{n+1} \neq 0$, então A é invertível e $A^{-1} = q(A)$ para algum polinômio q de grau menor do que n.

4 MATRIZES AUTO-ADJUNTAS

Vamos denotar por C^n o espaço vetorial de todas as n-uplas de números complexos. O conjunto C dos números complexos será o nosso corpo de escalares. Já vimos que uma matriz A com todos os elementos reais pode ter autovalores e autovetores complexos. Vamos estudar nesta seção matrizes com elementos complexos e considerar os análogos complexos de matrizes simétricas e ortogonais.

PRODUTOS INTERNOS COMPLEXOS

Se $\alpha = a + bi$ é um escalar complexo, o comprimento de α é dado por

$$|\alpha| = \sqrt{\bar{\alpha}\alpha} = \sqrt{a^2 + b^2}$$

O comprimento de um vetor $\mathbf{z} = (z_1, z_2, ..., z_n)^T$ em C^n é dado por

$$\begin{aligned}\|\mathbf{z}\| &= \left(|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2\right)^{1/2} \\ &= (\overline{z}_1 z_1 + \overline{z}_2 z_2 + \cdots + \overline{z}_n z_n)^{1/2} \\ &= \left(\overline{\mathbf{z}}^T \mathbf{z}\right)^{1/2}\end{aligned}$$

Por conveniência, denotaremos por $\mathbf{z}^H$ a transposta de $\overline{\mathbf{z}}$. Então

$$\overline{\mathbf{z}}^T = \mathbf{z}^H \qquad \text{e} \qquad \|\mathbf{z}\| = (\mathbf{z}^H \mathbf{z})^{1/2}$$

Definição. Seja V um espaço vetorial sobre os números complexos. Um **produto interno** em V é uma operação que associa a cada par de vetores $\mathbf{z}$ e $\mathbf{w}$ em V um número complexo $<\mathbf{z}, \mathbf{w}>$ satisfazendo as seguintes condições:

(i) $<\mathbf{z}, \mathbf{z}> \geq 0$ e a igualdade é válida se e somente se $\mathbf{z} = \mathbf{0}$;
(ii) $<\mathbf{z}, \mathbf{w}> = \overline{\langle \mathbf{w}, \mathbf{z} \rangle}$ quaisquer que sejam $\mathbf{z}$ e $\mathbf{w}$ em V;
(iii) $<\alpha\mathbf{z} + \beta\mathbf{w}, \mathbf{u}> = \alpha<\mathbf{z}, \mathbf{u}> + \beta<\mathbf{w}, \mathbf{u}>$.

Observe que, para um produto interno complexo, temos $<\mathbf{z}, \mathbf{w}> = \overline{\langle \mathbf{w}, \mathbf{z} \rangle}$, em vez de $<\mathbf{w}, \mathbf{z}>$. Se fizermos as modificações adequadas levando em consideração essa propriedade, os teoremas sobre produtos internos reais na Seção 5 do Cap. 5 serão todos válidos para produtos internos complexos. Em particular, vamos lembrar o Teorema 5.5.2: se $\{\mathbf{u}_1, ..., \mathbf{u}_n\}$ é uma base ortonormal para um espaço vetorial real V munido de um produto interno e se

$$\mathbf{x} = \sum_{i=1}^{n} c_i \mathbf{u}_i$$

então

$$c_i = \langle \mathbf{u}_i, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{u}_i \rangle \qquad \text{e} \qquad \|\mathbf{x}\|^2 = \sum_{i=1}^{n} c_i^2$$

No caso de um espaço complexo munido de um produto interno, se $\{\mathbf{w}_1, ..., \mathbf{w}_n\}$ é uma base ortonormal e

$$\mathbf{z} = \sum_{i=1}^{n} c_i \mathbf{w}_i$$

então

$$c_i = \langle \mathbf{z}, \mathbf{w}_i \rangle, \overline{c}_i = \langle \mathbf{w}_i, \mathbf{z} \rangle, \qquad \text{e} \qquad \|\mathbf{z}\|^2 = \sum_{i=1}^{n} c_i \overline{c}_i$$

Podemos definir um produto interno em C^n por

(1) $$\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{w}^H \mathbf{z}$$

para $\mathbf{z}$ e $\mathbf{w}$ em C^n. Deixamos a cargo do leitor verificar que (1) define, de fato, um produto interno em C^n. O espaço vetorial complexo C^n munido desse produto interno é bastante semelhante ao espaço vetorial real R^n munido do produto interno usual. A diferença principal é que, no caso complexo, é preciso conjugar antes de transpor ao efetuar o produto interno.

R^n	C^n
$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^T\mathbf{x}$	$\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{w}^H\mathbf{z}$
$\mathbf{x}^T\mathbf{y} = \mathbf{y}^T\mathbf{x}$	$\mathbf{z}^H\mathbf{w} = \overline{\mathbf{w}^H\mathbf{z}}$
$\|\mathbf{x}\|^2 = \mathbf{x}^T\mathbf{x}$	$\|\mathbf{z}\|^2 = \mathbf{z}^H\mathbf{z}$

EXEMPLO 1. Se

$$\mathbf{z} = \begin{pmatrix} 5+i \\ 1-3i \end{pmatrix} \quad \text{e} \quad \mathbf{w} = \begin{pmatrix} 2+i \\ -2+3i \end{pmatrix}$$

então

$$\mathbf{w}^H\mathbf{z} = (2-i,\ -2-3i)\begin{pmatrix} 5+i \\ 1-3i \end{pmatrix} = (11-3i) + (-11+3i) = 0$$

$$\mathbf{z}^H\mathbf{z} = |5+i|^2 + |1-3i|^2 = 36$$

$$\mathbf{w}^H\mathbf{w} = |2+i|^2 + |-2+3i|^2 = 18$$

Segue que $\mathbf{z}$ e $\mathbf{w}$ são ortogonais e

$$\|\mathbf{z}\| = 6, \qquad \|\mathbf{w}\| = 3\sqrt{2}$$

□

MATRIZES AUTO-ADJUNTAS

Seja $M = (m_{ij})$ uma matriz $n \times n$ com $m_{ij} = a_{ij} + ib_{ij}$ para cada i e j. Podemos escrever M na forma

$$M = A + iB$$

onde $A = (a_{ij})$ e $B = (b_{ij})$ tem elementos reais. Definimos a matriz conjugada de M por

$$\overline{M} = A - iB$$

Logo, $\overline{M}$ é formada conjugando-se todos os elementos de M. A transposta de $\overline{M}$ será denotada por M^H. O espaço vetorial de todas as matrizes $m \times n$ com elementos complexos será denotado por $C^{m\times n}$. Se A e C são elementos de $C^{m\times n}$ e se $B \in C^{n\times r}$, então as seguintes regras são facilmente verificáveis (ver Exercício 7):

I. $(A^H)^H = A$
II. $(\alpha A + \beta B)^H = \overline{\alpha}A^H + \overline{\beta}B^H$
III. $(AC)^H = C^H A^H$

Definição. Uma matriz M é dita **auto-adjunta** se $M = M^H$.

EXEMPLO 2. A matriz

$$M = \begin{pmatrix} 3 & 2-i \\ 2+i & 4 \end{pmatrix}$$

é auto-adjunta, já que

$$M^H = \begin{pmatrix} \overline{3} & \overline{2-i} \\ \overline{2+i} & \overline{4} \end{pmatrix}^T = \begin{pmatrix} 3 & 2-i \\ 2+i & 4 \end{pmatrix} = M$$

□

Se M é uma matriz com elementos reais, então $M^H = M^T$. Em particular, se M é uma matriz real simétrica, então M é auto-adjunta. Podemos, portanto, considerar matrizes auto-adjuntas como o análogo complexo de matrizes simétricas. Matrizes auto-adjuntas têm muitas propriedades interessantes, como veremos no próximo teorema.

Teorema 6.4.1. *Todos os autovalores de uma matriz auto-adjunta são reais. Além disso, autovetores associados a autovalores distintos são ortogonais.*

Demonstração. Seja A uma matriz auto-adjunta. Seja λ um autovalor de A e seja $\mathbf{x}$ um autovetor associado. Se $\alpha = \mathbf{x}^H A\mathbf{x}$, então

$$\overline{\alpha} = \alpha^H = (\mathbf{x}^H A\mathbf{x})^H = \mathbf{x}^H A\mathbf{x} = \alpha$$

Logo, α é real. Temos que

$$\alpha = \mathbf{x}^H A\mathbf{x} = \mathbf{x}^H \lambda\mathbf{x} = \lambda\|\mathbf{x}\|^2$$

e, portanto,

$$\lambda = \frac{\alpha}{\|\mathbf{x}\|^2}$$

é real. Se $\mathbf{x}_1$ e $\mathbf{x}_2$ são autovetores associados a autovalores distintos λ_1 e λ_2, respectivamente, então

$$(A\mathbf{x}_1)^H\mathbf{x}_2 = \mathbf{x}_1^H A^H\mathbf{x}_2 = \mathbf{x}_1^H A\mathbf{x}_2 = \lambda_2\mathbf{x}_1^H\mathbf{x}_2$$

e

$$(A\mathbf{x}_1)^H\mathbf{x}_2 = (\mathbf{x}_2^H A\mathbf{x}_1)^H = (\lambda_1\mathbf{x}_2^H\mathbf{x}_1)^H = \lambda_1\mathbf{x}_1^H\mathbf{x}_2$$

Logo,

$$\lambda_1\mathbf{x}_1^H\mathbf{x}_2 = \lambda_2\mathbf{x}_1^H\mathbf{x}_2$$

e, como $\lambda_1 \neq \lambda_2$, temos

$$\langle\mathbf{x}_2, \mathbf{x}_1\rangle = \mathbf{x}_1^H\mathbf{x}_2 = 0 \qquad \square$$

Definição. Uma matriz U $n \times n$ é dita **unitária** se seus vetores colunas formam um conjunto ortonormal em C^n.

Logo, U é unitária se e somente se $U^H U = I$. Se U é unitária, então, já que suas colunas são ortonormais, U tem posto n. Segue que

$$U^{-1} = IU^{-1} = U^H UU^{-1} = U^H$$

Uma matriz unitária real é uma matriz ortogonal.

Corolário 6.4.2. *Se os autovalores de uma matriz auto-adjunta A forem distintos, então existe uma matriz unitária U que diagonaliza A.*

Demonstração. Seja $\mathbf{x}_i$ um autovetor associado a λ_i para cada autovalor λ_i de A. Seja $\mathbf{u}_i = (1/\|\mathbf{x}_i\|)\mathbf{x}_i$. Então, $\mathbf{u}_i$ é um autovetor unitário associado a λ_i para cada i. Pelo Teorema 6.4.1, $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ é um conjunto ortonormal em C^n. Seja U a matriz cuja i-ésima coluna é o vetor $\mathbf{u}_i$ para cada i: U é unitária e diagonaliza A. $\square$

EXEMPLO 3. Seja

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

Encontre uma matriz unitária que diagonaliza A.

SOLUÇÃO. Os autovalores de A são $\lambda_1 = 3$ e $\lambda_2 = 0$ com autovetores associados $\mathbf{x}_1 = (1 - i, 1)^T$ e $\mathbf{x}_2 = (-1, 1 + i)^T$. Sejam

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{x}_1\|}\mathbf{x}_1 = \frac{1}{\sqrt{3}}(1 - i, 1)^T$$

e

$$\mathbf{u}_2 = \frac{1}{\|\mathbf{x}_2\|}\mathbf{x}_2 = \frac{1}{\sqrt{3}}(-1, 1+i)^T$$

Então

$$U = \frac{1}{\sqrt{3}}\begin{pmatrix} 1-i & -1 \\ 1 & 1+i \end{pmatrix}$$

e

$$\begin{aligned} U^HAU &= \frac{1}{3}\begin{pmatrix} 1+i & 1 \\ -1 & 1-i \end{pmatrix}\begin{pmatrix} 2 & 1-i \\ 1+i & 1 \end{pmatrix}\begin{pmatrix} 1-i & -1 \\ 1 & 1+i \end{pmatrix} \\ &= \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

□

De fato, o Corolário 6.4.2 é válido mesmo se os autovalores de A não forem distintos. Para mostrar isso, vamos provar primeiro o seguinte teorema:

Teorema 6.4.3 (Teorema de Schur). *Se A é uma matriz $n \times n$, então existe uma matriz unitária U tal que U^HAU é triangular superior.*

Demonstração. A demonstração é por indução em n. O resultado é óbvio para $n = 1$. Suponha que a hipótese é válida para matrizes $k \times k$ e seja A uma matriz $(k + 1) \times (k + 1)$. Sejam λ_1 um autovalor de A e $\mathbf{w}_1$ um autovetor unitário associado. Usando o processo de Gram-Schmidt, construa $\mathbf{w}_2, ..., \mathbf{w}_{k+1}$ de modo que $\{\mathbf{w}_1, ..., \mathbf{w}_{k+1}\}$ seja uma base ortonormal para C^{k+1}. Seja W a matriz cuja i-ésima coluna é o vetor $\mathbf{w}_i$ para $i = 1, ..., k + 1$. Então, por construção, W é unitária. A primeira coluna de W^HAW vai ser $W^HA\mathbf{w}_1$.

$$W^HA\mathbf{w}_1 = \lambda_1 W^H\mathbf{w}_1 = \lambda_1\mathbf{e}_1$$

Então W^HAW é uma matriz da forma

$$\left(\begin{array}{c|ccc} \lambda_1 & \times & \times \cdots & \times \\ \hline 0 & & & \\ \vdots & & M & \\ 0 & & & \end{array}\right)$$

onde M é uma matriz $k \times k$. Pela hipótese de indução, existe uma matriz $k \times k$ unitária V_1 tal que $V_1^HMV_1 = T_1$, onde T_1 é triangular. Seja

$$V = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & V_1 & \\ 0 & & & \end{array}\right)$$

Então V é unitária e

$$V^H W^H A W V = \left(\begin{array}{c|ccc} \lambda_1 & \times & \cdots & \times \\ \hline 0 & & & \\ \vdots & & V_1^H M V_1 & \\ 0 & & & \end{array}\right)$$

$$= \left(\begin{array}{c|ccc} \lambda_1 & \times & \cdots & \times \\ \hline 0 & & & \\ \vdots & & T_1 & \\ 0 & & & \end{array}\right)$$

$$= T$$

Seja $U = WV$. A matriz U é unitária, já que

$$U^H U = (WV)^H WV = V^H W^H WV = I$$

e $U^H AU = T$. □

A fatoração $A = UTU^H$ é conhecida como a *decomposição de Schur* de A. No caso em que A é auto-adjunta, a matriz T é diagonal.

Teorema 6.4.4 (Teorema Espectral). *Se* A *é auto-adjunta, então existe uma matriz unitária U que diagonaliza A.*

Demonstração. Pelo Teorema 6.4.3, existe uma matriz unitária U tal que $U^H AU = T$, onde T é triangular superior.

$$T^H = (U^H AU)^H = U^H A^H U = U^H AU = T$$

Logo, T é auto-adjunta e, portanto, tem que ser diagonal. □

No caso em que A é real e simétrica, seus autovalores e autovetores têm que ser reais. Assim, a matriz diagonalizante U deve ser ortogonal.

Corolário 6.4.5. *Se A é uma matriz real e simétrica, então existe uma matriz ortogonal U que diagonaliza A, isto é, tal que $U^T AU = D$, onde D é diagonal.*

EXEMPLO 4. Dada

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

encontre uma matriz ortogonal U que diagonaliza A.

SOLUÇÃO. O polinômio característico

$$p(\lambda) = -\lambda^3 + 3\lambda^2 + 9\lambda + 5 = (1 + \lambda)^2(5 - \lambda)$$

tem raízes $\lambda_1 = \lambda_2 = -1$, $\lambda_3 = 5$. Calculando os autovetores associados da maneira usual, vemos que $\mathbf{x}_1 = (1, 0, 1)^T$ e $\mathbf{x}_2 = (-2, 1, 0)^T$ formam uma base para o auto-espaço $N(A + I)$. Podemos aplicar o processo de Gram-Schmidt para obter uma base ortonormal para o auto-espaço associado a $\lambda_1 = \lambda_2 = -1$.

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{x}_1\|}\mathbf{x}_1 = \frac{1}{\sqrt{2}}(1, 0, 1)^T$$
$$\mathbf{p} = \left(\mathbf{x}_2^T\mathbf{u}_1\right)\mathbf{u}_1 = -\sqrt{2}\mathbf{u}_1 = (-1, 0, 1)^T$$
$$\mathbf{x}_2 - \mathbf{p} = (-1, 1, 1)^T$$
$$\mathbf{u}_2 = \frac{1}{\|\mathbf{x}_2 - \mathbf{p}\|}(\mathbf{x}_2 - \mathbf{p}) = \frac{1}{\sqrt{3}}(-1, 1, 1)^T$$

O auto-espaço associado a $\lambda_3 = 5$ é gerado por $\mathbf{x}_3 = (-1, -2, 1)^T$. Como $\mathbf{x}_3$ tem que ser ortogonal a $\mathbf{u}_1$ e $\mathbf{u}_2$ (Teorema 6.4.1), precisamos apenas normalizá-lo:

$$\mathbf{u}_3 = \frac{1}{\|\mathbf{x}_3\|}\mathbf{x}_3 = \frac{1}{\sqrt{6}}(-1, -2, 1)^T$$

Logo, $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ é um conjunto ortonormal e

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix}$$

diagonaliza A. □

Segue, do Teorema 6.4.4, que toda matriz auto-adjunta pode ser fatorada em um produto UDU^H, onde U é unitária e D é diagonal. Como U diagonaliza A, os elementos diagonais em D são os autovalores de A e as colunas de U são autovetores de A. Logo, A é diagonalizável e tem um conjunto completo de autovetores que formam uma base ortonormal para C^n. De certa forma, essa é a situação ideal. Vimos como expressar um vetor como uma combinação linear dos elementos em uma base ortonormal (Teorema 5.5.2) e a ação de A em qualquer combinação linear de autovetores pode ser determinada facilmente. Portanto, se A tem um conjunto ortonormal de autovetores $\{\mathbf{u}_1, ..., \mathbf{u}_n\}$ e se $\mathbf{x} = c_1\mathbf{u}_1 + ... + c_n\mathbf{u}_n$, temos

$$A\mathbf{x} = c_1\lambda_1\mathbf{u}_1 + \cdots + c_n\lambda_n\mathbf{u}_n$$

Além disso,

$$c_i = \langle \mathbf{x}, \mathbf{u}_i \rangle = \mathbf{u}_i^H\mathbf{x}$$

ou, de modo equivalente, $\mathbf{c} = U^h\mathbf{x}$. Logo,

$$A\mathbf{x} = \lambda_1(\mathbf{u}_1^H\mathbf{x})\mathbf{u}_1 + \cdots + \lambda_n(\mathbf{u}_n^H\mathbf{x})\mathbf{u}_n$$

MATRIZES NORMAIS

Existem matrizes não-auto-adjuntas que possuem um conjunto completo de autovetores ortonormais. Por exemplo, matrizes anti-simétricas e antiadjuntas* têm essa propriedade. (A é *antiadjunta* se $A^H = -A$.) Em geral, se A é uma matriz com um conjunto completo de autovetores ortonormais, então $A = UDU^H$, onde U é unitária e D é diagonal (cujos elementos diagonais podem ser complexos). Em geral, $D^H \neq D$ e, portanto,

$$A^H = UD^HU^H \neq A$$

*Essa terminologia não é padrão; usa-se, também, a terminologia *anti-hermitiana*. (*N. T.*)

No entanto,

$$AA^H = UDU^HUD^HU^H = UDD^HU^H$$

e

$$A^HA = UD^HU^HUDU^H = UD^HDU^H$$

Como

$$D^HD = DD^H = \begin{pmatrix} |\lambda_1|^2 & & & \\ & |\lambda_2|^2 & & \\ & & \ddots & \\ & & & |\lambda_n|^2 \end{pmatrix}$$

temos que

$$AA^H = A^HA$$

Definição. Uma matriz A é dita **normal** se $AA^H = A^HA$.

Mostramos que, se existe um conjunto ortonormal completo de autovetores de uma matriz, então ela é normal. A recíproca também é verdadeira.

Teorema 6.4.6. *Uma matriz A é normal se e somente existe um conjunto ortonormal completo de autovetores de A.*

Demonstração. Em vista das observações anteriores, precisamos mostrar, apenas, que existe um conjunto completo de autovetores para toda matriz normal A. Pelo Teorema 6.4.3, existem uma matriz unitária U e uma matriz triangular T tais que $T = U^HAU$. Vamos mostrar que T é normal.

$$T^HT = U^HA^HUU^HAU = U^HA^HAU$$

e

$$TT^H = U^HAUU^HA^HU = U^HAA^HU$$

Como $A^HA = AA^H$, temos que $T^HT = TT^H$. Comparando os elementos diagonais de TT^H e T^HT, vemos que

$$\begin{aligned} |t_{11}|^2 + |t_{12}|^2 + |t_{13}|^2 + \cdots + |t_{1n}|^2 &= |t_{11}|^2 \\ |t_{22}|^2 + |t_{23}|^2 + \cdots + |t_{2n}|^2 &= |t_{12}|^2 + |t_{22}|^2 \\ &\vdots \\ |t_{nn}|^2 &= |t_{1n}|^2 + |t_{2n}|^2 + |t_{3n}|^2 + \cdots + |t_{nn}|^2 \end{aligned}$$

Então $t_{ij} = 0$ sempre que $i \neq j$. Logo, U diagonaliza A e as colunas de U são autovetores de A. □

EXERCÍCIOS

1. Para cada um dos pares dados de vetores **z** e **w** em C^2, calcule (i) $\|\mathbf{z}\|$, (ii) $\|\mathbf{w}\|$, (iii) $\langle \mathbf{z}, \mathbf{w} \rangle$, (iv) $\langle \mathbf{w}, \mathbf{z} \rangle$.

(a) $\mathbf{z} = \begin{pmatrix} 4+2i \\ 4i \end{pmatrix}$, $\quad \mathbf{w} = \begin{pmatrix} -2 \\ 2+i \end{pmatrix}$

(b) $\mathbf{z} = \begin{pmatrix} 1+i \\ 2i \\ 3-i \end{pmatrix}$, $\quad \mathbf{w} = \begin{pmatrix} 2-4i \\ 5 \\ 2i \end{pmatrix}$

2. Sejam

$$\mathbf{z}_1 = \begin{pmatrix} \frac{1+i}{2} \\ \frac{1-i}{2} \end{pmatrix} \quad \text{e} \quad \mathbf{z}_2 = \begin{pmatrix} \frac{i}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

(a) Mostre que $\{\mathbf{z}_1, \mathbf{z}_2\}$ é um conjunto ortonormal em C^2.

(b) Escreva o vetor $\mathbf{z} = \begin{pmatrix} 2+4i \\ -2i \end{pmatrix}$ como uma combinação linear de $\mathbf{z}_1$ e $\mathbf{z}_2$.

3. Seja $\{\mathbf{u}_1, \mathbf{u}_2\}$ uma base ortonormal para C^2 e seja $\mathbf{z} = (4 + 2i)\mathbf{u}_1 + (6 - 5i)\mathbf{u}_2$.

(a) Quais os valores de $\mathbf{u}_1^H\mathbf{z}$, $\mathbf{z}^H\mathbf{u}_1$, $\mathbf{u}_2^H\mathbf{z}$ e $\mathbf{z}^H\mathbf{u}_2$?

(b) Determine o valor de $\|\mathbf{z}\|$.

4. Quais das matrizes a seguir são auto-adjuntas? Quais são normais?

(a) $\begin{pmatrix} 1-i & 2 \\ 2 & 3 \end{pmatrix}$ $\qquad$ (b) $\begin{pmatrix} 1 & 2-i \\ 2+i & -1 \end{pmatrix}$

(c) $\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$ $\qquad$ (d) $\begin{pmatrix} \frac{1}{\sqrt{2}}i & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}i \end{pmatrix}$

(e) $\begin{pmatrix} 0 & i & 1 \\ i & 0 & -2+i \\ -1 & 2+i & 0 \end{pmatrix}$ $\qquad$ (f) $\begin{pmatrix} 3 & 1+i & i \\ 1-i & 1 & 3 \\ -i & 3 & 1 \end{pmatrix}$

5. Encontre uma matriz diagonalizante ortogonal ou unitária para cada uma das matrizes a seguir.

(a) $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ $\qquad$ (b) $\begin{pmatrix} 1 & 3+i \\ 3-i & 4 \end{pmatrix}$

(c) $\begin{pmatrix} 2 & i & 0 \\ -i & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ $\qquad$ (d) $\begin{pmatrix} 2 & 1 & 1 \\ 1 & 3 & -2 \\ 1 & -2 & 3 \end{pmatrix}$

(e) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ $\qquad$ (f) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

(g) $\begin{pmatrix} 4 & 2 & -2 \\ 2 & 1 & -1 \\ -2 & -1 & 1 \end{pmatrix}$

6. Mostre que os elementos diagonais de uma matriz auto-adjunta têm que ser reais.

7. Sejam A e C matrizes em $C^{m \times n}$ e seja $B \in C^{n \times r}$. Prove cada uma das regras a seguir.

(a) $(A^H)^H = A$
(b) $(\alpha A + \beta C)^H = \overline{\alpha} A^H + \overline{\beta} C^H$
(c) $(AB)^H = B^H A^H$

8. Mostre que

$$\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{w}^H \mathbf{z}$$

define um produto interno em C^n.

9. Seja $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ uma base ortonormal para um espaço vetorial V munido de um produto interno complexo e sejam $\mathbf{z}$ e $\mathbf{w}$ elementos de V. Mostre que

$$\langle \mathbf{z}, \mathbf{w} \rangle = \sum_{i=1}^{n} (\langle \mathbf{z}, \mathbf{u}_i \rangle \cdot \langle \mathbf{u}_i, \mathbf{w} \rangle)$$

10. Dada

$$A = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & i \\ 0 & -i & 1 \end{pmatrix}$$

encontre uma matriz B tal que $B^H B = A$.

11. Seja U uma matriz unitária. Prove que:
(a) U é normal;
(b) $\|U\mathbf{x}\| = \|\mathbf{x}\|$ para todo $\mathbf{x}$ em C^n;
(c) Se λ é um autovalor de U, então $|\lambda| = 1$.

12. Seja $\mathbf{u}$ um vetor unitário em C^n e defina $U = I - 2\mathbf{u}\mathbf{u}^H$. Mostre que U é unitária e auto-adjunta, de modo que é igual à sua própria inversa.

13. Mostre que, se uma matriz U é unitária e auto-adjunta, então qualquer autovalor de U tem que ser igual a 1 ou a -1.

14. Seja A uma matriz 2×2 com decomposição de Schur UTU^H e suponha que $t_{12} \neq 0$. Mostre que:
(a) Os autovalores de A são $\lambda_1 = t_{11}$ e $\lambda_2 = t_{22}$;
(b) $\mathbf{u}_1$ é um autovetor associado a $\lambda_1 = t_{11}$;
(c) $\mathbf{u}_2$ não é um autovetor associado a $\lambda_2 = t_{22}$.

15. Mostre que $M = A + iB$ (A e B matrizes reais) é antiadjunta se e somente se A é anti-simétrica e B é simétrica.

16. Mostre que, se A é antiadjunta e se λ é um autovalor de A, então λ é um imaginário puro (isto é, $\lambda = bi$, onde b é real).

17. Seja A uma matriz 2×2 com a propriedade de que $a_{21}a_{12} > 0$.

(a) Faça $r = \sqrt{a_{21}/a_{12}}$, $S = \begin{pmatrix} r & 0 \\ 0 & 1 \end{pmatrix}$ e calcule $B = SAS^{-1}$.
(b) O que você pode concluir sobre os autovalores e autovetores de B? O que você pode concluir sobre os autovalores e autovetores de A? Explique.

18. Seja $p(x) = -x^3 + cx^2 + (c+3)x + 1$, onde c é um número real. Seja C a matriz companheira de $p(x)$,

$$C = \begin{pmatrix} c & c+3 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

e seja

$$A = \begin{pmatrix} -1 & 2 & -c-3 \\ 1 & -1 & c+2 \\ -1 & 1 & -c-1 \end{pmatrix}$$

(a) Calcule $A^{-1}CA$.
(b) Use o resultado do item (a) para provar que $p(x)$ tem apenas raízes reais, independentemente do valor de c.

19. Seja A uma matriz auto-adjunta com autovalores $\lambda_1, \ldots, \lambda_n$ e autovetores ortonormais $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Mostre que

$$A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^H + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^H + \cdots + \lambda_n \mathbf{u}_n \mathbf{u}_n^H$$

20. Seja

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Escreva A como uma soma $\lambda_1\mathbf{u}_1\mathbf{u}_1^T + \lambda_2\mathbf{u}_2\mathbf{u}_2^T$, onde λ_1 e λ_2 são autovalores e $\mathbf{u}_1$ e $\mathbf{u}_2$ autovetores ortonormais.

21. Seja A uma matriz auto-adjunta com autovalores $\lambda_1 \leq \ldots \leq \lambda_n$ e autovetores ortonormais $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Qualquer que seja o vetor não-nulo $\mathbf{x}$ em R^n, o *quociente de Rayleigh* $\rho(\mathbf{x})$ é definido por

$$\rho(\mathbf{x}) = \frac{\langle A\mathbf{x}, \mathbf{x}\rangle}{\langle \mathbf{x}, \mathbf{x}\rangle} = \frac{\mathbf{x}^H A\mathbf{x}}{\mathbf{x}^H\mathbf{x}}$$

(a) Se $\mathbf{x} = \alpha_1\mathbf{u}_1 + \ldots + \alpha_n\mathbf{u}_n$, mostre que

$$\rho(\mathbf{x}) = \frac{|\alpha_1|^2\lambda_1 + |\alpha_2|^2\lambda_2 + \cdots + |\alpha_n|^2\lambda_n}{\|\boldsymbol{\alpha}\|^2}$$

(b) Mostre que

$$\lambda_1 \leq \rho(\mathbf{x}) \leq \lambda_n$$

5 FORMAS QUADRÁTICAS

A essa altura o leitor já deve estar bem consciente do papel importante das matrizes no estudo de equações lineares. Nesta seção, vamos ver que as matrizes também têm um papel importante no estudo das equações quadráticas. A cada equação de segundo grau podemos associar uma função vetorial $f(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$. Tal função vetorial é chamada de uma "forma quadrática". Formas quadráticas aparecem em uma grande variedade de problemas aplicados. Elas são particularmente importantes na teoria de otimização.

Definição. Uma **equação quadrática** em duas variáveis x e y é uma equação da forma

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0 \tag{1}$$

A Equação (1) pode ser colocada na forma

$$(x \quad y)\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + (d \quad e)\begin{pmatrix} x \\ y \end{pmatrix} + f = 0 \tag{2}$$

Sejam

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{e} \quad A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

O termo

$$\mathbf{x}^T A \mathbf{x} = ax^2 + 2bxy + cy^2$$

é chamado de **forma quadrática** associada a (1).

SEÇÕES CÔNICAS

O gráfico de uma equação da forma (1) é uma *seção cônica.* [No caso em que não existem pares ordenados satisfazendo (1), dizemos que a equação representa uma cônica imaginária.] No caso em que o gráfico de (1) consiste em um único ponto, em uma reta ou em um par de retas, dizemos que (1) representa uma cônica degenerada. Os gráficos de cônicas não-degeneradas são círculos, elipses, parábolas ou hipérboles (ver Fig. 6.5.1). O gráfico de uma cônica é particularmente fácil de desenhar quando sua equação pode ser colocada em uma das seguintes formas:

$$\text{(i)}\quad x^2 + y^2 = r^2 \qquad \text{(círculo)}$$

$$\text{(ii)}\quad \frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 1 \qquad \text{(elipse)}$$

$$\text{(iii)}\quad \frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} = 1 \quad \text{ou} \quad \frac{y^2}{\alpha^2} - \frac{x^2}{\beta^2} = 1 \qquad \text{(hipérbole)}$$

$$\text{(iv)}\quad x^2 = \alpha y \quad \text{ou} \quad y^2 = \alpha x \qquad \text{(parábola)}$$

onde α, β e r são números reais não-nulos. Observe que o círculo é um caso especial da elipse ($\alpha = \beta = r$). Uma cônica está em *posição canônica* se sua equação pode ser colocada em uma das quatro formas acima. Os gráficos de (i), (ii) e (iii) vão ser simétricos em relação a ambos os eixos coordenados e em relação à origem. Uma parábola em posição canônica terá seu vértice na origem e será simétrica em relação a um dos eixos coordenados.

E as cônicas que não estão em posição canônica? Vamos considerar os seguintes casos:

Caso 1

A seção cônica foi transladada horizontalmente da posição canônica. Isso acontece quando ambos os termos em x^2 e em x na Equação (1) têm coeficientes não-nulos.

Caso 2

A seção cônica foi transladada verticalmente da posição canônica. Isso acontece quando os termos em y^2 e em y na Equação (1) têm coeficientes não-nulos (isto é, $c \neq 0$ e $e \neq 0$).

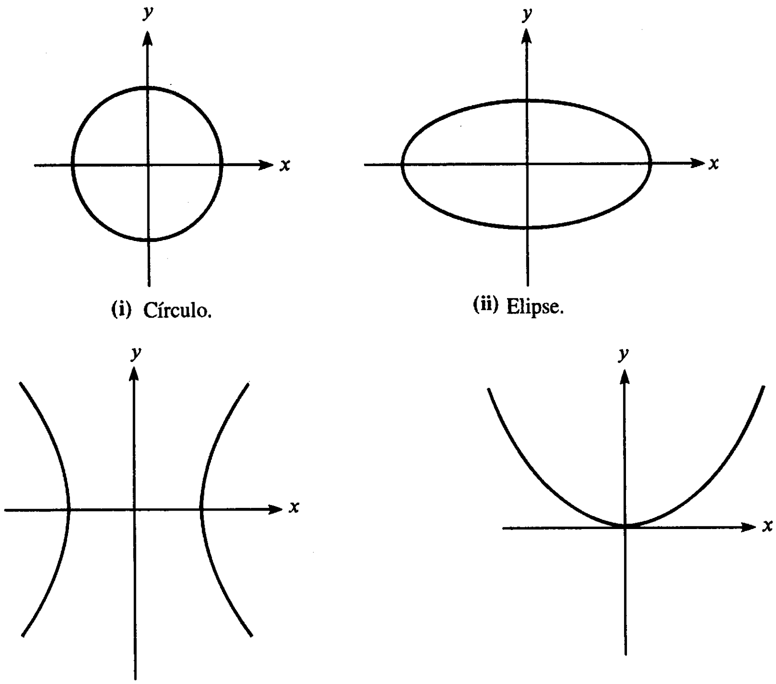
Caso 3

A seção cônica foi rodada da posição canônica por um ângulo θ diferente de 90°. Isso acontece quando o coeficiente de xy é não-nulo (isto é, $b \neq 0$).

Em geral, podemos ter qualquer um, ou uma combinação, desses três casos. Para desenhar uma cônica que não está em posição canônica, normalmente procuramos um outro conjunto de eixos x' e y' de modo que a cônica esteja em posição canônica em relação a esses novos eixos. Isso não é difícil se a cônica tiver sido transladada horizontal ou verticalmente: nesse caso, os novos eixos podem ser encontrados completando-se os quadrados. O exemplo a seguir ilustra como isso é feito.

EXEMPLO 1. Desenhe o gráfico da equação

$$9x^2 - 18x + 4y^2 + 16y - 11 = 0$$



(i) Círculo. (ii) Elipse. (iii) Hipérbole. (iv) Parábola.

FIG. 6.5.1

SOLUÇÃO. Para escolher os novos eixos, completamos os quadrados.

$$9(x^2 - 2x + 1) + 4(y^2 + 4y + 4) - 11 = 9 + 16$$

Essa equação pode ser simplificada para

$$\frac{(x-1)^2}{2^2} + \frac{(y+2)^2}{3^2} = 1$$

Definindo

$$x' = x - 1 \quad \text{e} \quad y' = y + 2$$

temos

$$\frac{(x')^2}{2^2} + \frac{(y')^2}{3^2} = 1$$

que é a forma canônica em relação às variáveis x' e y'. Logo, o gráfico, como mostra a Fig. 6.5.2, é uma elipse na posição canônica em relação aos eixos x' e y'. O eixo da elipse é a origem no plano $x'y'$ [isto é, o ponto $(x, y) = (1, -2)$]. O eixo x' é, simplesmente, $y' = 0$, que é a equação da reta $y = -2$ no plano xy. Analogamente, o eixo dos y' coincide com a reta $x = 1$. □

Não tem muito problema se o vértice da cônica tiver sido transladado. Se, no entanto, a cônica também tiver sido rodada de sua posição canônica, é necessário mudar para novas coordenadas x' e y' de modo que a equação não tenha termo $x'y'$. Sejam $\mathbf{x} = (x, y)^T$ e $\mathbf{x}' = (x', y')^T$. Como as novas coordenadas diferem das antigas por uma rotação, temos

$$\mathbf{x} = Q\mathbf{x}' \qquad \text{ou} \qquad \mathbf{x}' = Q^T\mathbf{x}$$

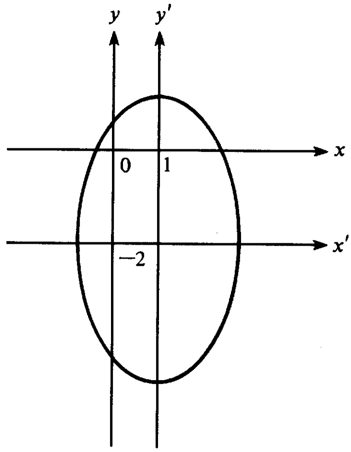


FIG. 6.5.2

onde

$$Q = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix} \quad \text{ou} \quad Q^T = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

Se $0 < \theta < \pi$, a matriz Q corresponde a uma rotação de θ radianos no sentido trigonométrico e Q^T corresponde a uma rotação de θ radianos no sentido horário (ver Exemplo 2 na Seção 2 do Cap. 4). Com essa mudança de variáveis, a Equação (2) fica

$$(3) \qquad (\mathbf{x}')^T(Q^TAQ)\mathbf{x}' + \begin{pmatrix} d' & e' \end{pmatrix}\mathbf{x}' + f = 0$$

onde $(d' \quad e') = (d \quad e)Q$. Essa equação não tem termo em $x'y'$ se e somente se Q^TAQ é diagonal. Como A é simétrica, é possível encontrar um par de autovetores ortonormais $\mathbf{q}_1 = (x_1, -y_1)^T$ e $\mathbf{q}_2 = (y_1, x_1)^T$. Logo, fazendo $\cos\theta = x_1$ e $\operatorname{sen}\theta = y_1$, temos que

$$Q = \begin{pmatrix} \mathbf{q}_1 & \mathbf{q}_2 \end{pmatrix} = \begin{pmatrix} x_1 & y_1 \\ -y_1 & x_1 \end{pmatrix}$$

diagonaliza A e (3) pode ser simplificada para

$$\lambda_1(x')^2 + \lambda_2(y')^2 + d'x' + e'y' + f = 0$$

EXEMPLO 2. Considere a seção cônica

$$3x^2 + 2xy + 3y^2 - 8 = 0$$

Essa equação pode ser colocada na forma

$$\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 8$$

A matriz

$$\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$

tem autovalores $\lambda = 2$ e $\lambda = 4$ com autovetores associados

$$\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)^T \quad \text{e} \quad \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T$$

Seja

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \cos 45° & \operatorname{sen} 45° \\ -\operatorname{sen} 45° & \cos 45° \end{pmatrix}$$

e faça

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

Então

$$Q^T A Q = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$$

e a equação da cônica fica

$$2(x')^2 + 4(y')^2 = 8$$

ou

$$\frac{(x')^2}{4} + \frac{(y')^2}{2} = 1$$

No novo sistema de coordenadas, a direção e o sentido do eixo dos x' são determinados pelo ponto $x' = 1$, $y' = 0$. Colocando esse ponto no sistema de coordenadas xy, temos

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = \mathbf{q}_1$$

O eixo x' tem a mesma direção e o mesmo sentido que $\mathbf{q}_1$. Analogamente, para encontrar a direção e o sentido do eixo dos y', multiplicamos

$$Q\mathbf{e}_2 = \mathbf{q}_2$$

Os autovetores que formam as colunas de Q nos dão a direção e o sentido dos novos eixos coordenados (ver Fig. 6.5.3). □

EXEMPLO 3. Dada a equação quadrática

$$3x^2 + 2xy + 3y^2 + 8\sqrt{2}y - 4 = 0$$

encontre uma mudança de coordenadas de modo que a equação resultante represente uma cônica em posição canônica.

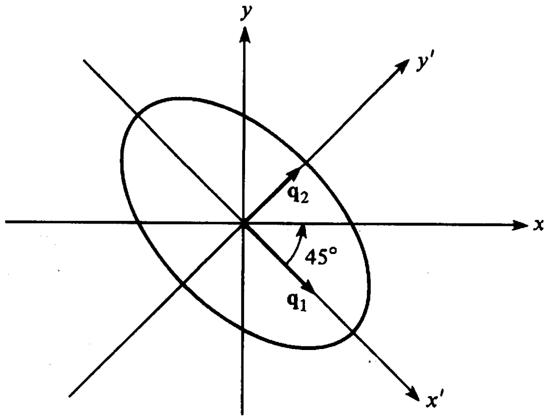


FIG. 6.5.3

SOLUÇÃO. O termo xy é eliminado da mesma maneira que no Exemplo 2. Nesse caso, usando a matriz de rotação

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

a equação fica

$$2(x')^2 + 4(y')^2 + \begin{pmatrix} 0 & 8\sqrt{2} \end{pmatrix} Q \begin{pmatrix} x' \\ y' \end{pmatrix} = 4$$

ou

$$(x')^2 - 4x' + 2(y')^2 + 4y' = 2$$

Completando os quadrados, obtemos

$$(x' - 2)^2 + 2(y' + 1)^2 = 8$$

Fazendo $x'' = x' - 2$ e $y'' = y' + 1$ (ver Fig. 6.5.4), a equação simplifica para

$$\frac{(x'')^2}{8} + \frac{(y'')^2}{4} = 1$$

□

Resumindo, uma equação quadrática nas variáveis x e y pode ser colocada na forma

$$\mathbf{x}^T A\mathbf{x} + B\mathbf{x} + f = 0$$

onde $\mathbf{x} = (x, y)^T$, A é uma matriz simétrica 2×2, B é uma matriz 1×2 e f é um escalar. Se A for invertível, é possível, através de rotação e translação dos eixos, colocar a equação na forma

(4) $$\lambda_1(x')^2 + \lambda_2(y')^2 + f' = 0$$

onde λ_1 e λ_2 são os autovalores de A. Se (4) representar uma cônica real não-degenerada, ela vai ser uma elipse ou uma hipérbole, dependendo se λ_1 e λ_2 têm o mesmo sinal ou sinais diferentes. Se A for singular

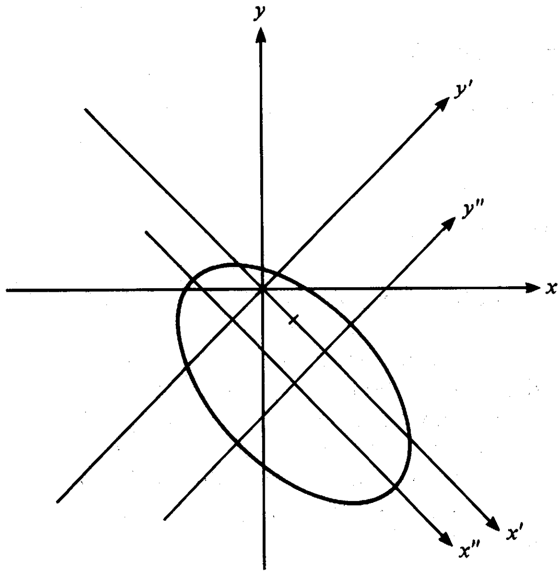


FIG. 6.5.4

e exatamente um de seus autovalores for nulo, a equação quadrática pode ser colocada em uma das seguintes formas:

$$\lambda_1(x')^2 + e'y' + f' = 0 \qquad \text{ou} \qquad \lambda_2(y')^2 + d'x' + f' = 0$$

Essas equações representam parábolas, desde que e' e d' não sejam iguais a zero.

Não há razão para nos restringirmos a duas variáveis. Poderíamos ter equações do segundo grau e formas quadráticas com qualquer número de variáveis. De fato, uma *equação quadrática em n variáveis* $x_1, \ldots, x_n$ é uma equação da forma

(5) $$\mathbf{x}^T A\mathbf{x} + B\mathbf{x} + \alpha = 0$$

onde $\mathbf{x} = (x_1, \ldots, x_n)^T$, A é uma matriz simétrica $n \times n$, B é uma matriz $1 \times n$ e α é um escalar. A função vetorial

$$f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} x_j \right) x_i$$

é a *forma quadrática em n variáveis* associada à equação quadrática.

No caso de três variáveis, se

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \qquad A = \begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix}, \qquad B = \begin{pmatrix} g \\ h \\ i \end{pmatrix}$$

(5) fica

$$ax^2 + by^2 + cz^2 + 2dxy + 2exz + 2fyz + gx + hy + iz + \alpha = 0$$

O gráfico de uma equação quadrática em três variáveis é uma *quádrica*. Existem quatro tipos básicos de quádricas não-degeneradas:

1. elipsóides;
2. hiperbolóides (de uma ou duas folhas);

3. cones;
4. parabolóides (elípticos ou hiperbólicos).

Como no caso bidimensional, podemos usar translações e rotações para colocar a equação em forma canônica,

$$\lambda_1(x')^2 + \lambda_2(y')^2 + \lambda_3(z')^2 + \alpha = 0$$

onde λ_1, λ_2, λ_3 são os autovalores de A.

OTIMIZAÇÃO: UMA APLICAÇÃO AO CÁLCULO

Vamos considerar o problema de encontrar máximos e mínimos de funções de várias variáveis. Em particular, gostaríamos de determinar a natureza dos pontos críticos de uma função vetorial com valores reais $w = F(\mathbf{x})$. Se a função é uma forma quadrática, $w = \mathbf{x}^T A\mathbf{x}$, então $\mathbf{0}$ é um ponto crítico. Se esse é um ponto de máximo, de mínimo ou de sela, depende dos autovalores de A. Mais geralmente, se a função da qual queremos calcular os extremos for suficientemente diferenciável, ela vai se comportar localmente como uma forma quadrática. Logo, cada ponto crítico pode ser testado através dos sinais dos autovalores da matriz associada à forma quadrática.

Definição. Seja $F(\mathbf{x})$ uma função com valores reais definida em R^n. Um ponto $\mathbf{x}_0$ em R^n é dito um **ponto estacionário** de F se todas as derivadas parciais de F em $\mathbf{x}_0$ existem e são iguais a zero.

Se $F(\mathbf{x})$ tem um máximo local ou um mínimo local em $\mathbf{x}_0$ e se as derivadas parciais de F existem em $\mathbf{x}_0$, então elas são todas nulas. Logo, se $F(\mathbf{x})$ tiver derivadas parciais em todos os pontos, seus máximos e mínimos locais vão ocorrer em pontos estacionários.

Considere a forma quadrática

$$f(x, y) = ax^2 + 2bxy + cy^2$$

As derivadas parciais de primeira ordem de f são

$$f_x = 2ax + 2by$$
$$f_y = 2bx + 2cy$$

Igualando a zero, vemos que (0, 0) é um ponto estacionário. Além disso, se a matriz

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

for invertível, esse vai ser o único ponto crítico*. Logo, se A for invertível, f vai ter um mínimo global, ou um máximo global ou um ponto de sela em (0, 0).

Vamos escrever f na forma

$$f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} \quad \text{onde} \quad \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$$

Como $f(\mathbf{0}) = 0$, f vai ter um mínimo global em $\mathbf{0}$ se e somente se

$$\mathbf{x}^T A\mathbf{x} > 0 \quad \text{para todo} \quad \mathbf{x} \neq \mathbf{0}$$

e f vai ter um máximo global em $\mathbf{0}$ se e somente se

$$\mathbf{x}^T A\mathbf{x} < 0 \quad \text{para todo} \quad \mathbf{x} \neq \mathbf{0}$$

---

* $\mathbf{x}_0$ é um ponto crítico de f se todas as derivadas parciais de f se anulam em $\mathbf{x}_0$ ou se alguma(s) das derivadas parciais não existe(m) em $\mathbf{x}_0$, embora exista(m) em todos os pontos diferentes de $\mathbf{x}_0$ pertencentes a uma vizinhança de $\mathbf{x}_0$. *(N. T.)*

Se $\mathbf{x}^T A\mathbf{x}$ mudar de sinal, $\mathbf{0}$ vai ser um ponto de sela.

Em geral, se f é uma forma quadrática em n variáveis, então, para todo $\mathbf{x} \in R^n$,

$$f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$$

onde A é uma matriz simétrica $n \times n$.

Definição. Uma forma quadrática $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ é dita **definida** se ela tem o mesmo sinal para todos os vetores não-nulos $\mathbf{x}$ em R^n. A forma é **positiva definida** se $\mathbf{x}^T A\mathbf{x} > 0$ para todos os vetores não-nulos $\mathbf{x}$ em R^n e é **negativa definida** se $\mathbf{x}^T A\mathbf{x} < 0$ para todos os vetores não-nulos $\mathbf{x}$ em R^n. Se a forma muda de sinal, dizemos que ela é **indefinida.** Se $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} \geq 0$ e tem valor zero para algum $\mathbf{x} \neq \mathbf{0}$, dizemos que $f(\mathbf{x})$ é **positiva semidefinida**. Se $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} \leq 0$ e tem valor zero para algum $\mathbf{x} \neq \mathbf{0}$, dizemos que $f(\mathbf{x})$ é **negativa semidefinida.**

Se a forma quadrática é positiva definida ou negativa definida, depende da matriz A. Se a forma quadrática é positiva definida, dizemos simplesmente que A é positiva definida. Podemos, então, colocar a definição acima de outra maneira:

Definição. Uma matriz simétrica real A é dita

(i) **positiva definida** se $\mathbf{x}^T A\mathbf{x} > 0$ para todo $\mathbf{x}$ em R^n diferente de zero;
(ii) **negativa definida** se $\mathbf{x}^T A\mathbf{x} < 0$ para todo $\mathbf{x}$ em R^n diferente de zero;
(iii) **positiva semidefinida** se $\mathbf{x}^T A\mathbf{x} \geq 0$ para todo $\mathbf{x}$ em R^n diferente de zero;
(iv) **negativa semidefinida** se $\mathbf{x}^T A\mathbf{x} \leq 0$ para todo $\mathbf{x}$ em R^n diferente de zero.

Se A for invertível, $\mathbf{0}$ vai ser o único ponto estacionário de $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$. Esse ponto vai ser um mínimo global se A for positiva definida e um máximo global se A for negativa definida. Se A for indefinida, $\mathbf{0}$ vai ser um ponto de sela. Para classificar o ponto estacionário, precisamos, portanto, classificar a matriz A. Existem diversas maneiras de se decidir se uma matriz é ou não positiva definida. Vamos estudar alguns desses métodos na próxima seção. O teorema a seguir nos dá o que é, talvez, a caracterização mais importante de matrizes positivas definidas.

Teorema 6.5.1. *Seja A uma matriz real simétrica $n \times n$. Então, A é positiva definida se e somente se todos os seus autovalores são positivos.*

Demonstração. Se A é positiva definida e se λ é um autovalor de A, então, para qualquer autovetor $\mathbf{x}$ associado a λ,

$$\mathbf{x}^T A\mathbf{x} = \lambda\mathbf{x}^T\mathbf{x} = \lambda\|\mathbf{x}\|^2$$

Logo,

$$\lambda = \frac{\mathbf{x}^T A\mathbf{x}}{\|\mathbf{x}\|^2} > 0$$

Reciprocamente, suponha que todos os autovalores de A são positivos. Seja $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ um conjunto ortonormal de autovetores de A. Se $\mathbf{x}$ é um vetor não-nulo em R^n, então $\mathbf{x}$ pode ser escrito como

$$\mathbf{x} = \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_n\mathbf{x}_n$$

onde

$$\alpha_i = \mathbf{x}^T\mathbf{x}_i \quad \text{para} \quad i = 1, \ldots, n \quad \text{e} \quad \sum_{i=1}^{n}(\alpha_i)^2 = \|\mathbf{x}\|^2 > 0$$

Temos, então, que

$$\begin{aligned}\mathbf{x}^T A\mathbf{x} &= (\alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n)^T(\alpha_1\lambda_1\mathbf{x}_1 + \cdots + \alpha_n\lambda_n\mathbf{x}_n)\\ &= \sum_{i=1}^{n}(\alpha_i)^2\lambda_i\\ &\geq (\min \lambda_i)\|\mathbf{x}\|^2 > 0\end{aligned}$$

e, portanto, A é positiva definida. □

Se os autovalores de A forem todos negativos, então $-A$ tem que ser positiva definida e, conseqüentemente, A tem que ser negativa definida. Se A tem autovalores com sinais diferentes, então A é indefinida. De fato, se λ_1 é um autovalor positivo de A com autovetor associado $\mathbf{x}_1$, então

$$\mathbf{x}_1^T A\mathbf{x}_1 = \lambda_1\mathbf{x}_1^T\mathbf{x}_1 = \lambda_1\|\mathbf{x}_1\|^2 > 0$$

e, se λ_2 é um autovalor negativo de A com autovetor associado $\mathbf{x}_2$, então

$$\mathbf{x}_2^T A\mathbf{x}_2 = \lambda_2\mathbf{x}_2^T\mathbf{x}_2 = \lambda_2\|\mathbf{x}_2\|^2 < 0$$

EXEMPLO 4. Seja $f(x, y) = 2x^2 - 4xy + 5y^2$. Determine a natureza do ponto estacionário (0, 0).

SOLUÇÃO. A matriz A associada à forma quadrática é

$$\begin{pmatrix} 2 & -2\\ -2 & 5\end{pmatrix}$$

Os autovalores são $\lambda_1 = 6$ e $\lambda_2 = 1$. Como ambos os autovalores são positivos, A é positiva definida e, portanto, (0, 0) é um ponto de mínimo global. □

Suponha agora que temos uma função $F(x, y)$ com um ponto estacionário (x_0, y_0). Se F tem derivadas parciais contínuas em uma vizinhança de (x_0, y_0), então ela pode ser expandida em uma série de Taylor em torno desse ponto:

$$\begin{aligned}&F(x_0 + h, y_0 + k)\\ &= F(x_0, y_0) + \left[hF_x(x_0, y_0) + kF_y(x_0, y_0)\right]\\ &\quad + \tfrac{1}{2}\left[h^2F_{xx}(x_0, y_0) + 2hkF_{xy}(x_0, y_0) + k^2F_{yy}(x_0, y_0)\right] + R\\ &= F(x_0, y_0) + \tfrac{1}{2}(ah^2 + 2bhk + ck^2) + R\end{aligned}$$

onde

$$a = F_{xx}(x_0, y_0), \qquad b = F_{xy}(x_0, y_0), \qquad c = F_{yy}(x_0, y_0)$$

e o resto R é dado por

$$\begin{aligned}R &= \tfrac{1}{6}\left[h^3F_{xxx}(\mathbf{z}) + 3h^2kF_{xyy}(\mathbf{z}) + 3hk^2F_{xyy}(\mathbf{z}) + k^3F_{yyy}(\mathbf{z})\right]\\ \mathbf{z} &= (x_0 + \theta h, y_0 + \theta k), \qquad 0 < \theta < 1\end{aligned}$$

Se h e k são suficientemente pequenos, $|R|$ vai ser menor do que o módulo de $\frac{1}{2}\{ah^2 + 2bhk + ck^2\}$ e, portanto, $[F(x_0 + h, y_0 + k) - F(x_0, y_0)]$ vai ter o mesmo sinal que $(ah^2 + 2bhk + ck^2)$. A expressão

$$f(h, k) = ah^2 + 2bhk + ck^2$$

é uma forma quadrática nas variáveis h e k. Logo, $F(x, y)$ vai ter um mínimo (máximo) local em (x_0, y_0) se e somente se $f(h, k)$ tem um mínimo (máximo) em $(0, 0)$. Seja

$$H = \begin{pmatrix} a & b \\ b & c \end{pmatrix} = \begin{pmatrix} F_{xx}(x_0, y_0) & F_{xy}(x_0, y_0) \\ F_{xy}(x_0, y_0) & F_{yy}(x_0, y_0) \end{pmatrix}$$

e sejam λ_1 e λ_2 os autovalores de H. Se H for invertível, então λ_1 e λ_2 são diferentes de zero e podemos classificar os pontos estacionários da seguinte maneira:

(i) F tem um mínimo em (x_0, y_0) se $\lambda_1 > 0$ e $\lambda_2 > 0$;
(ii) F tem um máximo em (x_0, y_0) se $\lambda_1 < 0$ e $\lambda_2 < 0$;
(iii) F tem um ponto de sela em (x_0, y_0) se λ_1 e λ_2 têm sinais diferentes.

EXEMPLO 5. Encontre e classifique todos os pontos estacionários da função

$$F(x, y) = \tfrac{1}{3}x^3 + xy^2 - 4xy + 1$$

SOLUÇÃO. As derivadas parciais de primeira ordem de F são

$$F_x = x^2 + y^2 - 4y$$

$$F_y = 2xy - 4x = 2x(y - 2)$$

TABELA 1

Ponto Estacionário (x_0, y_0)	λ_1	λ_2	Descrição
(0, 0)	4	−4	Ponto de sela
(0, 4)	4	−4	Ponto de sela
(2, 2)	4	4	Mínimo local
(−2, 2)	−4	−4	Máximo local

Igualando F_y a 0, obtemos $x = 0$ ou $y = 2$. Fazendo $F_x = 0$, vemos que, se $x = 0$, então y tem que ser 0 ou 4 e, se $y = 2$, então $x = \pm 2$. Logo, $(0, 0)$, $(0, 4)$, $(2, 2)$ e $(-2, 2)$ são os pontos estacionários de F. Para classificar os pontos estacionários, vamos calcular as derivadas parciais de segunda ordem:

$$F_{xx} = 2x, \qquad F_{xy} = 2y - 4, \qquad F_{yy} = 2x$$

Para cada ponto estacionário (x_0, y_0), determinamos os autovalores de

$$\begin{pmatrix} 2x_0 & 2y_0 - 4 \\ 2y_0 - 4 & 2x_0 \end{pmatrix}$$

A Tabela 1 mostra esses autovalores. □

Podemos generalizar nosso método de classificação de pontos estacionários para funções de mais de duas variáveis. Seja $F(\mathbf{x}) = F(x_1, \ldots, x_n)$ uma função com valores reais e derivadas parciais contínuas de terceira ordem. Seja $\mathbf{x}_0$ um ponto estacionário de F e defina a matriz $H = H(\mathbf{x}_0)$ por

$$h_{ij} = F_{x_i x_j}(\mathbf{x}_0)$$

$H(\mathbf{x}_0)$ é chamada de *matriz hessiana* de F em $\mathbf{x}_0$.

O ponto estacionário pode ser classificado da seguinte maneira:

(i) $\mathbf{x}_0$ é um ponto de mínimo local de F se $H(\mathbf{x}_0)$ é positiva definida;
(ii) $\mathbf{x}_0$ é um ponto de máximo local de F se $H(\mathbf{x}_0)$ é negativa definida;
(iii) $\mathbf{x}_0$ é um ponto de sela de F se $H(\mathbf{x}_0)$ é indefinida.

EXEMPLO 6. Encontre os pontos de mínimo local para a função

$$F(x, y, z) = x^2 + xz - 3\cos y + z^2$$

SOLUÇÃO. As derivadas parciais de primeira ordem de F são

$$F_x = 2x + z$$

$$F_y = 3\,\text{sen}\,y$$

$$F_z = x + 2z$$

Temos que (x, y, z) é um ponto estacionário de F se e somente se $x = z = 0$ e $y = n\pi$, onde n é um inteiro. Seja $\mathbf{x}_0 = (0, 2k\pi, 0)^T$. A matriz hessiana de F em $\mathbf{x}_0$ é dada por

$$H(\mathbf{x}_0) = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$$

Os autovalores de $H(\mathbf{x}_0)$ são 3, 3 e 1. Como os autovalores são todos positivos, $H(\mathbf{x}_0)$ é positiva definida e, portanto, F tem um mínimo local em $\mathbf{x}_0$. Por outro lado, em um ponto estacionário da forma $\mathbf{x}_1 = (0, (2k - 1)\pi, 0)^T$, a matriz hessiana é

$$H(\mathbf{x}_1) = \begin{pmatrix} 2 & 0 & 1 \\ 0 & -3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$$

Os autovalores de $H(\mathbf{x}_1)$ são -3, 3 e 1, logo $H(\mathbf{x}_1)$ é indefinida e $\mathbf{x}_1$ é um ponto de sela de F. □

EXERCÍCIOS

1. Encontre a matriz associada a cada uma das formas quadráticas a seguir.

(a) $3x^2 - 5xy + y^2$
(b) $2x^2 + 3y^2 + z^2 + xy - 2xz + 3yz$
(c) $x^2 + 2y^2 + z^2 + xy - 2xz + 3yz$

2. Reordene os autovalores no Exemplo 2 de modo que $\lambda_1 = 4$ e $\lambda_2 = 2$ e refaça o exemplo. Em que quadrantes vão ficar os semi-eixos positivos de x' e y'? Esboce o gráfico e compare-o com a Fig. 6.5.3.

3. Para cada um dos itens a seguir, encontre uma mudança apropriada de coordenadas (isto é, uma rotação e/ou uma translação), de modo que a cônica resultante esteja em forma canônica; identifique a curva e esboce seu gráfico.

(a) $x^2 + xy + y^2 - 6 = 0$
(b) $3x^2 + 8xy + 3y^2 + 28 = 0$
(c) $-3x^2 + 6xy + 5y^2 - 24 = 0$
(d) $x^2 + 2xy + y^2 + 3x + y - 1 = 0$

4. Sejam λ_1 e λ_2 os autovalores de

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

Que espécie de seção cônica a equação

$$ax^2 + 2bxy + cy^2 = 1$$

vai representar se $\lambda_1\lambda_2 < 0$? Explique.

5. Seja A uma matriz simétrica 2×2 e seja α um escalar não-nulo para o qual a equação $\mathbf{x}^T A\mathbf{x} = \alpha$ é compatível. Mostre que a seção cônica correspondente vai ser não-degenerada se e somente se A for invertível.

6. Quais das matrizes a seguir são positivas definidas? Quais são negativas definidas? Quais são indefinidas?

(a) $\begin{pmatrix} 3 & 2 \\ 2 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 3 & 4 \\ 4 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 3 & \sqrt{2} \\ \sqrt{2} & 4 \end{pmatrix}$

(d) $\begin{pmatrix} -2 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & -2 \end{pmatrix}$ (e) $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ (f) $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 3 \\ 0 & 3 & 5 \end{pmatrix}$

7. Para cada uma das funções a seguir, determine se o ponto estacionário dado é um ponto de mínimo local, de máximo local ou de sela.

(a) $f(x, y) = 3x^2 - xy + y^2 \quad (0, 0)$
(b) $f(x, y) = \operatorname{sen} x + y^3 + 3xy + 2x - 3y \quad (0, -1)$
(c) $f(x, y) = \frac{1}{3}x^3 - \frac{1}{3}y^3 + 3xy + 2x - 2y \quad (1, -1)$
(d) $f(x, y) = \dfrac{y}{x^2} + \dfrac{x}{y^2} + xy \quad (1, 1)$
(e) $f(x, y, z) = x^3 + xyz + y^2 - 3x \quad (1, 0, 0)$
(f) $f(x, y, z) = -\frac{1}{4}(x^{-4} + y^{-4} + z^{-4}) + yz - x - 2y - 2z \quad (1, 1, 1)$

8. Mostre que, se A é positiva definida, então $\det(A) > 0$. Dê um exemplo de uma matriz 2×2 com determinante positivo que não é positiva definida.

9. Mostre que, se A é uma matriz simétrica positiva definida, então A é invertível e A^{-1} também é positiva definida.

10. Seja A uma matriz singular $n \times n$. Mostre que A^TA é positiva semidefinida, mas não é positiva definida.

11. Seja A uma matriz simétrica $n \times n$ com autovalores $\lambda_1, \ldots, \lambda_n$. Mostre que existe um conjunto ortonormal de vetores $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ tal que

$$\mathbf{x}^T A\mathbf{x} = \sum_{i=1}^{n} \lambda_i \left(\mathbf{x}^T \mathbf{x}_i\right)^2$$

para todo $\mathbf{x} \in R^n$.

12. Seja A uma matriz simétrica positiva definida. Mostre que os elementos diagonais de A têm que ser todos positivos.

13. Seja A uma matriz simétrica positiva definida $n \times n$ e seja S uma matriz invertível $n \times n$. Mostre que S^TAS é positiva definida.

14. Seja A uma matriz simétrica positiva definida $n \times n$. Mostre que A pode ser fatorada em um produto QQ^T, onde Q é uma matriz $n \times n$ cujas colunas são ortogonais duas a duas.

[***Sugestão:*** ver Corolário 6.4.5.]

6 MATRIZES POSITIVAS DEFINIDAS

Vimos na Seção 5 que uma matriz simétrica é positiva definida se e somente se seus autovalores são todos positivos. Tais matrizes aparecem em uma grande variedade de aplicações. Elas aparecem com freqüência na solução numérica de problemas de valor inicial pelo método de diferenças finitas ou pelo método de elementos finitos. Devido à sua importância em matemática aplicada, devotamos esta seção ao estudo de suas propriedades.

Lembre-se de que uma matriz simétrica A $n \times n$ é positiva definida se $\mathbf{x}^T A\mathbf{x} > 0$ para todos os vetores não-nulos $\mathbf{x}$ em R^n. No Teorema 6.5.1, as matrizes simétricas positivas definidas foram caracterizadas pela condição de que todos os seus autovalores são positivos. Essa caracterização pode ser usada para se provar as propriedades a seguir.

I. Se A é uma matriz simétrica positiva definida, então A é invertível.
II. Se A é uma matriz simétrica positiva definida, então $\det(A) > 0$.

Se A fosse singular, $\lambda = 0$ seria um autovalor. Como todos os autovalores de A são positivos, A tem que ser invertível. A segunda propriedade também é conseqüência do Teorema 6.5.1, pois

$$\det(A) = \lambda_1 \cdots \lambda_n > 0$$

Dada uma matriz A $n \times n$, vamos denotar por A_r a matriz obtida retirando-se as últimas $n - r$ linhas e colunas de A. A_r é chamada a *submatriz principal* de A de ordem r. Podemos agora enunciar a terceira propriedade de matrizes positivas definidas.

III. Se A é uma matriz simétrica positiva definida, então as submatrizes principais $A_1, ..., A_n$ de A são todas positivas definidas.

Demonstração. Para mostrar que A_r é positiva definida, $1 \le r \le n$, seja $\mathbf{x}_r = (x_1, ..., x_r)^T$ um vetor não-nulo arbitrário em R^r e seja $\mathbf{x} = (x_1, ..., x_r, 0, ..., 0)^T$. Como

$$\mathbf{x}_r^T A_r \mathbf{x}_r = \mathbf{x}^T A \mathbf{x} > 0$$

A_r é positiva definida. □

Uma conseqüência imediata das propriedades **I**, **II** e **III** é que, se A_r é uma submatriz principal de uma matriz simétrica positiva definida A, então A_r é invertível e $\det(A_r) > 0$. Isso é importante no método de Gauss. Em geral, se A é uma matriz $n \times n$ cujas submatrizes principais são todas invertíveis, então A pode ser reduzida a uma forma triangular superior usando-se apenas a operação elementar III; em outras palavras, como os elementos diagonais nunca serão iguais a zero ao se usar o método de Gauss, o processo pode ser completado sem haver troca de linhas.

IV. Se A é uma matriz simétrica positiva definida, então A pode ser reduzida a uma forma triangular superior usando-se apenas a operação elementar III e os pivôs serão todos positivos.

$$\underbrace{\begin{pmatrix} a_{11} & \times & \times & \times \\ \times & a_{22} & \times & \times \\ \times & \times & a_{33} & \times \\ \times & \times & \times & a_{44} \end{pmatrix}}_{A} \xrightarrow{1} \underbrace{\begin{pmatrix} a_{11} & \times & \times & \times \\ 0 & a_{22}^{(1)} & \times & \times \\ 0 & \times & a_{33}^{(1)} & \times \\ 0 & \times & \times & a_{44}^{(1)} \end{pmatrix}}_{A^{(1)}} \xrightarrow{2} \underbrace{\begin{pmatrix} a_{11} & \times & \times & \times \\ 0 & a_{22}^{(1)} & \times & \times \\ 0 & 0 & a_{33}^{(2)} & \times \\ 0 & 0 & \times & a_{44}^{(2)} \end{pmatrix}}_{A^{(2)}} \xrightarrow{3} \underbrace{\begin{pmatrix} a_{11} & \times & \times & \times \\ 0 & a_{22}^{(1)} & \times & \times \\ 0 & 0 & a_{33}^{(2)} & \times \\ 0 & 0 & 0 & a_{44}^{(3)} \end{pmatrix}}_{A^{(3)} = U}$$

FIG. 6.6.1

Vamos ilustrar a propriedade **IV** no caso de uma matriz simétrica positiva definida 4×4. Em primeiro lugar,

$$a_{11} = \det(A_1) > 0$$

de modo que a_{11} pode ser usado como pivô e a primeira linha é a linha do pivô. Vamos denotar por $a_{22}^{(1)}$ o elemento na posição (2, 2) após os três últimos elementos da primeira coluna terem ficado nulos (ver Fig. 6.6.1). Nessa etapa, a submatriz A_2 foi reduzida a

$$\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22}^{(1)} \end{pmatrix}$$

Como isso foi feito usando-se apenas a operação elementar III, o determinante não se alterou, logo

$$\det(A_2) = a_{11}a_{22}^{(1)}$$

e, portanto,

$$a_{22}^{(1)} = \frac{\det(A_2)}{a_{11}} = \frac{\det(A_2)}{\det(A_1)} > 0$$

Como $a_{22}^{(1)} \neq 0$, esse elemento pode ser usado como pivô na segunda etapa do método de Gauss. Após essa segunda etapa, a matriz A_3 foi transformada em

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & 0 & a_{33}^{(2)} \end{pmatrix}$$

Como foi usada apenas a operação elementar III,

$$\det(A_3) = a_{11}a_{22}^{(1)}a_{33}^{(2)}$$

e, portanto,

$$a_{33}^{(2)} = \frac{\det(A_3)}{a_{11}a_{22}^{(1)}} = \frac{\det(A_3)}{\det(A_2)} > 0$$

Logo, $a_{33}^{(2)}$ pode ser usado como pivô na última etapa. Após essa etapa, o elemento restante na diagonal será

$$a_{44}^{(3)} = \frac{\det(A_4)}{\det(A_3)} > 0$$

Em geral, se uma matriz A $n \times n$ puder ser colocada em forma triangular superior U sem troca de linhas, então A pode ser fatorada em um produto LU, onde L é triangular inferior e tem todos os elementos da diagonal iguais a 1. O elemento (i, j) de L abaixo da diagonal principal vai ser o múltiplo da i-ésima linha que foi subtraído da j-ésima linha durante o método de Gauss. Vamos ilustrar isso com um exemplo 3×3.

EXEMPLO 1. Seja

$$A = \begin{pmatrix} 4 & 2 & -2 \\ 2 & 10 & 2 \\ -2 & 2 & 5 \end{pmatrix}$$

A matriz L é determinada da seguinte maneira. Na primeira etapa do método de Gauss, subtraímos $\frac{1}{2}$ vez a primeira linha da segunda e subtraímos $-\frac{1}{2}$ vez a primeira linha da terceira. Correspondendo a essas operação, fazemos $l_{21} = \frac{1}{2}$ e $l_{31} = -\frac{1}{2}$. Após a primeira etapa, obtemos a matriz

$$A^{(1)} = \begin{pmatrix} 4 & 2 & -2 \\ 0 & 9 & 3 \\ 0 & 3 & 4 \end{pmatrix}$$

Na etapa final subtraímos $\frac{1}{3}$ vezes a segunda linha da terceira. Correspondendo a essa etapa, fazemos $l_{32} = \frac{1}{3}$. Terminamos com uma matriz triangular superior

$$U = A^{(2)} = \begin{pmatrix} 4 & 2 & -2 \\ 0 & 9 & 3 \\ 0 & 0 & 3 \end{pmatrix}$$

A matriz L é dada por

$$L = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & \frac{1}{3} & 1 \end{pmatrix}$$

e podemos verificar que $LU = A$.

$$\begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & \frac{1}{3} & 1 \end{pmatrix} \begin{pmatrix} 4 & 2 & -2 \\ 0 & 9 & 3 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 2 & -2 \\ 2 & 10 & 2 \\ -2 & 2 & 5 \end{pmatrix}$$

Para ver por que essa fatoração funciona, vamos analisar o processo em termos de matrizes elementares. A operação elementar III foi utilizada três vezes durante o processo. Isso é equivalente a multiplicar A à esquerda por três matrizes elementares E_1, E_2, E_3. Então, $E_3E_2E_1A = U$.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{3} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & 2 & -2 \\ 2 & 10 & 2 \\ -2 & 2 & 5 \end{pmatrix}$$
$$= \begin{pmatrix} 4 & 2 & -2 \\ 0 & 9 & 3 \\ 0 & 0 & 3 \end{pmatrix}$$

Como as matrizes elementares são invertíveis, temos

$$A = (E_1^{-1}E_2^{-1}E_3^{-1})U$$

Quando as inversas das matrizes elementares são multiplicadas nessa ordem, o resultado é uma matriz L triangular inferior com todos os elementos da diagonal iguais a 1. Os elementos abaixo da diagonal são, simplesmente, os múltiplos que foram subtraídos durante o processo.

$$E_1^{-1}E_2^{-1}E_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{3} & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & \frac{1}{3} & 1 \end{pmatrix}$$

□

Dada uma fatoração LU de uma matriz A, é possível prosseguir e fatorar U em um produto DU_1, onde D é diagonal e U_1 é triangular superior com todos os elementos diagonais iguais a 1.

$$DU_1 = \begin{pmatrix} u_{11} & & & \\ & u_{22} & & \\ & & \ddots & \\ & & & u_{nn} \end{pmatrix} \begin{pmatrix} 1 & \frac{u_{12}}{u_{11}} & \frac{u_{13}}{u_{11}} & \cdots & \frac{u_{1n}}{u_{11}} \\ & 1 & \frac{u_{23}}{u_{22}} & \cdots & \frac{u_{2n}}{u_{22}} \\ & & & & \vdots \\ & & & & 1 \end{pmatrix}$$

Temos, então, que $A = LDU_1$. Em geral, se A puder ser fatorada em um produto da forma LDU com L triangular inferior, D diagonal, U triangular superior e L e U tendo todos os elementos diagonais iguais a 1, então essa fatoração é única (ver Exercício 6).

Se A é uma matriz simétrica positiva definida, então A pode ser fatorada em um produto $LU = LDU_1$. Os elementos diagonais de D são os elementos $u_{11}, \ldots, u_{nn}$ que foram os pivôs durante o método de Gauss. Pela propriedade IV, todos esses elementos são positivos. Além disso, como A é simétrica,

$$LDU_1 = A = A^T = (LDU_1)^T = U_1^T D^T L^T$$

Pela singularidade da fatoração LDU, temos que $L^T = U_1$. Logo,

$$A = LDL^T$$

Essa fatoração importante é utilizada muitas vezes em cálculos numéricos. Existem algoritmos eficientes que usam essa fatoração na resolução de sistemas lineares simétricos positivos definidos.

V. Se A é uma matriz simétrica positiva definida, então A pode ser fatorada em um produto LDL^T, onde L é triangular inferior com todos os elementos diagonais iguais a 1 e D é diagonal com todos os elementos diagonais positivos.

EXEMPLO 2. Vimos, no Exemplo 1, que

$$A = \begin{pmatrix} 4 & 2 & -2 \\ 2 & 10 & 2 \\ -2 & 2 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & \frac{1}{3} & 1 \end{pmatrix} \begin{pmatrix} 4 & 2 & -2 \\ 0 & 9 & 3 \\ 0 & 0 & 3 \end{pmatrix} = LU$$

Fatorando U, obtemos

$$A = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & \frac{1}{3} & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 1 \end{pmatrix} = LDL^T$$

□

Como os elementos diagonais $u_{11}, \ldots, u_{nn}$ são todos positivos, é possível prosseguir com a fatoração. Seja

$$D^{1/2} = \begin{pmatrix} \sqrt{u_{11}} & & & \\ & \sqrt{u_{22}} & & \\ & & \ddots & \\ & & & \sqrt{u_{nn}} \end{pmatrix}$$

e defina $L_1 = LD^{1/2}$. Então

$$A = LDL^T = LD^{1/2}(D^{1/2})^T L^T = L_1 L_1^T$$

Essa fatoração é conhecida como a *decomposição de Cholesky* de A.

VI. (Decomposição de Cholesky) Se A é uma matriz simétrica positiva definida, então A pode ser fatorada em um produto LL^T, onde L é triangular inferior com elementos diagonais positivos.

EXEMPLO 3. Seja A a matriz dos Exemplos 1 e 2. Definindo

$$L_1 = LD^{1/2} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & \frac{1}{3} & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & \sqrt{3} \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ -1 & 1 & \sqrt{3} \end{pmatrix}$$

temos

$$L_1 L_1^T = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ -1 & 1 & \sqrt{3} \end{pmatrix} \begin{pmatrix} 2 & 1 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & \sqrt{3} \end{pmatrix}$$

$$= \begin{pmatrix} 4 & 2 & -2 \\ 2 & 10 & 2 \\ -2 & 2 & 5 \end{pmatrix} = A$$

□

A matriz $A = LL^T$ também pode ser escrita em termos da matriz triangular superior $R = L^T$. De fato, se $R = L^T$, então $A = LL^T = R^T R$. Por outro lado, não é difícil mostrar que qualquer produto $B^T B$ é positivo definido desde que B seja invertível. Colocando todos esses resultados juntos, temos o teorema a seguir.

Teorema 6.6.1. *Seja A uma matriz simétrica $n \times n$. As seguintes afirmações são equivalentes:*
(a) *A é positiva definida;*
(b) *Todas as submatrizes principais $A_1, \ldots, A_n$ têm determinante positivo;*
(c) *A pode ser reduzida por linhas a uma forma triangular superior usando apenas operações elementares do tipo III com todos os pivôs positivos;*
(d) *A tem uma decomposição de Cholesky LL^T (onde L é triangular inferior com todos os elementos diagonais positivos);*
(e) *A pode ser fatorada em um produto $B^T B$ para alguma matriz invertível B.*

Demonstração. Já mostramos que (a) implica (b), (b) implica (c) e (c) implica (d). Para ver que (d) implica (e), suponha que $A = LL^T$. Definindo $B = L^T$, temos que B é invertível e

$$A = LL^T = B^T B$$

Finalmente, para mostrar que (e) $\rightarrow$ (a), suponha que $A = B^T B$, onde B é invertível. Seja $\mathbf{x}$ um vetor não-nulo arbitrário em R^n e seja $\mathbf{y} = B\mathbf{x}$. Como B é invertível, $\mathbf{y} \neq \mathbf{0}$ e

$$\mathbf{x}^T A \mathbf{x} = \mathbf{x}^T B^T B \mathbf{x} = \mathbf{y}^T \mathbf{y} = \|\mathbf{y}\|^2 > 0$$

Logo, A é positiva definida.

□

Resultados análogos ao Teorema 6.6.1 não são válidos para matrizes positivas semidefinidas. Por exemplo, considere a matriz

$$A = \begin{pmatrix} 1 & 1 & -3 \\ 1 & 1 & -3 \\ -3 & -3 & 5 \end{pmatrix}$$

Todas as submatrizes principais têm determinante não-negativo

$$\det(A_1) = 1, \qquad \det(A_2) = 0, \qquad \det(A_3) = 0$$

mas A não é positiva semidefinida, pois tem um autovalor negativo $\lambda = -1$. De fato, $\mathbf{x} = (1, 1, 1)^T$ é um autovetor associado a $\lambda = -1$ e

$$\mathbf{x}^T A \mathbf{x} = -3$$

EXERCÍCIOS

1. Para cada item a seguir, calcule o determinante de todas as submatrizes principais e use-os para determinar se a matriz é ou não positiva definida.

(a) $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 3 & 4 \\ 4 & 2 \end{pmatrix}$

(c) $\begin{pmatrix} 6 & 4 & -2 \\ 4 & 5 & 3 \\ -2 & 3 & 6 \end{pmatrix}$ (d) $\begin{pmatrix} 4 & 2 & 1 \\ 2 & 3 & -2 \\ 1 & -2 & 5 \end{pmatrix}$

2. Seja A uma matriz simétrica positiva definida 3×3 e suponha que $\det(A_1) = 3$, $\det(A_2) = 6$ e $\det(A_3) = 8$. Quais seriam os pivôs no processo de redução de A a uma forma triangular superior, supondo que se utiliza apenas a operação elementar III no processo?

3. Para cada item a seguir, fatore a matriz dada em um produto LDL^T, onde L é triangular inferior com todos os elementos diagonais iguais a 1 e D é uma matriz diagonal.

(a) $\begin{pmatrix} 4 & 2 \\ 2 & 10 \end{pmatrix}$ (b) $\begin{pmatrix} 9 & -3 \\ -3 & 2 \end{pmatrix}$

(c) $\begin{pmatrix} 16 & 8 & 4 \\ 8 & 6 & 0 \\ 4 & 0 & 7 \end{pmatrix}$ (d) $\begin{pmatrix} 9 & 3 & -6 \\ 3 & 4 & 1 \\ -6 & 1 & 9 \end{pmatrix}$

4. Encontre a decomposição de Cholesky LL^T para cada uma das matrizes no Exercício 3.

5. Seja A uma matriz simétrica positiva definida $n \times n$. Para cada par $\mathbf{x}, \mathbf{y} \in R^n$, defina

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A \mathbf{y}$$

Mostre que $<\, , >$ define um produto interno em R^n.

6. Seja A uma matriz invertível $n \times n$ e suponha que $A = L_1 D_1 U_1 = L_2 D_2 U_2$, onde L_1 e L_2 são triangulares inferiores, D_1 e D_2 são diagonais, U_1 e U_2 são triangulares superiores, e L_1, L_2, U_1, U_2 têm todos os elementos diagonais iguais a 1. Mostre que $L_1 = L_2$, $D_1 = D_2$ e $U_1 = U_2$.

[***Sugestão:*** L_2^{-1} é triangular inferior e U_1^{-1} é triangular superior; compare ambos os lados da equação $D_2^{-1} L_2^{-1} L_1 D_1 = U_2 U_1^{-1}$.]

7. Seja A uma matriz simétrica positiva definida e seja Q uma matriz ortogonal que diagonaliza A. Use a fatoração $A = QDQ^T$ para encontrar uma matriz invertível B tal que $B^T B = A$.

8. Seja B uma matriz $m \times n$ de posto n. Mostre que $B^T B$ é positiva definida.

9. Seja A uma matriz simétrica $n \times n$. Mostre que e^A é simétrica e positiva definida.

10. Mostre que, se B é uma matriz invertível simétrica, então B^2 é positiva definida.

11. Sejam

$$A = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

(a) Mostre que A é positiva definida e que $\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T B\mathbf{x}$ para todo $\mathbf{x} \in R^2$.
(b) Mostre que B é positiva definida mas que B^2 não é positiva definida.

12. Seja A uma matriz simétrica negativa definida $n \times n$.
(a) Qual o sinal do det(A) se n for par? E se n for ímpar?
(b) Mostre que as submatrizes principais de A são negativas definidas.
(c) Mostre que os determinantes das submatrizes principais têm sinais alternados.

13. Seja A uma matriz simétrica positiva definida $n \times n$.
(a) Se $k < n$, as submatrizes principais A_k e A_{k+1} são ambas positivas definidas e, portanto, têm decomposições de Cholesky $L_k L_k^T$ e $L_{k+1}L_{k+1}^T$. Se A_{k+1} puder ser colocada na forma

$$A_{k+1} = \begin{pmatrix} A_k & \mathbf{y}_k \\ \mathbf{y}_k^T & \beta_k \end{pmatrix}$$

onde $\mathbf{y}_k \in R^k$ e β_k é um escalar, mostre que L_{k+1} é da forma

$$L_{k+1} = \begin{pmatrix} L_k & \mathbf{0} \\ \mathbf{x}_k^T & \alpha_k \end{pmatrix}$$

e determine $\mathbf{x}_k$ e α_k em função de L_k, $\mathbf{y}_k$ e β_k.

(b) A submatriz principal A_1 tem decomposição de Cholesky $L_1 L_1^T$, onde $L_1 = (\sqrt{a_{11}})$. Explique como o item (a) pode ser usado para se calcular sucessivamente as decomposições de Cholesky de $A_2, \ldots, A_n$. Escreva um algoritmo que calcula, em um único ciclo (*loop*), $L_2, L_3, \ldots, L_n$. Como $A = A_n$, a decomposição de Cholesky de A vai ser $L_n L_n^T$. (Esse algoritmo é eficiente no sentido de que usa aproximadamente metade dos cálculos aritméticos necessários, em geral, para calcular uma fatoração LU.)

7 MATRIZES NÃO-NEGATIVAS

Em muitos tipos de sistemas lineares que aparecem nas aplicações, os elementos da matriz de coeficientes representam quantidades não-negativas. Esta seção trata do estudo de tais matrizes e algumas de suas propriedades.

Definição. Uma matriz A $n \times n$ com elementos reais é dita **não-negativa** se $a_{ij} \geq 0$ para todo i e j e é dita **positiva** se $a_{ij} > 0$ para todo i e j.

Analogamente, um vetor $\mathbf{x} = (x_1, \ldots, x_n)^T$ é **não-negativo** se $x_i \geq 0$ para todo i e **positivo** se $x_i > 0$ para todo i.

Para exemplificar uma das aplicações de matrizes não-negativas, vamos considerar os modelos econômicos de Leontief.

APLICAÇÃO 1: O MODELO ABERTO

Suponha que existem n indústrias produzindo n produtos. Cada indústria necessita de produtos de outras indústrias e, possivelmente, de seus próprios produtos. No modelo aberto supõe-se que haja uma demanda adicional externa para cada produto. O problema é determinar a produção necessária de cada indústria para atender à demanda total.

Vamos mostrar que esse problema pode ser representado por um sistema de equações lineares e que o sistema tem uma única solução não-negativa. Vamos denotar por a_{ij} a quantidade de insumos da i-ésima indústria necessária para produzir uma unidade da indústria j. Por unidade de insumo ou de produção queremos dizer a quantidade que corresponde ao valor de um real. Então, a produção correspondente a um real do j-ésimo produto envolve um custo de $\sum_{i=1}^{n} a_{ij}$ reais. É claro que a produção do j-ésimo produto não dará lucro, a menos que $\sum_{i=1}^{n} a_{ij} < 1$. Vamos denotar por d_i a demanda no setor aberto para o i-ésimo produto. Finalmente, vamos denotar por x_i a quantidade necessária do i-ésimo produto para atender a demanda total. Para que a j-ésima indústria produza x_j, ela vai precisar de $a_{ij}x_j$ unidades da i-ésima indústria. Então, a demanda total para o i-ésimo produto vai ser

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + d_i$$

e, portanto, precisamos que

$$x_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + d_i$$

para $i = 1, ..., n$. Isso nos leva ao sistema

$$\begin{array}{rcrcccrcl} (1 - a_{11})x_1 & + & (-a_{12})x_2 & + & \cdots & + & (-a_{1n})x_n & = & d_1 \\ (-a_{21})x_1 & + & (1 - a_{22})x_2 & + & \cdots & + & (-a_{2n})x_n & = & d_2 \\ \vdots & & & & & & & & \\ (-a_{n1})x_1 & + & (-a_{n2})x_2 & + & \cdots & + & (1 - a_{nn})x_n & = & d_n \end{array}$$

que pode ser escrito na forma

(1) $$(I - A)\mathbf{x} = \mathbf{d}$$

Os elementos de A têm duas propriedades importantes:

(i) $a_{ij} \geq 0$ para cada i e cada j

(ii) $\sum_{i=1}^{n} a_{ij} < 1$

Além de ser solução do sistema (1), o vetor $\mathbf{x}$ também precisa ser não-negativo. (Não faria sentido ter uma produção negativa.)

Para mostrar que o sistema tem uma única solução não-negativa, vamos precisar usar uma nova norma para matrizes, chamada *norma* 1 ou *norma* l_1, e denotada por $\|\cdot\|_1$. A definição e as propriedades dessa norma serão estudadas na Seção 4 do Cap. 7. Nesta seção, vamos mostrar que, qualquer que seja a matriz B $m \times n$,

(2) $$\|B\|_1 = \max_{1 \leq j \leq n} \left(\sum_{i=1}^{m} |b_{ij}| \right)$$

Vamos mostrar, também, que a norma 1 satisfaz as seguintes propriedades multiplicativas:

(3) $$\begin{array}{rcll} \|BC\|_1 & \leq & \|B\|_1 \|C\|_1 & \text{para qualquer matriz } C \in R^{n \times r} \\ \|B\mathbf{x}\|_1 & \leq & \|B\|_1 \|\mathbf{x}\|_1 & \text{para todo } \mathbf{x} \in R^n \end{array}$$

Em particular, se A é uma matriz $n \times n$ satisfazendo as condições (i) e (ii), segue de (2) que $\|A\|_1 < 1$. Além disso, se λ é um autovalor de A com autovetor associado $\mathbf{x}$, então

$$|\lambda|\|\mathbf{x}\|_1 = \|\lambda\mathbf{x}\|_1 = \|A\mathbf{x}\|_1 \leq \|A\|_1\|\mathbf{x}\|_1$$

e, portanto,

$$|\lambda| \leq \|A\|_1 < 1$$

Logo, 1 não é um autovalor de A. Então $I - A$ é invertível e o sistema (1) tem uma única solução

$$\mathbf{x} = (I - A)^{-1}\mathbf{d}$$

Gostaríamos de mostrar que essa solução tem que ser não-negativa. Para fazer isso, vamos mostrar que $(I - A)^{-1}$ é não-negativa. Em primeiro lugar, observe que, em conseqüência da propriedade multiplicativa (3), temos

$$\|A^m\|_1 \leq \|A\|_1^m$$

Como $\|A\|_1 < 1$, obtemos

$$\|A^m\|_1 \to 0 \quad \text{quando} \quad m \to \infty$$

e, portanto, A^m tende à matriz nula quando $m \to \infty$.

Como

$$(I - A)(I + A + \cdots + A^m) = I - A^{m+1}$$

segue que

$$I + A + \cdots + A^m = (I - A)^{-1} - (I - A)^{-1}A^{m+1}$$

e, portanto, $I + A + \ldots + A^m$ tende a $(I - A)^{-1}$ quando $m \to \infty$. Pela condição (i), $I + A + \ldots + A^m$ é não-negativa para todo m e, portanto, $(I - A)^{-1}$ tem que ser não-negativa. Como $\mathbf{d}$ é não-negativo, a solução $\mathbf{x}$ tem que ser não-negativa. Vemos, então, que as condições (i) e (ii) garantem que o sistema (1) tem uma única solução não-negativa $\mathbf{x}$.

Como você já deve ter desconfiado, existe também uma versão fechada do modelo de Leontief. Na versão fechada, supõe-se que cada indústria deve produzir o suficiente para as necessidades de insumo das outras indústrias e para suas próprias necessidades. O setor aberto é ignorado. Então, em vez do sistema (1), temos

$$(I - A)\mathbf{x} = \mathbf{0}$$

e necessitamos que $\mathbf{x}$ seja uma solução positiva. A existência de tal $\mathbf{x}$ nesse caso é um resultado mais profundo do que no modelo aberto e necessita de resultados mais avançados.

Teorema 6.7.1 (Perron). *Se A é uma matriz positiva $n \times n$, então A tem um autovalor real positivo r com as seguintes propriedades:*

(i) *r é uma raiz simples da equação característica;*
(ii) *r tem um autovetor positivo associado $\mathbf{x}$;*
(iii) *se λ é um outro autovalor qualquer de A, então $|\lambda| < r$.*

O teorema de Perron pode ser considerado um caso particular de um teorema mais geral de Frobenius. O teorema de Frobenius se aplica a matrizes não-negativas "irredutíveis".

Definição. Uma matriz não-negativa é dita **redutível** se existir uma partição do conjunto de índices $\{1, 2, \ldots, n\}$ em conjuntos disjuntos não-vazios I_1 e I_2 tal que $a_{ij} = 0$ sempre que $i \in I_1$ e $j \in I_2$. Caso contrário, A é dita **irredutível**.

EXEMPLO 1. Seja A uma matriz da forma

$$\begin{pmatrix} \times & \times & 0 & 0 & \times \\ \times & \times & 0 & 0 & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & 0 & 0 & \times \end{pmatrix}$$

Sejam $I_1 = \{1, 2, 5\}$ e $I_2 = \{3, 4\}$. Então $I_1 \cup I_2 = \{1, 2, 3, 4, 5\}$ e $a_{ij} = 0$ sempre que $i \in I_1$ e $j \in I_2$. Logo, A é redutível. Observe que, se P é a matriz de permutação que corresponde à troca de ordem entre as terceira e quinta linhas da matriz identidade I, então

$$PA = \begin{pmatrix} \times & \times & 0 & 0 & \times \\ \times & \times & 0 & 0 & \times \\ \times & \times & 0 & 0 & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{pmatrix}$$

e

$$PAP^T = \left(\begin{array}{ccc|cc} \times & \times & \times & 0 & 0 \\ \times & \times & \times & 0 & 0 \\ \times & \times & \times & 0 & 0 \\ \hline \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{array}\right)$$

Em geral, é possível mostrar que uma matriz A $n \times n$ é redutível se e somente se existe uma matriz de permutação P tal que PAP^T é da forma

$$\left(\begin{array}{c|c} B & O \\ \hline X & C \end{array}\right)$$

onde B e C são matrizes quadradas. □

Teorema 6.7.2 (Frobenius). *Se A é uma matriz não-negativa irredutível, então A tem um autovalor real positivo r com as seguintes propriedades:*

(i) *r tem um autovetor associado positivo* $\mathbf{x}$*;*

(ii) *se λ é um outro autovalor qualquer de A, então $|\lambda| \le r$. Os autovalores de módulo igual a r são as raízes simples da equação característica. De fato, se existirem m autovalores de módulo igual a r, eles têm que ser da forma*

$$\lambda_k = r\left[\exp\left(\frac{2k\pi i}{m}\right)\right] \qquad k = 0, 1, \ldots, m-1$$

A demonstração desse teorema está além do escopo deste livro. O leitor interessado deve consultar o livro de Gantmacher [2, Vol. 2]. O teorema de Perron é um caso particular do teorema de Frobenius.

APLICAÇÃO 2: O MODELO FECHADO

No modelo fechado de Leontief, supomos que não existe demanda do setor aberto e queremos encontrar a produção que satisfaça a demanda de todas as n indústrias. Definindo, então, x_i e a_{ij} como no modelo aberto, temos

$$x_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$$

para $i = 1, \ldots, n$. O sistema resultante pode ser escrito na forma

(4) $$(A - I)\mathbf{x} = \mathbf{0}$$

Como anteriormente, temos a condição

(i) $$a_{ij} \ge 0$$

Como não existe setor aberto, a produção da j-ésima indústria deve ser a mesma que os insumos totais da indústria. Logo,

$$x_j = \sum_{i=1}^{n} a_{ij} x_j$$

e, portanto, temos nossa segunda condição

$$\text{(ii)} \qquad \sum_{i=1}^{n} a_{ij} = 1 \qquad j = 1, \ldots, n$$

A condição (ii) implica que $A - I$ é singular, já que a soma de seus vetores linhas é $\mathbf{0}$. Então 1 é um autovalor de A e, como $\|A\|_1 = 1$, todos os autovalores de A têm módulo menor ou igual a 1. Vamos supor que um número suficiente dos coeficientes de A seja diferente de zero, de modo que A seja irredutível. Então, pelo Teorema 6.7.2, $\lambda = 1$ tem um autovetor associado positivo $\mathbf{x}$. Logo, qualquer múltiplo positivo de $\mathbf{x}$ vai ser uma solução positiva de (4).

EXERCÍCIOS

1. Encontre os autovalores de cada uma das matrizes a seguir e verifique que as condições (i), (ii) e (iii) do Teorema 6.7.1 são válidas.

(a) $\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 4 & 2 \\ 2 & 7 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 2 & 4 \\ 2 & 4 & 1 \\ 1 & 2 & 4 \end{pmatrix}$

2. Encontre os autovalores de cada uma das matrizes a seguir e verifique que as condições (i) e (ii) do Teorema 6.7.2 são válidas.

(a) $\begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$ (c) $\begin{pmatrix} 0 & 0 & 8 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$

3. Encontre o vetor solução $\mathbf{x}$ para a versão aberta do modelo de Leontief se

$$A = \begin{pmatrix} 0{,}2 & 0{,}4 & 0{,}4 \\ 0{,}4 & 0{,}2 & 0{,}2 \\ 0{,}0 & 0{,}2 & 0{,}2 \end{pmatrix} \qquad \text{e} \qquad \mathbf{d} = \begin{pmatrix} 16.000 \\ 8.000 \\ 24.000 \end{pmatrix}$$

4. Considere a versão fechada do modelo de Leontief com matriz

$$A = \begin{pmatrix} 0{,}5 & 0{,}4 & 0{,}1 \\ 0{,}5 & 0{,}0 & 0{,}5 \\ 0{,}0 & 0{,}6 & 0{,}4 \end{pmatrix}$$

Se $\mathbf{x} = (x_1, x_2, x_3)^T$ é qualquer vetor solução para esse modelo, qual a relação entre as coordenadas x_1, x_2 e x_3?

5. Demonstre: se $A^m = O$ para algum inteiro positivo m, então $I - A$ é invertível.

6. Seja

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & -1 & 1 \end{pmatrix}$$

(a) Calcule $(I - A)^{-1}$.
(b) Calcule A^2 e A^3. Verifique que $(I - A)^{-1} = I + A + A^2$.

7. Quais das matrizes a seguir são redutíveis? Para cada matriz redutível, encontre uma matriz de permutação P tal que PAP^T seja da forma

$$\left(\begin{array}{c|c} B & O \\ \hline X & C \end{array}\right)$$

onde B e C são matrizes quadradas.

(a) $\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$ (d) $\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}$

8. Seja A uma matriz 3×3 não-negativa irredutível cujos autovalores satisfazem $\lambda_1 = 2 = |\lambda_2| = |\lambda_3|$. Determine λ_2 e λ_3.

9. Seja

$$A = \left(\begin{array}{c|c} B & O \\ \hline O & C \end{array}\right)$$

onde B e C são matrizes quadradas.

(a) Se λ é um autovalor de B com autovetor associado $\mathbf{x} = (x_1, ..., x_k)^T$, mostre que λ também é um autovalor de A com autovetor associado $\tilde{\mathbf{x}} = (x_1, ..., x_k, 0, ..., 0)^T$.

(b) Se B e C são matrizes positivas, mostre que A tem um autovalor real positivo r com a propriedade de que $|\lambda| < r$, qualquer que seja o autovalor $\lambda \neq r$. Mostre, também, que a multiplicidade de r é no máximo 2 e que r tem um autovetor associado não-negativo.

(c) Se $B = C$, mostre que o autovalor r do item (b) tem multiplicidade 2 e tem um autovetor positivo associado.

10. Prove que uma matriz A 2×2 é redutível se e somente se $a_{12}a_{21} = 0$.

11. Prove o teorema de Frobenius no caso em que A é uma matriz 2×2.

EXERCÍCIOS COM O MATLAB PARA O CAPÍTULO 6

MATRIZES DIAGONALIZÁVEIS E NÃO-DIAGONALIZÁVEIS

1. Construa uma matriz simétrica A fazendo

$$A = \texttt{round}(5 * \texttt{rand}(6)); \qquad A = A + A'$$

Calcule os autovalores de A fazendo $\mathbf{e} = \texttt{eig}(A)$.

(a) O traço de A pode ser calculado usando-se o comando do MATLAB `trace`(A) e a soma dos autovalores de A pode ser calculada através do comando `sum`($\mathbf{e}$). Calcule ambas as quantidades e compare os resultados. Use o comando `prod`($\mathbf{e}$) para calcular o produto dos autovalores de A e compare o resultado com `det`(A).

(b) Calcule os autovetores de A fazendo $[X, D] = \texttt{eig}(A)$. Use o MATLAB para calcular $X^{-1}AX$ e compare o resultado com D. Calcule, também, A^{-1} e $XD^{-1}X^{-1}$ e compare os resultados.

2. Defina

$$A = \texttt{ones}(10) + \texttt{eye}(10)$$

(a) Qual o posto de $A - I$? Por que $\lambda = 1$ tem que ser um autovalor de multiplicidade nove? Calcule o traço de A usando a função `trace` do MATLAB. O autovalor restante λ_{10} tem que ser igual a 11. Por quê? Explique. Calcule os autovalores de A digitando $\mathbf{e} = \texttt{eig}(A)$. Examine os autovalores usando `format long`. Os autovalores calculados têm quantos dígitos de precisão?

(b) A rotina do MATLAB para calcular autovalores é baseada no algoritmo QR descrito na Seção 7 do Cap. 7. Podemos, também, calcular os autovalores de A calculando as raízes de seu polinômio característico. Para determinar os coeficientes do polinômio característico de A, defina $\mathbf{p} = \texttt{poly}(A)$. O polinômio característico de A deveria ter coeficientes inteiros. Por quê? Explique. Se definirmos $\mathbf{p} = \texttt{round}(\mathbf{p})$, deveríamos obter os coeficientes exatos do polinômio característico de A. Calcule as raízes de $\mathbf{p}$ fazendo

$$\mathbf{r} = \texttt{roots}(\mathbf{p})$$

e apresente os resultados usando `format long`. Os resultados obtidos têm quantos dígitos de precisão? Qual método para calcular autovalores é mais preciso, o que utiliza a função `eig` ou o que calcula as raízes do polinômio característico?

3. Considere as matrizes

$$A = \begin{pmatrix} 5 & -3 \\ 3 & -5 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 5 & -3 \\ 3 & 5 \end{pmatrix}$$

Observe que as duas matrizes são iguais, exceto pelo elemento na posição (2, 2).

(a) Use o MATLAB para calcular os autovalores de A e de B. Essas matrizes têm o mesmo tipo de autovalores? Para ver como os autovalores mudam ao se mudar o elemento (2, 2), vamos considerar matrizes com um elemento variável na posição (2, 2) da forma

$$C_t = \begin{pmatrix} 5 & -3 \\ 3 & t-5 \end{pmatrix}$$

Quando t varia de 0 a 10, os elementos (2, 2) dessas matrizes variam de -5 a 5. Podemos usar o MATLAB para encontrar o valor de t onde os autovalores passam de real para complexo. Para fazer isso, defina $a = 0$ e $b = 10$ e calcule t e a matriz correspondente $C = C_1$ usando os seguintes comandos:

```
while b - a > 10 * eps
    t = (a + b)/2;
    C = [5 -3; 3  t - 5];
    e = eig(C);
    if imag(e) == zeros(2,1)
        a = t;
    else
        b = t;
    end
end
```

Imprima os valores de t, C e $\mathbf{e}$. Determine, também, uma matriz X formada pelos autovetores de C. Use o MATLAB para testar se X é ou não singular. A matriz C é diagonalizável? Explique.

(b) Para uma descrição geométrica do que está acontecendo, vamos examinar os gráficos dos polinômios característicos dessas matrizes. Para obter o gráfico do polinômio característico de A, defina

$$\mathbf{p} = \texttt{poly}(A); \qquad \mathbf{x} = -10 : 0.1 : 10;$$
$$\mathbf{z} = \texttt{zeros}(\mathbf{x}); \qquad \mathbf{y} = \texttt{polyval}(\mathbf{p}, \mathbf{x});$$

e depois digite

$$\texttt{plot}(\mathbf{x}, \mathbf{z}, \mathbf{x}, \mathbf{y})$$

Como podemos localizar os autovalores de A nesse gráfico? Explique. Como $A = C_0$ e $B = C_{10}$, para ver geometricamente como os autovalores estão mudando, podemos desenhar os gráficos dos polinômios característicos de C_t para $t = 1, \ldots, 10$. Use o comando **hold** do MATLAB para congelar a tela contendo o gráfico atual, de modo que todos os gráficos apareçam no sistema atual de eixos. Os gráficos podem ser gerados através dos seguintes comandos:

```
for t = 1 : 10
      p = poly([5  -3;  3   t - 5]);
      y = polyval(p, x);
      plot(x, y), pause(1)
end
```

Explique geometricamente o que está acontecendo com as raízes dos polinômios característicos quando t varia de 0 a 10.

4. Defina

$$B = \texttt{toeplitz}(0{:}-1{:}-3, 0{:}3)$$

A matriz B não é simétrica e, portanto, pode não ser diagonalizável. Use o MATLAB para verificar que o posto de B é igual a 2. Explique por que 0 tem que ser um autovalor de B e por que o auto-espaço associado tem que ter dimensão 2. Faça $[X, D] = \texttt{eig}(B)$. Calcule $X^{-1}BX$ e compare o resultado com D. Calcule, também, XD^5X^{-1} e compare o resultado com B^5.

5. Defina

$$C = \texttt{triu}(\texttt{ones}(4), 1) + \texttt{diag}([1, -1], -2)$$
$$[X, D] = \texttt{eig}(C)$$

Calcule $X^{-1}CX$ e compare o resultado com D. C é diagonalizável? Calcule o posto de X e o número condicional de X. Se o número condicional de X for muito grande, os valores calculados para os autovalores podem não ser precisos. Calcule a forma escada reduzida por linhas de C. Explique por que 0 tem que ser um autovalor de C e por que o auto-espaço correspondente tem que ter dimensão 1. Use o MATLAB para calcular C^4. Deveria ser a matriz nula. Sabendo que $C^4 = O$, o que você pode dizer sobre os valores dos outros autovalores de C? Explique. C é diagonalizável? Explique.

6. Construa uma matriz não-diagonalizável definindo

$$A = \texttt{ones}(6); \qquad A = A - \texttt{tril}(A) - \texttt{triu}(A, 2)$$

É fácil ver que $\lambda = 0$ é o único autovalor de A e que seu auto-espaço é gerado por $\mathbf{e}_1$. Verifique isso usando o MATLAB para calcular os autovalores e autovetores de A. Examine os autovetores usando `format long`. Os autovetores calculados são múltiplos de $\mathbf{e}_1$? Faça agora uma transformação em A, obtendo uma matriz semelhante. Defina

$$Q = \texttt{orth(rand(6));} \qquad \text{e} \qquad B = Q' * A * Q$$

Se os cálculos tivessem sido feitos em aritmética exata, a matriz B seria semelhante a A e, portanto, não seria diagonalizável. Use o MATLAB para calcular os autovalores de B e uma matriz X formada por autovetores de B. Determine o posto de X. A matriz calculada B é diagonalizável? Devido aos erros de arredondamento, uma pergunta mais razoável é se a matriz calculada B está ou não próxima de uma matriz não-diagonalizável (isto é, os vetores colunas de X estão próximos de serem linearmente dependentes?). Para responder a essa pergunta, use o MATLAB para calcular `rcond`(X), o inverso do número condicional de X. Um valor de `rcond` próximo de zero indica que X não tem posto máximo.

7. Gere uma matriz A definindo

$$B = \texttt{[} -1, \ -1; \ 1, \ 1 \texttt{]}, \qquad A = \texttt{[zeros(2), eye(2); eye(2),} \ B\texttt{]}$$

(a) A matriz A deveria ter autovalores $\lambda_1 = 1$ e $\lambda_2 = -1$. Use o MATLAB para verificar isso calculando as formas escadas reduzidas por linhas de $A - I$ e de $A + I$. Quais as dimensões dos auto-espaços associados a λ_1 e a λ_2?

(b) É fácil ver que `trace`$(A) = 0$ e `det`$(A) = 1$. Verifique esses resultados usando o MATLAB. Use os valores do traço e do determinante para provar que tanto 1 quanto -1 são, de fato, autovalores duplos. A é diagonalizável? Explique.

(c) Faça $\mathbf{e}$ = `eig`(A) e examine os autovalores usando `format long`. Os autovalores calculados têm quantos dígitos de precisão? Defina $[X, D]$ = `eig`(A) e calcule o número condicional de X. O logaritmo do número condicional de X dá uma estimativa de quantos dígitos de precisão são perdidos no cálculo dos autovalores de A.

(d) Calcule o posto de X. Os autovetores calculados são linearmente independentes? Use o MATLAB para calcular $X^{-1}AX$. A matriz calculada X diagonaliza A?

APLICAÇÃO: GENES LIGADOS AO SEXO

8. Suponha que 10.000 homens e 10.000 mulheres imigram para uma ilha no Pacífico aberta para desenvolvimento. Suponha, também, que um estudo médico dos imigrantes descobre que 200 dos homens são daltônicos e apenas 9 das mulheres são daltônicas. Denote por $x(1)$ a proporção de genes para o daltonismo na população masculina e por $x(2)$ a proporção na população feminina. Suponha que $x(1)$ é igual à proporção de homens daltônicos e que $x(2)^2$ é igual à proporção de mulheres daltônicas. Encontre $x(1)$ e $x(2)$ e coloque os valores no MATLAB como um vetor coluna $\mathbf{x}$. Coloque também a matriz A da Aplicação 2 na Seção 3. Configure o MATLAB para usar `format long` e use a matriz A para calcular as proporções de genes para daltonismo nas 5.ª, 10.ª, 20.ª e 40.ª gerações. Quais os percentuais limites de genes para o daltonismo nessa população? A longo prazo, qual o percentual de homens e mulheres daltônicos?

SEMELHANÇA

9. Defina

$$\begin{array}{ll} S = \texttt{round(10} * \texttt{rand(5));} & S = \texttt{triu}(S, 1) + \texttt{eye(5)} \\ S = S' * S & T = \texttt{inv}(S) \end{array}$$

(a) A inversa exata de S deveria ter elementos inteiros. Por quê? Explique. Verifique os elementos de T usando `format long`. Arredonde os elementos de T para o inteiro mais próximo fazendo T = `round`(T). Calcule $T*S$ e compare com `eye(5)`.

(b) Defina

$$A = \texttt{triu(ones}(5), 1) + \texttt{diag}(1:5), \qquad B = S * A * T$$

Ambas as matrizes A e B têm autovalores 1, 2, 3, 4, 5. Use o MATLAB para calcular os autovalores de B. Os autovalores calculados têm quantos dígitos de precisão? Use o MATLAB para calcular e comparar cada par de quantidades a seguir.

(i) $\det(A)$ e $\det(B)$
(ii) $\texttt{trace}(A)$ e $\texttt{trace}(B)$
(iii) SA^2T e B^2
(iv) $SA^{-1}T$ e B^{-1}

MATRIZES AUTO-ADJUNTAS

10. Construa uma matriz auto-adjunta complexa definindo

$$j = \texttt{sqrt}(-1); \quad A = \texttt{rand}(5) + j * \texttt{rand}(5); \quad A = (A + A')/2$$

(a) Os autovalores de A deveriam ser reais. Por quê? Calcule os autovalores e examine seus resultados usando `format long`. Os autovalores calculados são reais? Calcule também os autovetores fazendo

$$[X, D] = \texttt{eig}(A)$$

Você esperaria que X fosse uma matriz de que tipo? Use o comando $X' * X$ do MATLAB para calcular X^HX. Esses resultados coincidem com o que você esperava?

(b) Defina

$$E = D + j * \texttt{eye}(5) \quad \text{e} \quad B = X * E/X$$

Você esperaria que B fosse uma matriz de que tipo? Use o MATLAB para calcular B^HB e BB^H. Qual a relação entre essas duas matrizes?

MATRIZES POSITIVAS DEFINIDAS

11. Defina

$$C = \texttt{ones}(6) + 7 * \texttt{eye}(6) \quad \text{e} \quad [X, D] = \texttt{eig}(C)$$

(a) Embora $\lambda = 7$ seja um autovalor de multiplicidade 5, a matriz C tem que ser diagonalizável. Por quê? Explique. Verifique que C é diagonalizável calculando o posto de X. Calcule também X^TX. X é que tipo de matriz? Explique. Calcule também o posto de $C - 7I$. O que você pode concluir sobre a dimensão do auto-espaço associado a $\lambda = 7$? Explique.

(b) A matriz C deveria ser simétrica positiva definida. Por quê? Explique. Logo, C deveria ter uma decomposição de Cholesky LL^T. Use o comando `flops(0)` do MATLAB para zerar a contagem das operações aritméticas de ponto flutuante. O comando $R = \texttt{chol}(C)$ do MATLAB vai gerar uma matriz triangular superior R que é igual a L^T. Calcule R dessa maneira e verifique o valor de `flops` para ver quantas operações de ponto flutuante foram necessárias. Faça $L = R'$ e use o MATLAB para verificar que

$$C = LL^T = R^TR$$

(c) Coloque de novo a contagem das operações de ponto flutuante em 0 e calcule a fatoração LU de C fazendo

$$[L\ \ U] = \texttt{lu}(C)$$

Verifique o valor de `flops` para ver quantas operações de ponto flutuante foram necessárias. Qual dos cálculos foi mais eficiente, a decomposição de Cholesky ou a fatoração LU? Defina

$$D = \texttt{diag(sqrt(diag(}U\texttt{)))} \quad \text{e} \quad W = (L * D)'$$

Qual a relação entre R e W?

12. Para qualquer inteiro positivo n, o comando $P = \texttt{pascal}(n)$ do MATLAB gera uma matriz P $n \times n$ cujos elementos são dados por

$$p_{ij} = \begin{cases} 1 & \text{se } i = 1 \quad \text{ou} \quad j = 1 \\ p_{i-1,j} + p_{i,j-1} & \text{se } i > 1 \quad \text{e} \quad j > 1 \end{cases}$$

O nome `pascal` refere-se ao triângulo de Pascal, um arranjo triangular de números usado para gerar os coeficientes binomiais. Os elementos da matriz P formam um triângulo de Pascal.

(a) Defina

$$P = \texttt{pascal}(6)$$

e calcule o valor de seu determinante. Agora subtraia 1 do elemento (6, 6) de P fazendo

$$P(6,6) = P(6,6) - 1$$

e calcule o determinante da nova matriz P. Qual o efeito geral de subtrair 1 do elemento (6, 6) de uma matriz de Pascal 6×6?

(b) Vimos no item (a) que o determinante de uma matriz de Pascal 6×6 é 1 mas que, se subtrairmos 1 do elemento na posição (6, 6), a matriz fica singular. Isso acontece em geral para matrizes de Pascal $n \times n$? Para responder essa pergunta, considere os casos $n = 4$, 8, 12. Em cada caso, faça $P = \texttt{pascal}(n)$ e calcule seu determinante. A seguir, subtraia 1 do elemento (n, n) e calcule o determinante da matriz resultante. A propriedade que descobrimos no item (a) é válida para matrizes de Pascal em geral?

(c) Defina

$$\text{P} = \texttt{pascal (8)}$$

e examine suas submatrizes principais. Supondo que todas as matrizes de Pascal têm determinante igual a 1, por que P tem que ser positiva definida? Calcule o fator triangular superior R na decomposição de Cholesky de P. Como os elementos não-nulos de R podem ser gerados como um triângulo de Pascal? Em geral, qual a relação entre o determinante de uma matriz positiva definida e o determinante de um de seus fatores na decomposição de Cholesky? Por que $\det(P) = 1$?

(d) Defina

$$R(8,8) = 0 \quad \text{e} \quad Q = R' * R$$

A matriz Q deveria ser singular. Por quê? Explique. Por que as matrizes P e Q deveriam ser iguais exceto pelo elemento (8, 8)? Por que $q_{88} = p_{88} - 1$? Explique. Verifique a relação entre P e Q calculando a diferença $P - Q$.

CAPÍTULO 7

ÁLGEBRA LINEAR NUMÉRICA

Neste capítulo, vamos considerar métodos computacionais para resolver problemas de álgebra linear. Para compreender esses métodos, é necessário entender o tipo de sistema numérico usado em computadores. Quando são colocados dados no computador, eles são transformados para um sistema numérico finito próprio do computador. Essa transformação normalmente envolve erros de arredondamento. Outros erros de arredondamento ocorrem quando se efetuam as operações algébricas contidas nos algoritmos. Por essa razão, não podemos esperar obter a solução exata do problema original. O melhor que se pode esperar obter é uma boa aproximação para uma ligeira perturbação do problema original. Suponha, por exemplo, que se queira resolver $A\mathbf{x} = \mathbf{b}$. Ao alimentarmos o computador com os elementos de A e de $\mathbf{b}$, vamos ter, em geral, erros de arredondamento. Então, o programa vai estar tentando, na verdade, encontrar uma boa aproximação para a solução de um sistema da forma $(A + E)\mathbf{x} = \hat{\mathbf{b}}$. Um algoritmo é dito *estável* se fornece uma boa aproximação para uma ligeira perturbação de um problema. Algoritmos que normalmente convergiriam para a solução se usássemos aritmética exata podem perfeitamente não ser estáveis, devido ao crescimento do erro nos processos algébricos. Mesmo usando um algoritmo estável, podemos encontrar problemas sensíveis a perturbações. Por exemplo, se A for "quase singular", as soluções exatas de $A\mathbf{x} = \mathbf{b}$ e de $(A + E)\mathbf{x} = \mathbf{b}$ podem ser muito diferentes, embora todos os elementos de E sejam pequenos. A maior parte deste capítulo é dedicada a métodos numéricos para a solução de sistemas lineares. Vamos dar atenção especial ao crescimento do erro e à sensibilidade de sistemas a pequenas perturbações.

Um outro problema muito importante em aplicações numéricas é o de encontrar os autovalores de uma matriz. Apresentaremos dois métodos iterativos para calcular autovalores na Seção 7. O segundo desses métodos é o poderoso algoritmo *QR*, que utiliza os tipos especiais de transformações ortogonais apresentadas na Seção 5.

Em muitos cálculos matriciais é importante saber se a matriz está próxima, ou não, de ser singular. Na Seção 6, definiremos o conceito de valores singulares de matrizes. O menor valor singular não-nulo pode ser usado para medir quão perto a matriz está de uma matriz de posto menor. Na Seção 8, vamos mostrar como usar valores singulares na solução de problemas de mínimos quadráticos. Apresentaremos, também nessa seção, o algoritmo de Golub-Reinsch para calcular valores singulares.

1 NÚMEROS EM PONTO FLUTUANTE

Ao resolver um problema numérico em um computador, não esperamos, em geral, obter a resposta exata. Algum erro é inevitável. Erros de arredondamento podem ocorrer inicialmente quando os dados são representados no sistema numérico finito utilizado pelo computador. Erros adicionais podem ocorrer quando se efetuam operações aritméticas. Como um câncer, esses erros podem crescer de tal modo que a solução calculada numericamente pode estar completamente errada. Para evitar isso, é preciso entender como aparecem os erros computacionais. Para isso, é necessário familiaridade com o tipo de números usado pelo computador.

Definição. Um **número em ponto flutuante** em base b é um número da forma

$$\pm\left(\frac{d_1}{b}+\frac{d_2}{b^2}+\cdots+\frac{d_t}{b^t}\right)\times b^e$$

onde t, d_1, d_2, ..., d_t, b, e são todos inteiros e

$$0 \le d_i \le b-1 \qquad i = 1,\ldots,t$$

O inteiro t é o número de dígitos e depende do tamanho da palavra no computador. O expoente e fica restrito a um intervalo, $L \le e \le U$, que também depende do computador particular. A maioria dos computadores usa base 2, embora alguns usem outras bases, como 8 e 16. Calculadoras portáteis usam, geralmente, base 10.

EXEMPLO 1. Os números a seguir são números decimais (base 10) em ponto flutuante com cinco dígitos:

$$0{,}53216 \times 10^{-4}$$
$$-0{,}81724 \times 10^{21}$$
$$0{,}00112 \times 10^{8}$$
$$0{,}11200 \times 10^{6}$$

Note que os números $0{,}00112 \times 10^8$ e $0{,}11200 \times 10^6$ são iguais. Logo, a representação de um número em ponto flutuante não precisa ser única. Números em ponto flutuante escritos sem zeros logo após a vírgula são ditos normalizados. □

EXEMPLO 2. $(0{,}236)_8 \times 8^2$ e $(0{,}132)_8 \times 8^4$ são números em ponto flutuante em base 8 normalizados e com três dígitos. Assim, $(0{,}236)_8$ representa

$$\frac{2}{8}+\frac{3}{8^2}+\frac{6}{8^3}$$

Logo, $(0{,}236)_8 \times 8^2$ é a representação em ponto flutuante e base 8 do número decimal 19,75. Analogamente,

$$(0{,}132)_8 \times 8^4 = \left(\frac{1}{8}+\frac{3}{8^2}+\frac{2}{8^3}\right)\times 8^4 = 720$$ □

Para entender melhor o tipo de sistema numérico com que estamos trabalhando, vamos olhar um exemplo bem simples.

EXEMPLO 3. Suponha que $t = 1$, $L = -1$, $U = 1$ e $b = 10$. Existem ao todo 55 números em ponto flutuante nesse sistema com um dígito. Eles são

$$0, \pm 0{,}1 \times 10^{-1}, \pm 0{,}2 \times 10^{-1}, \ldots, \pm 0{,}9 \times 10^{-1}$$
$$\pm 0{,}1 \times 10^{0}, \pm 0{,}2 \times 10^{0}, \ldots, \pm 0{,}9 \times 10^{0}$$
$$\pm 0{,}1 \times 10^{1}, \pm 0{,}2 \times 10^{1}, \ldots, \pm 0{,}9 \times 10^{1}$$

Embora todos esses números estejam no intervalo $[-9, 9]$, mais de um terço deles tem valor absoluto menor do que 0,1 e mais do que dois terços têm valor absoluto menor do que 1. A Fig. 7.1.1 ilustra a distribuição no intervalo [0, 2] desses números em ponto flutuante. □

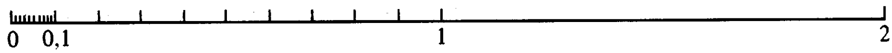


FIG. 7.1.1

Em sua maior parte, os números reais precisam ser arredondados para serem representados em um sistema numérico de pontos flutuantes com t dígitos. A diferença entre os valores do número x' em ponto flutuante e do número original x é chamada de *erro de arredondamento.* O tamanho do erro de arredondamento faz mais sentido, talvez, se comparado ao tamanho do número original.

Definição. Se x é um número real e x' é sua aproximação em ponto flutuante, então a diferença $x' - x$ é chamada de **erro absoluto**, e o quociente $(x' - x)/x$ é chamado de **erro relativo**.

Quando se efetuam operações aritméticas com números em ponto flutuante, podem ocorrer erros de arredondamento adicionais.

Número Real, x	Número Decimal em Ponto Flutuante com Quatro Dígitos, x'	Erro Absoluto, $x' - x$	Erro Relativo $(x' - x)/x$
62.133	$0{,}6213 \times 10^5$	-3	$\dfrac{-3}{62.133} \approx -4{,}8 \times 10^{-5}$
0,12658	$0{,}1266 \times 10^0$	2×10^{-5}	$\dfrac{1}{6329} \approx 1{,}6 \times 10^{-4}$
47,213	$0{,}4721 \times 10^2$	$-3{,}0 \times 10^{-3}$	$\dfrac{-0{,}003}{47{,}213} \approx -6{,}4 \times 10^{-5}$
π	$0{,}3142 \times 10^1$	$3{,}142 - \pi \approx 4 \times 10^{-4}$	$\dfrac{3{,}142 - \pi}{\pi} \approx 1{,}3 \times 10^{-4}$

EXEMPLO 4. Sejam $a' = 0{,}263 \times 10^4$ e $b' = 0{,}466 \times 10^1$ números decimais em ponto flutuante com três dígitos. Se esses números forem somados, a soma exata vai ser

$$a' + b' = 0{,}263446 \times 10^4$$

No entanto, a representação em ponto flutuante desse número é $0{,}263 \times 10^4$. Essa deve ser, portanto, a soma calculada. Vamos denotar a soma em ponto flutuante por $fl(a' + b')$. O erro absoluto na soma é

$$fl(a' + b') - (a' + b') = -4{,}46$$

e o erro relativo é

$$\frac{-4{,}46}{0{,}26344 \times 10^4} \approx -0{,}17 \times 10^{-2}$$

O valor exato de $a'b'$ é 11.729,8; no entanto, $fl(a'b')$ é $0{,}117 \times 10^5$. O erro absoluto no produto é $-29{,}8$ e o erro relativo é aproximadamente $-0{,}25 \times 10^{-2}$. Subtração e divisão podem ser feitas de maneira análoga. □

Em geral, o erro relativo na aproximação de um número x por sua representação em ponto flutuante x' é denotado por δ. Então

$$(1) \qquad \delta = \frac{x' - x}{x} \qquad \text{ou} \qquad x' = x(1 + \delta)$$

$|\delta|$ pode ser limitado por uma constante positiva ε, chamada de *precisão da máquina* ou *épsilon da máquina*. O épsilon da máquina é definido como o menor número em ponto flutuante tal que

$$fl(1 + \epsilon) > 1$$

Por exemplo, se o computador usa números decimais com três dígitos, então

$$fl(1 + 0{,}499 \times 10^{-2}) = 1$$

enquanto

$$fl(1 + 0{,}500 \times 10^{-2}) = 1{,}01$$

Nesse caso, o épsilon da máquina seria $0{,}500 \times 10^{-2}$.

Da Equação (1), vemos que, se a' e b' são dois números em ponto flutuante, então

$$fl(a' + b') = (a' + b')(1 + \delta_1)$$
$$fl(a'b') = (a'b')(1 + \delta_2)$$
$$fl(a' - b') = (a' - b')(1 + \delta_3)$$
$$fl(a' \div b') = (a' \div b')(1 + \delta_4)$$

Os δ_i são erros relativos e todos vão ter valor absoluto menor do que ε. Note que, no Exemplo 4, $\delta_1 \approx -0{,}17 \times 10^{-2}$, $\delta_2 \approx -0{,}25 \times 10^{-2}$ e $\varepsilon = 0{,}5 \times 10^{-2}$.

Se os números com os quais você está trabalhando envolverem alguns pequenos erros, as operações aritméticas podem aumentar esses erros. Se dois números são iguais até a k-ésima casa decimal e um dos números é subtraído do outro, haverá uma perda de dígitos significativos em sua resposta. Nesse caso, o erro relativo da diferença pode ser muitas vezes maior do que o erro relativo em qualquer dos números.

EXEMPLO 5. Sejam $c = 3{,}4215298$ e $d = 3{,}4213851$. Calcule $c - d$ e $1/(c - d)$ usando aritmética de ponto flutuante com seis dígitos.

SOLUÇÃO

(i) O primeiro passo é representar c e d em ponto flutuante com seis dígitos.

$$c' = 0{,}342153 \times 10^1$$
$$d' = 0{,}342139 \times 10^1$$

Os erros relativos em c e d são

$$\frac{c' - c}{c} \approx 0{,}6 \times 10^{-7} \qquad \text{e} \qquad \frac{d' - d}{d} \approx 1{,}4 \times 10^{-6}$$

(ii) $fl(c' - d') = c' - d' = 0{,}140000 \times 10^{-3}$. O valor real de $c - d$ é $0{,}1447 \times 10^{-3}$. Os erros absoluto e relativo na aproximação de $c - d$ por $fl(c' - d')$ são

$$fl(c' - d') - (c - d) = -0{,}47 \times 10^{-5}$$

e

$$\frac{fl(c' - d') - (c - d)}{c - d} \approx -3{,}2 \times 10^{-2}$$

Note que o erro relativo da diferença é da ordem de mais de 10^4 vezes o erro relativo na representação de c ou de d.

(iii) $fl([1/(c' - d')] = 0{,}714286 \times 10^4$ e a resposta correta com seis dígitos significativos é

$$\frac{1}{c-d} \approx 0{,}691085 \times 10^4$$

Os erros absolutos e relativos são aproximadamente 232 e 0,03. □

EXERCÍCIOS

1. Encontre a representação em ponto flutuante com três dígitos para cada um dos números a seguir.

(a) 2312 (b) 32,56 (c) 0,01277 (d) 82.431

2. Encontre os erros absoluto e relativo na representação de cada um dos números no Exercício 1 em ponto flutuante decimal com três dígitos.

3. Represente cada um dos números a seguir como um número binário em ponto flutuante com cinco dígitos.

(a) 21 (b) $\frac{3}{8}$ (c) 9,872 (d) $-0{,}1$

4. Efetue cada uma das operações a seguir usando a aritmética de pontos flutuantes decimais com quatro dígitos e calcule os erros absolutos e relativos em suas respostas.

(a) $10.420 + 0{,}0018$ (b) $10.424 - 10.416$
(c) $0{,}12347 - 0{,}12342$ (d) $(3626{,}6) \cdot (22{,}656)$

5. Sejam $x_1 = 94.210$, $x_2 = 8631$, $x_3 = 1440$, $x_4 = 133$ e $x_5 = 34$. Efetue cada uma das operações a seguir usando a aritmética de pontos flutuantes decimais com quatro dígitos.

(a) $(((x_1 + x_2) + x_3) + x_4) + x_5$
(b) $x_1 + ((x_2 + x_3) + (x_4 + x_5))$
(c) $(((x_5 + x_4) + x_3) + x_2) + x_1$

6. Qual seria o épsilon de um computador que trabalha com aritmética de ponto flutuante em base 2 com cinco dígitos?

7. Quantos números em ponto flutuante existem em um sistema com $t = 2$, $L = -2$, $U = 2$ e $b = 2$?

2 MÉTODO DE GAUSS

Nesta seção, vamos discutir o problema de resolver um sistema com n equações e n incógnitas usando o método de Gauss. O método de Gauss é considerado, em geral, o método computacional mais eficiente, já que envolve o menor número de operações aritméticas.

MÉTODO DE GAUSS SEM MUDANÇAS DE ORDEM DAS LINHAS

Seja $A = A^{(1)} = (a_{ij}^{(1)})$ uma matriz invertível. Então, A pode ser colocada em forma triangular usando as operações elementares I e III. Para fins de simplificação, vamos supor que isso pode ser feito usando-se apenas a operação elementar III.

$$A^{(1)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ \boxed{\begin{matrix} a_{21}^{(1)} \\ \vdots \\ a_{n1}^{(1)} \end{matrix}} & \begin{matrix} a_{22}^{(1)} \\ \\ a_{n2}^{(1)} \end{matrix} & \begin{matrix} \cdots \\ \\ \cdots \end{matrix} & \begin{matrix} a_{2n}^{(1)} \\ \\ a_{nn}^{(1)} \end{matrix} \end{pmatrix}$$

Etapa 1

Sejam $m_{k1} = a_{k1}^{(1)}/a_{11}^{(1)}$ para $k = 2, \ldots, n$ (por hipótese, $a_{11}^{(1)} \neq 0$). O primeiro passo no processo é usar a operação elementar III $n - 1$ vezes para anular os elementos abaixo da diagonal na primeira coluna de A. Note que m_{k1} é o número pelo qual devemos multiplicar a primeira linha para ela ser subtraída da k-ésima. A nova matriz obtida é

$$A^{(2)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & & & \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{pmatrix}$$

onde

$$a_{kj}^{(2)} = a_{kj}^{(1)} - m_{k1}a_{1j}^{(1)} \qquad (2 \leq k \leq n,\ 2 \leq j \leq n)$$

A primeira etapa do método de Gauss efetua $n - 1$ divisões, $(n - 1)^2$ multiplicações e $(n - 1)^2$ adições ou subtrações.

Etapa 2

Se $a_{22}^{(2)} \neq 0$, então esse elemento pode ser usado como pivô para anular $a_{32}^{(2)}, \ldots, a_{n2}^{(2)}$. Para $k = 3, \ldots, n$, defina

$$m_{k2} = \frac{a_{k2}^{(2)}}{a_{22}^{(2)}}$$

e subtraia m_{k2} vezes a segunda linha de $A^{(2)}$ da k-ésima linha. A nova matriz obtida é

$$A^{(3)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_{n3}^{(3)} & \cdots & a_{nn}^{(3)} \end{pmatrix}$$

A segunda etapa efetua $n - 2$ divisões, $(n - 2)^2$ multiplicações e $(n - 2)^2$ adições ou subtrações.

Após $n - 1$ etapas, terminaremos com uma matriz triangular $U = A^{(n)}$. O número de operações efetuadas durante todo o processo pode ser determinado da seguinte maneira:

Divisões:

$$(n-1) + (n-2) + \cdots + 1 = \frac{n(n-1)}{2}$$

Multiplicações:

$$(n-1)^2 + (n-2)^2 + \cdots + 1^2 = \frac{n(2n-1)(n-1)}{6}$$

Adições e/ou subtrações:

$$(n-1)^2 + \cdots + 1^2 = \frac{n(2n-1)(n-1)}{6}$$

O método de Gauss é resumido no algoritmo a seguir.

Algoritmo 7.2.1 (Método de Gauss sem Troca de Ordem das Linhas)

Para $i = 1, 2, \ldots, n-1$

Para $k = i+1, \ldots, n$

Defina $m_{ki} = \dfrac{a_{ki}^{(i)}}{a_{ii}^{(i)}}$ (*desde que* $a_{ii}^{(i)} \neq 0$)

Para $j = i+1, \ldots, n$

Defina $a_{kj}^{(i+1)} = a_{kj}^{(i)} - m_{ki}a_{ij}^{(i)}$

→ *Fim*

→ *Fim*

→ *Fim*

Para resolver o sistema $A\mathbf{x} = \mathbf{b}$, poderíamos aumentar A usando $\mathbf{b}$. Assim, $\mathbf{b}$ ficaria como uma coluna extra de A. O método de Gauss poderia ser usado através do Algoritmo 7.2.1 e fazendo-se j variar de $i+1$ até $n+1$, em vez de $i+1$ até n. O sistema triangular poderia ser resolvido através de substituição de baixo para cima.

A maior parte do trabalho envolvido na solução de um sistema $A\mathbf{x} = \mathbf{b}$ se dá durante a redução de A a uma forma triangular. Suponha que, após resolver o sistema $A\mathbf{x} = \mathbf{b}$, queremos resolver um outro, $A\mathbf{x} = \mathbf{b}_1$. Conhecemos a forma triangular U do primeiro sistema e, portanto, gostaríamos de resolver o novo sistema sem ter que refazer todo o processo. Podemos fazer isso se usarmos a fatoração LU discutida na Seção 6 do Cap. 6 (ver o Exemplo 1 daquela seção). Para calcular a fatoração LU, guardamos os números m_{ki} usados no Algoritmo 7.2.1. Esses números são chamados de *multiplicadores*. O multiplicador m_{ki} é o número pelo qual multiplicamos a i-ésima linha para subtraí-la da k-ésima durante a i-ésima etapa do método de Gauss. Para ver como os multiplicadores podem ser usados para resolver $A\mathbf{x} = \mathbf{b}_1$, vamos olhar o método de Gauss em termos de multiplicações matriciais.

FATORAÇÃO TRIANGULAR

A primeira etapa no método de Gauss consiste em multiplicar A por $n-1$ matrizes elementares,

$$A^{(2)} = E_{n1} \cdots E_{31}E_{21}A^{(1)}$$

onde

$$E_{k1} = \begin{pmatrix} 1 & & & & & \\ 0 & 1 & & & & \\ \vdots & & \ddots & & & \\ -m_{k1} & 0 & \cdots & 1 & & \\ \vdots & & & & \ddots & \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

Cada E_i é invertível, com inversa

$$E_{k1}^{-1} = \begin{pmatrix} 1 & & & & & \\ 0 & 1 & & & & \\ \vdots & & \ddots & & & \\ m_{k1} & 0 & \cdots & 1 & & \\ \vdots & & & & \ddots & \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

Seja

$$M_1 = E_{n1} \cdots E_{31} E_{21} = \begin{pmatrix} 1 & & & & \\ -m_{21} & 1 & & & \\ -m_{31} & 0 & 1 & & \\ \vdots & & & \ddots & \\ -m_{n1} & 0 & 0 & \cdots & 1 \end{pmatrix}$$

Logo, $A^{(2)} = M_1 A$. A matriz M_1 é invertível e

$$M_1^{-1} = E_{21}^{-1} E_{31}^{-1} \cdots E_{n1}^{-1} = \begin{pmatrix} 1 & & & & \\ m_{21} & 1 & & & \\ m_{31} & 0 & 1 & & \\ \vdots & & & \ddots & \\ m_{n1} & 0 & 0 & \cdots & 1 \end{pmatrix}$$

Analogamente,

$$\begin{aligned} A^{(3)} &= E_{n2} \cdots E_{42} E_{32} A^{(2)} \\ &= M_2 A^{(2)} \\ &= M_2 M_1 A \end{aligned}$$

onde

$$M_2 = E_{n2} \cdots E_{32} = \begin{pmatrix} 1 & & & & \\ 0 & 1 & & & \\ 0 & -m_{32} & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & -m_{n2} & 0 & \cdots & 1 \end{pmatrix}$$

e

$$M_2^{-1} = \begin{pmatrix} 1 & & & & \\ 0 & 1 & & & \\ 0 & m_{32} & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & m_{n2} & 0 & \cdots & 1 \end{pmatrix}$$

Ao final do processo, temos

$$U = A^{(n)} = M_{n-1} \cdots M_2 M_1 A$$

Segue que

$$A = M_1^{-1} M_2^{-1} \cdots M_{n-1}^{-1} U$$

O produto dos M_j^{-1} nos dá a seguinte matriz triangular inferior, quando multiplicados nessa ordem:

$$L = M_1^{-1} M_2^{-1} \cdots M_{n-1}^{-1} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ m_{21} & 1 & 0 & \cdots & 0 \\ m_{31} & m_{32} & 1 & \cdots & 0 \\ \vdots & & & & \\ m_{n1} & m_{n2} & m_{n3} & \cdots & 1 \end{pmatrix}$$

Logo, $A = LU$, onde L é triangular inferior e U é triangular superior.

EXEMPLO 1. Seja

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 4 & 1 & 4 \\ 3 & 4 & 6 \end{pmatrix}$$

O método de Gauss pode ser feito em duas etapas:

$$\begin{pmatrix} 2 & 3 & 1 \\ 4 & 1 & 4 \\ 3 & 4 & 6 \end{pmatrix} \xrightarrow{1} \begin{pmatrix} 2 & 3 & 1 \\ 0 & -5 & 2 \\ 0 & -\frac{1}{2} & \frac{9}{2} \end{pmatrix} \xrightarrow{2} \begin{pmatrix} 2 & 3 & 1 \\ 0 & -5 & 2 \\ 0 & 0 & 4{,}3 \end{pmatrix}$$

Os multiplicadores utilizados foram $m_{21} = 2$, $m_{31} = \dfrac{3}{2}$ (etapa 1) e $m_{32} = \dfrac{1}{10}$ (etapa 2). Sejam

$$L = \begin{pmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & m_{32} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ \frac{3}{2} & \frac{1}{10} & 1 \end{pmatrix}$$

e

$$U = \begin{pmatrix} 2 & 3 & 1 \\ 0 & -5 & 2 \\ 0 & 0 & 4{,}3 \end{pmatrix}$$

O leitor pode verificar que $LU = A$. □

Uma vez que se colocou A em forma triangular e se encontrou a fatoração LU, o sistema $A\mathbf{x} = \mathbf{b}$ pode ser resolvido em duas etapas.

Etapa 1. Substituição de Cima para Baixo

O sistema $A\mathbf{x} = \mathbf{b}$ pode ser escrito na forma

$$LU\mathbf{x} = \mathbf{b}$$

Seja $\mathbf{y} = U\mathbf{x}$. Temos

$$L\mathbf{y} = LU\mathbf{x} = \mathbf{b}$$

Logo, podemos encontrar **y** resolvendo o sistema triangular inferior

$$\begin{aligned}
&y_1 &&= b_1\\
&m_{21}y_1 + y_2 &&= b_2\\
&m_{31}y_1 + m_{32}y_2 + y_3 &&= b_3\\
&\vdots\\
&m_{n1}y_1 + m_{n2}y_2 + m_{n3}y_3 + \cdots + y_n &&= b_n
\end{aligned}$$

Pela primeira equação, temos $y_1 = b_1$. Esse valor pode ser usado na segunda equação para se encontrar y_2. Os valores de y_1 e y_2 podem ser usados na terceira equação para se encontrar y_3, e assim por diante. Esse método de solução para um sistema triangular inferior é chamado de *substituição de cima para baixo.*

Etapa 2. Substituição de Baixo para Cima

Uma vez encontrado **y**, precisamos apenas resolver o sistema triangular superior $U\mathbf{x} = \mathbf{y}$ para encontrar a solução **x** do sistema original. O sistema triangular superior é resolvido por substituição, começando pela última equação.

EXEMPLO 2. Resolva o sistema

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

SOLUÇÃO. A matriz de coeficientes para esse sistema é a matriz A do Exemplo 1. Como já encontramos L e U, o sistema pode ser resolvido usando substituição, primeiro de cima para baixo e, depois, de baixo para cima.

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & -4\\ 2 & 1 & 0 & 9\\ \frac{3}{2} & \frac{1}{10} & 1 & 0 \end{array}\right) \qquad \begin{aligned} y_1 &= -4\\ y_2 &= 9 - 2y_1 = 17\\ y_3 &= 0 - \tfrac{3}{2}y_1 - \tfrac{1}{10}y_2 = 4{,}3 \end{aligned}$$

$$\left(\begin{array}{ccc|c} 2 & 3 & 1 & -4\\ 0 & -5 & 2 & 17\\ 0 & 0 & 4{,}3 & 4{,}3 \end{array}\right) \qquad \begin{aligned} 2x_1 + 3x_2 + x_3 &= -4\\ -5x_2 + 2x_3 &= 17\\ 4{,}3x_3 &= 4{,}3 \end{aligned} \qquad \begin{aligned} x_1 &= 2\\ x_2 &= -3\\ x_3 &= 1 \end{aligned}$$

A solução do sistema é $\mathbf{x} = (2, -3, 1)^T$. □

Algoritmo 7.2.2 (Substituição de Cima para Baixo e de Baixo para Cima)

Para $k = 1, \ldots, n$

$$\textit{Defina } y_k = b_k - \sum_{i=1}^{k-1} m_{ki}y_i$$

Fim

Para $k = n, n-1, \ldots, 1$

$$\textit{Defina } x_k = \frac{y_k - \sum_{j=k+1}^{n} u_{kj}x_j}{u_{kk}}$$

Fim

Contagem das Operações. O Algoritmo 7.2.2 efetua n divisões, $n(n-1)$ multiplicações e $n(n-1)$ somas e/ou subtrações. Então, o número total de operações para se resolver um sistema $A\mathbf{x} = \mathbf{b}$, usando os Algoritmos 7.2.1 e 7.2.2, é

$$\begin{array}{ll} \text{Multiplicações/divisões:} & \frac{1}{3}n^3 + n^2 - \frac{1}{3}n \\ \text{Adições/subtrações:} & \frac{1}{3}n^3 + \frac{1}{2}n^2 - \frac{5}{6}n \end{array}$$

Em ambos os casos, $\dfrac{1}{3}n^3$ é o termo dominante. Diremos que resolver um sistema pelo método de Gauss envolve aproximadamente $\dfrac{1}{3}n^3$ multiplicações/divisões e $\dfrac{1}{3}n^3$ somas/subtrações.

Armazenamento. Não é necessário armazenar os multiplicadores em uma matriz separada L. Cada multiplicador pode ser armazenado na matriz A no lugar do elemento $a_{ki}^{(i)}$ que foi anulado. Ao final do método, A está sendo usada para se guardar os m_{ki} e os u_{ij}.

$$\begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1,n-1} & u_{1n} \\ m_{21} & u_{22} & \cdots & u_{2,n-1} & u_{2n} \\ \vdots & & & & \\ m_{n1} & m_{n2} & \cdots & m_{n,n-1} & u_{nn} \end{pmatrix}$$

O Algoritmo 7.2.1 não funciona se, em qualquer etapa, $a_{kk}^{(k)}$ for 0. Se isso acontecer, vai ser necessário trocar a ordem das linhas. Na próxima seção veremos como incorporar troca de ordem das linhas no nosso algoritmo para o método de Gauss.

EXERCÍCIOS

1. Seja

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 1 \\ -3 & 1 & -2 \end{pmatrix}$$

Fatore A em um produto LU, onde L é triangular inferior com todos os elementos diagonais iguais a 1 e U é triangular superior.

2. Seja A a matriz do Exercício 1. Use a fatoração LU para resolver $A\mathbf{x} = \mathbf{b}$ para cada uma das escolhas de $\mathbf{b}$ a seguir.

(a) $(4, 3, -13)^T$ (b) $(3, 1, -10)^T$ (c) $(7, 23, 0)^T$

3. Sejam A e B matrizes $n \times n$ e seja $\mathbf{x} \in R^n$.

(a) Quantas somas e multiplicações escalares são necessárias para se calcular $A\mathbf{x}$?

(b) Quantas somas e multiplicações escalares são necessárias para se calcular o produto AB?

(c) Quantas somas e multiplicações escalares são necessárias para se calcular $(AB)\mathbf{x}$? E para $A(B\mathbf{x})$?

4. Sejam $A \in R^{m\times n}$, $B \in R^{n\times r}$ e $\mathbf{x}, \mathbf{y} \in R^n$. Suponha que o produto $A\mathbf{x}\mathbf{y}^TB$ é calculado de cada uma das seguintes maneiras:

(i) $(A(\mathbf{x}\mathbf{y}^T))B$ (ii) $(A\mathbf{x})(\mathbf{y}^TB)$ (iii) $((A\mathbf{x})\mathbf{y}^T)B$

(a) Quantas somas e multiplicações escalares são necessárias para se efetuar cada um desses cálculos?

(b) Compare o número de somas e multiplicações escalares para cada um dos três métodos no caso em que $m = 5$, $n = 4$ e $r = 3$. Qual o método mais eficiente nesse caso?

5. Seja E_{ki} a matriz elementar obtida subtraindo-se α vezes a i-ésima linha da matriz identidade de sua k-ésima linha.

(a) Mostre que $E_{ki} = I - \alpha \mathbf{e}_k \mathbf{e}_i^T$.

(b) Seja $E_{ji} = I - \beta \mathbf{e}_j \mathbf{e}_i^T$. Mostre que $E_{ji}E_{ki} = I - (\alpha \mathbf{e}_k + \beta \mathbf{e}_j)\mathbf{e}_i^T$.

(c) Mostre que $E_{ki}^{-1} = I + \alpha \mathbf{e}_k \mathbf{e}_i^T$.

6. Seja A uma matriz $n \times n$ com fatoração triangular LU. Mostre que

$$\det(A) = u_{11}u_{22}\cdots u_{nn}$$

7. Se A é uma matriz simétrica $n \times n$ com fatoração triangular LU, então A pode ser fatorada ainda mais em um produto LDL^T (onde D é diagonal). Escreva um algoritmo, semelhante ao algoritmo 7.2.2, para resolver $LDL^T\mathbf{x} = \mathbf{b}$.

8. Escreva um algoritmo para resolver o sistema tridiagonal

$$\begin{pmatrix} a_1 & b_1 & & & \\ c_1 & a_2 & \ddots & & \\ & \ddots & \ddots & & \\ & & \ddots & a_{n-1} & b_{n-1} \\ & & & c_{n-1} & a_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_{n-1} \\ d_n \end{pmatrix}$$

usando o método de Gauss com os elementos diagonais como pivôs. Quantas somas/subtrações e multiplicações/divisões são necessárias?

9. Seja $A = LU$, onde L é triangular inferior com todos os elementos diagonais iguais a 1 e U é triangular superior.

(a) Quantas somas e multiplicações escalares são necessárias para se resolver $L\mathbf{y} = \mathbf{e}_j$ por substituição?

(b) Quantas somas e multiplicações escalares são necessárias para se resolver $A\mathbf{x} = \mathbf{e}_j$? A solução $\mathbf{x}_j$ de $A\mathbf{x} = \mathbf{e}_j$ vai ser a j-ésima coluna de A^{-1}.

(c) Dada a fatoração $A = LU$, quantas somas/subtrações e multiplicações/divisões adicionais são necessárias para se calcular A^{-1}?

10. Suponha que já temos A^{-1} e a fatoração $A = LU$. Quantas somas e multiplicações escalares são necessárias para se calcular $A^{-1}\mathbf{b}$? Compare esse com o número de operações necessárias para resolver $LU\mathbf{x} = \mathbf{b}$ usando o Algoritmo 7.2.2. Suponha que temos uma série de sistemas para resolver com a mesma matriz de coeficientes A. Vale a pena calcular A^{-1}? Explique.

11. Seja A uma matriz 3×3 e suponha que A pode ser transformada em uma matriz triangular inferior L usando-se apenas operações elementares nas colunas de A do tipo III, isto é

$$AE_1E_2E_3 = L$$

onde E_1, E_2 e E_3 são matrizes elementares do tipo III. Seja

$$U = (E_1E_2E_3)^{-1}$$

Mostre que U é triangular superior com todos os elementos diagonais iguais a 1 e que $A = LU$. (Isso ilustra uma versão com operações nas colunas do método de Gauss.)

3 ESTRATÉGIAS DE PIVÔ

Nesta seção, vamos apresentar um algoritmo para o método de Gauss com troca de ordem entre linhas. Em cada etapa do algoritmo será necessário escolher a linha do pivô. Podemos, muitas vezes, evitar grandes acumulações de erros desnecessários escolhendo a linha do pivô de maneira razoável.

MÉTODO DE GAUSS COM TROCA DE ORDEM ENTRE LINHAS

Considere o seguinte exemplo:

EXEMPLO 1. Seja

$$A = \begin{pmatrix} 6 & -4 & 2 \\ 4 & 2 & 1 \\ 2 & -1 & 1 \end{pmatrix}$$

Queremos colocar A em forma triangular usando as operações elementares I e III. Para guardar as trocas de ordem entre linhas, vamos usar um vetor linha $\mathbf{p}$. As coordenadas de $\mathbf{p}$ serão denotadas por $p(1), p(2)$ e $p(3)$. Inicialmente, fazemos $\mathbf{p} = (1, 2, 3)$. Suponha que, na primeira etapa do método de Gauss, escolhemos a terceira como a linha do pivô. Em vez de permutar a primeira e terceira linhas, permutamos o primeiro e o terceiro elementos de $\mathbf{p}$. Fazendo $p(1) = 3$ e $p(3) = 1$, o vetor $\mathbf{p}$ fica $(3, 2, 1)$. O vetor $\mathbf{p}$ é usado para se guardar as mudanças de ordem entre as linhas. Podemos pensar no vetor $\mathbf{p}$ como uma troca de números das linhas. A mudança de ordem das linhas pode ficar para o final.

$$\begin{array}{l} \text{linha} \\ p(3) \\ p(2) \\ p(1) \end{array} \begin{array}{c} \\ \begin{pmatrix} \boxed{6} & -4 & 2 \\ \boxed{4} & 2 & 1 \\ 2 & -1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & -1 & -1 \\ 0 & 4 & -1 \\ 2 & -1 & 1 \end{pmatrix} \end{array}$$

Se, na segunda etapa, escolhe-se o pivô na linha $p(3)$, os elementos de $p(3)$ e $p(2)$ são trocados. Depois efetua-se a última etapa do método de Gauss.

$$\begin{array}{l} p(2) \\ p(3) \\ p(1) \end{array} \begin{pmatrix} 0 & -1 & -1 \\ 0 & \boxed{4} & -1 \\ 2 & -1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & -1 & -1 \\ 0 & 0 & -5 \\ 2 & -1 & 1 \end{pmatrix}$$

Se as linhas são colocadas na ordem $(p(1), p(2), p(3)) = (3, 1, 2)$, a matriz resultante estará em forma triangular,

$$\begin{array}{l} p(1) = 3 \\ p(2) = 1 \\ p(3) = 2 \end{array} \begin{pmatrix} 2 & -1 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & -5 \end{pmatrix}$$

Se as linhas estivessem na ordem $(3, 1, 2)$ no início, o método teria sido exatamente o mesmo, exceto que não haveria troca de ordem entre as linhas. Colocar as linhas de A na ordem $(3, 1, 2)$ é o mesmo que multiplicar a matriz A à esquerda pela matriz de permutação

$$P = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Vamos usar o método em A e PA simultaneamente e analisar os resultados. Os multiplicadores utilizados no processo foram 3, 2 e -4. Eles serão guardados nos lugares dos elementos anulados e colocados em caixas para diferenciá-los dos outros elementos da matriz.

$$A = \begin{pmatrix} 6 & -4 & 2 \\ 4 & 2 & 1 \\ 2 & -1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} \boxed{3} & -1 & -1 \\ \boxed{2} & 4 & -1 \\ 2 & -1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} \boxed{3} & -1 & -1 \\ \boxed{2} & \boxed{-4} & -5 \\ 2 & -1 & 1 \end{pmatrix}$$

$$PA = \begin{pmatrix} 2 & -1 & 1 \\ 6 & -4 & 2 \\ 4 & 2 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & -1 & 1 \\ \boxed{3} & -1 & -1 \\ \boxed{2} & 4 & -1 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & -1 & 1 \\ \boxed{3} & -1 & -1 \\ \boxed{2} & \boxed{-4} & -5 \end{pmatrix}$$

Ao se trocar a ordem das linhas da matriz obtida de A, as duas matrizes resultantes ficarão iguais. A forma reduzida de PA contém toda a informação necessária para se obter sua fatoração triangular. Na verdade,

$$PA = LU$$

onde

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & -4 & 1 \end{pmatrix} \quad \text{e} \quad U = \begin{pmatrix} 2 & -1 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & -5 \end{pmatrix} \qquad \square$$

No computador não há necessidade de se trocar, de fato, a ordem das linhas de A. Basta tratar a linha $p(k)$ como se fosse a linha k e usar $a_{p(k),j}$ no lugar de a_{kj}.

Algoritmo 7.3.1 (Método de Gauss com Troca de Ordem entre Linhas)

Para $i = 1, \ldots, n$
  Defina $p(i) = i$
Fim

Para $i = 1, \ldots, n$
  (1) *Escolha um pivô* $a_{p(j),i}$ *entre os elementos*

$$a_{p(i),i}, a_{p(i+1),i}, \ldots, a_{p(n),i}$$

  (Discutiremos estratégias para fazer isso mais adiante nesta sessão.)
  (2) *Troque as* i-ésimas *e* j-ésimas *elementos de* p.
  (3) *Para* $k = i + 1, \ldots, n$
    Defina $m_{p(k),i} = a_{p(k),i}/a_{p(i),i}$
    Para $j = i + 1, \ldots, n$
      Defina $a_{p(k),j} = a_{p(k),j} - m_{p(k),i}a_{p(i),j}$
    Fim
  Fim
Fim

Observações

1. O multiplicador $m_{p(k),i}$ é guardado na posição do elemento $a_{p(k),i}$ que está sendo anulado.
2. O vetor $\mathbf{p}$ pode ser usado para se formar a matriz de permutação P, cuja i-ésima linha é a $p(i)$-ésima linha da matriz identidade.
3. A matriz PA pode ser fatorada em um produto LU, onde

$$l_{ki} = \begin{cases} m_{p(k),i} & \text{se } k > i \\ 1 & \text{se } k = i \\ 0 & \text{se } k < i \end{cases} \quad \text{e} \quad u_{ki} = \begin{cases} a_{p(k),i} & \text{se } k \leq i \\ 0 & \text{se } k > i \end{cases}$$

4. Como P é invertível, o sistema $A\mathbf{x} = \mathbf{b}$ é equivalente ao sistema $PA\mathbf{x} = P\mathbf{b}$. Seja $\mathbf{c} = P\mathbf{b}$. Como $PA = LU$, o sistema é equivalente a

$$LU\mathbf{x} = \mathbf{c}$$

5. Se $PA = LU$, então $A = P^{-1}LU = P^TLU$

Segue das Observações 4 e 5 que, se $A = P^TLU$, então o sistema $A\mathbf{x} = \mathbf{b}$ pode ser resolvido em três etapas:

Etapa 1. Reordenação. Permute os elementos de $\mathbf{b}$ para obter $\mathbf{c} = P\mathbf{b}$.
Etapa 2. Substituição de Cima para Baixo. Resolva o sistema $L\mathbf{y} = \mathbf{c}$ para $\mathbf{y}$.
Etapa 3. Substituição de Baixo para Cima. Resolva $U\mathbf{x} = \mathbf{y}$.

EXEMPLO 2. Resolva o sistema

$$\begin{aligned} 6x_1 - 4x_2 + 2x_3 &= -2 \\ 4x_1 + 2x_2 + x_3 &= 4 \\ 2x_1 - x_2 + x_3 &= -1 \end{aligned}$$

SOLUÇÃO. A matriz de coeficientes desse sistema é a matriz A do Exemplo 1. P, L e U já foram determinadas e podem ser usadas para resolver o sistema da seguinte maneira:

Etapa 1. $\mathbf{c} = P\mathbf{b} = (-1, -2, 4)^T$

Etapa 2.
$$\begin{array}{rcl@{\qquad}l} y_1 & = & -1 & y_1 = -1 \\ 3y_1 + y_2 & = & -2 & y_2 = -2+3 = 1 \\ 2y_1 - 4y_2 + y_3 & = & 4 & y_3 = 4+2+4 = 10 \end{array}$$

Etapa 3.
$$\begin{array}{rcl@{\qquad}l} 2x_1 - x_2 + x_3 & = & -1 & x_1 = 1 \\ - x_2 - x_3 & = & 1 & x_2 = 1 \\ - 5x_3 & = & 10 & x_3 = -2 \end{array}$$

A solução do sistema é $\mathbf{x} = (1, 1, -2)^T$. □

É possível usar o método de Gauss sem troca de ordem entre linhas se os elementos diagonais $a_{ii}^{(i)}$ são diferentes de zero em cada etapa. No entanto, em aritmética de precisão finita, pivôs $a_{ii}^{(i)}$ próximos de 0 podem causar problemas.

EXEMPLO 3. Considere o sistema

$$\begin{aligned} 0{,}0001x_1 + 2x_2 &= 4 \\ x_1 + x_2 &= 3 \end{aligned}$$

A solução exata desse sistema é

$$\mathbf{x} = \left(\frac{2}{1{,}9999}, \frac{3{,}9997}{1{,}9999}\right)^T$$

Arredondando e usando quatro casas decimais, a solução fica $(1{,}0001, 1{,}9999)^T$. Vamos resolver o sistema usando aritmética de ponto flutuante com três dígitos decimais.

$$\left(\begin{array}{cc|c} 0{,}0001 & 2 & 4 \\ 1 & 1 & 3 \end{array}\right) \to \left(\begin{array}{cc|c} 0{,}0001 & 2 & 4 \\ 0 & -0{,}200 \times 10^5 & -0{,}400 \times 10^5 \end{array}\right)$$

A solução calculada é $\mathbf{x}' = (0, 2)^T$. Tem um erro de 100% na coordenada x_1. Por outro lado, se trocarmos a ordem das linhas para evitar o pivô próximo de zero, então a aritmética com três dígitos decimais nos dá

$$\left(\begin{array}{cc|c} 1 & 1 & 3 \\ 0{,}0001 & 2 & 4 \end{array}\right) \rightarrow \left(\begin{array}{cc|c} 1 & 1 & 3 \\ 0 & 2{,}00 & 4{,}00 \end{array}\right),$$

Nesse caso, a solução calculada é $\mathbf{x}' = (1, 2)^T$. □

Se o pivô $a_{ii}^{(i)}$ for pequeno, em valor absoluto, então os multiplicadores $m_{ki} = a_{ki}^{(i)}/a_{ii}^{(i)}$ serão grandes, em valor absoluto. Se existir um erro no valor calculado de $a_{ii}^{(i)}$, ele será multiplicado por m_{ki}. Em geral, multiplicadores com módulo grande contribuem para a propagação de erros. Por outro lado, multiplicadores com módulo menor do que 1 em geral retardam o crescimento de erro. Selecionando cuidadosamente os pivôs, podemos tentar evitar pivôs pequenos e, ao mesmo tempo, manter os multiplicadores com módulo menor do que 1. A estratégia mais comum para isso é uma técnica a que chamaremos *pivô parcial**.

PIVÔ PARCIAL

Na i-ésima etapa do método de Gauss, existem $n - i + 1$ candidatos a pivô:

$$a_{p(i),i},\ a_{p(i+1),i},\ \ldots,\ a_{p(n),i}$$

Escolha o candidato $a_{p(j),i}$ com módulo máximo,

$$|a_{p(j),i}| = \max_{i \le k \le n} |a_{p(k),i}|$$

e troque a ordem dos elementos i e j de $\mathbf{p}$. O elemento $a_{p(i),i}$ tem a propriedade

$$|a_{p(i),i}| \ge |a_{p(k),i}|$$

para $k = i + 1, \ldots, n$. Logo, todos os multiplicadores satisfazem

$$|m_{p(k),i}| = \left|\frac{a_{p(k),i}}{a_{p(i),i}}\right| \le 1$$

Poderíamos ir um passo adiante e usar o que chamaremos de *pivô total*: essa técnica consiste em escolher o pivô com módulo máximo entre todos os elementos das linhas e colunas restantes. Nesse caso, é preciso guardar tanto a ordem das linhas como a das colunas. Na i-ésima etapa escolhe-se o elemento $a_{p(j)q(k)}$ de modo que

$$|a_{p(j)q(k)}| = \max_{\substack{i \le s \le n \\ i \le t \le n}} |a_{p(s)q(t)}|$$

Permutam-se os elementos i e j de $\mathbf{p}$ e os elementos i e k de $\mathbf{q}$. O novo pivô vai ser $a_{p(i)q(i)}$. A maior desvantagem do pivô total é que, em cada etapa, precisamos procurar o pivô entre $(n - i + 1)^2$ elementos de A. Isso pode ser muito caro em termos de tempo de computador.

EXERCÍCIOS

1. Sejam

$$A = \begin{pmatrix} 0 & 3 & 1 \\ 1 & 2 & -2 \\ 2 & 5 & 4 \end{pmatrix} \quad \text{e} \quad \mathbf{b} = \begin{pmatrix} 1 \\ 7 \\ -1 \end{pmatrix}$$

(a) Coloque as linhas de $(A|\mathbf{b})$ na ordem (2, 3, 1) e depois resolva o sistema com essa nova ordem.
(b) Fatore A em um produto P^TLU, onde P é a matriz de permutação correspondente à nova ordem em (a).

*Essa terminologia não é padrão; em inglês, o termo utilizado é *partial pivoting*. (*N.T.*)

2. Seja A a matriz do Exercício 1. Use a fatoração P^TLU para resolver $A\mathbf{x} = \mathbf{c}$ para cada uma das escolhas de $\mathbf{c}$ a seguir.

(a) $(8, 1, 20)^T$ (b) $(-9, -2, -7)^T$ (c) $(4, 1, 11)^T$

3. Sejam

$$A = \begin{pmatrix} 1 & 8 & 6 \\ -1 & -4 & 5 \\ 2 & 4 & -6 \end{pmatrix} \quad \text{e} \quad \mathbf{b} = \begin{pmatrix} 8 \\ 1 \\ 4 \end{pmatrix}$$

Resolva o sistema $A\mathbf{x} = \mathbf{b}$ usando a técnica de pivô parcial. Se P é a matriz de permutação correspondente à estratégia usada, fatore PA em um produto LU.

4. Sejam

$$A = \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix} \quad \text{e} \quad \mathbf{b} = \begin{pmatrix} 5 \\ -2 \end{pmatrix}$$

Resolva o sistema $A\mathbf{x} = \mathbf{b}$ usando a técnica de pivô total. Seja P a matriz de permutação correspondente às permutações de linhas e Q a matriz de permutação determinada pelas colunas. Fatore PAQ em um produto LU.

5. Seja A a matriz do Exercício 4 e seja $\mathbf{c} = (6, -4)^T$. Resolva o sistema $A\mathbf{x} = \mathbf{c}$ em duas etapas:
(a) Defina $\mathbf{z} = Q^T\mathbf{x}$ e resolva $LU\mathbf{z} = P\mathbf{c}$ para $\mathbf{z}$;
(b) Calcule $\mathbf{x} = Q\mathbf{z}$.

6. Dados

$$A = \begin{pmatrix} 5 & 4 & 7 \\ 2 & -4 & 3 \\ 2 & 8 & 6 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 2 \\ -5 \\ 4 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} 5 \\ -4 \\ 2 \end{pmatrix}$$

(a) Resolva o sistema $A\mathbf{x} = \mathbf{b}$ usando a técnica de pivô total.
(b) Seja P a matriz de permutação determinada pelas linhas dos pivôs e Q a matriz de permutação determinada pelas colunas dos pivôs. Fatore PAQ em um produto LU.
(c) Use a fatoração LU do item (b) para resolver o sistema $A\mathbf{x} = \mathbf{c}$.

7. A solução exata do sistema

$$0{,}6000x_1 + 2000x_2 = 2003$$

$$0{,}3076x_1 - 0{,}4010x_2 = 1{,}137$$

é $\mathbf{x} = (5, 1)^T$. Suponha que o valor calculado de x_2 é $x_2' = 1 + \varepsilon$. Use esse valor na primeira equação e resolva para x_1. Qual vai ser o erro? Calcule o erro relativo em x_1 se $\varepsilon = 0{,}001$.

8. Resolva o sistema no Exercício 7 usando aritmética de ponto flutuante com quatro dígitos decimais e o método de Gauss com pivô parcial.

9. Resolva o sistema no Exercício 7 usando aritmética de ponto flutuante com quatro dígitos decimais e o método de Gauss com pivô total.

10. Use aritmética de ponto flutuante com quatro dígitos decimais e mude a escala do sistema no Exercício 7 multiplicando a primeira equação por 1/2000 e a segunda por 1/0,4010. Resolva o sistema nessa nova escala usando a técnica de pivô parcial.

4 NORMAS DE MATRIZES E NÚMEROS CONDICIONAIS

Nesta seção, vamos nos preocupar com a precisão das soluções calculadas de sistemas lineares. Qual a precisão que podemos esperar das soluções calculadas e como podemos testar essa precisão? A resposta a essas perguntas depende da sensibilidade da matriz de coeficientes do sistema a pequenas perturbações. A sensibilidade da matriz pode ser medida através do seu *número condicional.* O número condicional de uma matriz invertível é definido em termos de sua norma e da norma de sua inversa. Antes de discutirmos números condicionais, precisamos estabelecer alguns resultados importantes sobre alguns tipos canônicos de normas de matrizes.

NORMAS DE MATRIZES

Assim como normas de vetores são usadas para se medir o tamanho de vetores, normas de matrizes podem ser usadas para se medir o tamanho das matrizes. Na Seção 3 do Cap. 5 definimos uma norma em $R^{m\times n}$ induzida por um produto interno em $R^{m\times n}$. Essa norma foi chamada de norma de Frobenius e denotada por $\|\cdot\|_F$. Mostramos que a norma de Frobenius de uma matriz A podia ser calculada pela raiz quadrada da soma dos quadrados de todos os seus elementos.

(1) $$\|A\|_F = \left(\sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}^2\right)^{1/2}$$

Na verdade, a equação (1) define uma família de normas de matrizes, já que define uma norma em $R^{m\times n}$ para qualquer escolha de m e n. A norma de Frobenius tem propriedades importantes que listamos a seguir.

I. Se $\mathbf{a}_j$ representa a j-ésima coluna de A, então

$$\|A\|_F = \left(\sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}^2\right)^{1/2} = \left(\sum_{j=1}^{n} \|\mathbf{a}_j\|_2^2\right)^{1/2}$$

II. Se $\mathbf{a}(i,:)$ representa a i-ésima linha de A, então

$$\|A\|_F = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right)^{1/2} = \left(\sum_{i=1}^{m} \|\mathbf{a}(i,:)^T\|_2^2\right)^{1/2}$$

III. Se $\mathbf{x} \in R^n$, então

$$\begin{aligned}\|A\mathbf{x}\|_2 &= \left[\sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j\right)^2\right]^{1/2} = \left[\sum_{i=1}^{m} (\mathbf{a}(i,:)\mathbf{x})^2\right]^{1/2} \\ &\le \left[\sum_{i=1}^{m} \|\mathbf{x}\|_2^2 \|\mathbf{a}(i,:)^T\|_2^2\right]^{1/2} \qquad \text{(Cauchy–Schwarz)} \\ &= \|A\|_F\,\|\mathbf{x}\|_2\end{aligned}$$

IV. Se $B = (\mathbf{b}_1, \ldots, \mathbf{b}_r)$ é uma matriz $n \times r$, então, pelas propriedades I e III,

$$\begin{aligned}\|AB\|_F &= \|(A\mathbf{b}_1, A\mathbf{b}_2, \ldots, A\mathbf{b}_r)\|_F \\ &= \left(\sum_{i=1}^{r} \|A\mathbf{b}_i\|_2^2\right)^{1/2} \\ &\le \|A\|_F \left(\sum_{i=1}^{r} \|\mathbf{b}_i\|_2^2\right)^{1/2} \\ &= \|A\|_F\|B\|_F\end{aligned}$$

Existem muitas outras normas que poderíamos usar para $R^{m\times n}$, além da norma de Frobenius. Qualquer norma usada tem que satisfazer as três condições que definem normas em geral.

(i) $\|A\| \geq 0$ e $\|A\| = 0$ se e somente se $= O$
(ii) $\|\alpha A\| = |\alpha|\|A\|$
(iii) $\|A + B\| \leq \|A\| + \|B\|$

As famílias de normas mais úteis satisfazem a propriedade adicional

(iv) $\|AB\| \leq \|A\|\ \|B\|$

Vamos, então, considerar apenas famílias de normas que satisfazem essa propriedade. Uma conseqüência importante da propriedade (iv) é

$$\|A^n\| \leq \|A\|^n$$

Em particular, se $\|A\| < 1$, então $\|A^n\| \to 0$ quando $n \to \infty$.

Em geral, uma norma matricial $\|\cdot\|_M$ em $R^{m\times n}$ e uma norma vetorial $\|\cdot\|_V$ em R^n são *compatíveis* se

$$\|A\mathbf{x}\|_V \leq \|A\|_M\|\mathbf{x}\|_V$$

para todo $x \in R^n$. Em particular, pela propriedade III, a norma matricial $\|\cdot\|_F$ e a norma vetorial $\|\cdot\|_2$ são compatíveis. Para cada uma das normas vetoriais canônicas podemos definir uma norma matricial compatível usando a norma vetorial para calcular uma norma de operador para a matriz. A norma matricial definida dessa maneira é dita *subordinada* à norma vetorial.

NORMAS MATRICIAIS SUBORDINADAS

Podemos pensar em cada matriz $m \times n$ como um operador linear de R^n em R^m. Para qualquer família de normas vetoriais, podemos definir uma *norma de operador* comparando $\|A\mathbf{x}\|$ e $\|\mathbf{x}\|$ para cada $\mathbf{x}$ não-nulo e definindo

$$\|A\| = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \tag{2}$$

Pode-se mostrar que existe um $\mathbf{x}_0$ particular em R^n que maximiza $\|A\mathbf{x}\| / \|\mathbf{x}\|$, mas a demonstração está além do escopo deste livro. Supondo que $\|A\mathbf{x}\| / \|\mathbf{x}\|$ sempre pode ser maximizada, vamos mostrar que (2) define, de fato, uma norma em $R^{m\times n}$. Para isso, precisamos verificar que cada uma das três condições é satisfeita.

(i) Para cada $\mathbf{x} \neq \mathbf{0}$,

$$\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \geq 0$$

e, portanto,

$$\|A\| = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \geq 0$$

Se $\|A\| = 0$, então $A\mathbf{x} = \mathbf{0}$ para todo $\mathbf{x} \in R^n$. Isso implica que

$$\mathbf{a}_j = A\mathbf{e}_j = \mathbf{0} \qquad \text{para} \qquad j = 1, \ldots, n$$

e, portanto, A tem que ser a matriz nula.

(ii) $\displaystyle \|\alpha A\| = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|\alpha A\mathbf{x}\|}{\|\mathbf{x}\|} = |\alpha| \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = |\alpha|\ \|A\|$

(iii) Se $\mathbf{x} \neq \mathbf{0}$, então

$$\begin{aligned}\|A+B\| &= \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|(A+B)\mathbf{x}\|}{\|\mathbf{x}\|} \\ &\leq \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|+\|B\mathbf{x}\|}{\|\mathbf{x}\|} \\ &\leq \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} + \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|B\mathbf{x}\|}{\|\mathbf{x}\|} \\ &= \|A\| + \|B\|\end{aligned}$$

Logo, (2) define uma norma em $R^{m\times n}$. Para cada família de normas vetoriais $\|\cdot\|$, podemos, então, definir uma família de normas matriciais por (2). As normas matriciais definidas por (2) são ditas *subordinadas* às normas vetoriais $\|\cdot\|$.

Teorema 7.4.1. *Se a família de normas matriciais $\|\cdot\|_M$ é subordinada à família de normas vetoriais $\|\cdot\|_V$, então $\|\cdot\|_M$ e $\|\cdot\|_V$ são compatíveis e as normas matriciais $\|\cdot\|_M$ satisfazem a propriedade* (iv).

Demonstração. Se $\mathbf{x}$ é um vetor qualquer não-nulo em R^n, então

$$\frac{\|A\mathbf{x}\|_V}{\|\mathbf{x}\|_V} \leq \max_{\mathbf{y}\neq\mathbf{0}} \frac{\|A\mathbf{y}\|_V}{\|\mathbf{y}\|_V} = \|A\|_M$$

e, portanto,

$$\|A\mathbf{x}\|_V \leq \|A\|_M\|\mathbf{x}\|_V$$

Como esta última desigualdade também é válida para $\mathbf{x} = \mathbf{0}$, temos que $\|\cdot\|_M$ e $\|\cdot\|_V$ são compatíveis. Se B é uma matriz $n \times r$, como $\|\cdot\|_M$ e $\|\cdot\|_V$ são compatíveis, temos

$$\|AB\mathbf{x}\|_V \leq \|A\|_M\|B\mathbf{x}\|_V \leq \|A\|_M\|B\|_M\|\mathbf{x}\|_V$$

Logo, para todo $\mathbf{x} \neq \mathbf{0}$,

$$\frac{\|AB\mathbf{x}\|_V}{\|\mathbf{x}\|_V} \leq \|A\|_M\|B\|_M$$

e, portanto,

$$\|AB\|_M = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|AB\mathbf{x}\|_V}{\|\mathbf{x}\|_V} \leq \|A\|_M\|B\|_M \qquad \square$$

É fácil calcular a norma de Frobenius de uma matriz. Por exemplo, se

$$A = \begin{pmatrix} 4 & 2 \\ 0 & 4 \end{pmatrix}$$

então

$$\|A\|_F = (4^2 + 0^2 + 2^2 + 4^2)^{1/2} = 6$$

Por outro lado, não é tão óbvio como calcular $\|A\|$ se $\|\cdot\|$ é uma norma matricial subordinada. A norma matricial

$$\|A\|_2 = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$$

é difícil de calcular, mas as normas

$$\|A\|_1 = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|_1}{\|\mathbf{x}\|_1}$$

e

$$\|A\|_\infty = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty}$$

podem ser calculadas com facilidade.

Teorema 7.4.2. *Se A é uma matriz m × n, então*

$$\|A\|_1 = \max_{1\le j\le n}\left(\sum_{i=1}^{m}|a_{ij}|\right)$$

e

$$\|A\|_\infty = \max_{1\le i\le m}\left(\sum_{j=1}^{n}|a_{ij}|\right)$$

Demonstração. Vamos provar que

$$\|A\|_1 = \max_{1\le j\le n}\left(\sum_{i=1}^{m}|a_{ij}|\right)$$

e deixar a demonstração da segunda igualdade como exercício. Seja

$$\alpha = \max_{1\le j\le n}\sum_{i=1}^{m}|a_{ij}| = \sum_{i=1}^{m}|a_{ik}|$$

Note que k é o índice da coluna onde ocorre o máximo. Seja $\mathbf{x}$ um vetor arbitrário em R^n. Então

$$A\mathbf{x} = \left(\sum_{j=1}^{n}a_{1j}x_j, \sum_{j=1}^{n}a_{2j}x_j, \dots, \sum_{j=1}^{n}a_{mj}x_j\right)^T$$

e temos que

$$\begin{aligned}
\|A\mathbf{x}\|_1 &= \sum_{i=1}^{m}\left|\sum_{j=1}^{n}a_{ij}x_j\right| \\
&\le \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}x_j| \\
&= \sum_{j=1}^{n}\left(|x_j|\sum_{i=1}^{m}|a_{ij}|\right) \\
&\le \alpha\sum_{j=1}^{n}|x_j| \\
&= \alpha\|\mathbf{x}\|_1
\end{aligned}$$

Logo, qualquer que seja o vetor não-nulo $\mathbf{x}$ em R^n,

$$\frac{\|A\mathbf{x}\|_1}{\|\mathbf{x}\|_1} \le \alpha$$

e, portanto,

$$\|A\|_1 = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|_1}{\|\mathbf{x}\|_1} \leq \alpha \tag{3}$$

Por outro lado,

$$\|A\mathbf{e}_k\|_1 = \|\mathbf{a}_k\|_1 = \alpha$$

Como $\|\mathbf{e}_k\|_1 = 1$, temos que

$$\|A\|_1 = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|_1}{\|\mathbf{x}\|_1} \geq \frac{\|A\mathbf{e}_k\|_1}{\|\mathbf{e}_k\|_1} = \alpha \tag{4}$$

As equações (3) e (4) juntas implicam que $\|A\|_1 = \alpha$. □

EXEMPLO 1. Seja

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

Então

$$\|A\|_1 = |4| + |-3| + |-6| + |1| = 14$$

e

$$\|A\|_\infty = |5| + |-2| + |-3| + |5| = 15$$

□

NÚMEROS CONDICIONAIS

Normas matriciais podem ser usadas para se estimar a sensibilidade de sistemas lineares a pequenas perturbações na matriz de coeficientes. Considere o seguinte exemplo.

EXEMPLO 2. Resolva o seguinte sistema linear:

$$\begin{aligned} 2{,}0000x_1 + 2{,}0000x_2 &= 6{,}0000 \\ 2{,}0000x_1 + 2{,}0005x_2 &= 6{,}0010 \end{aligned} \tag{5}$$

Se usarmos aritmética de ponto flutuante com cinco casas decimais, a solução calculada será a solução exata $\mathbf{x} = (1, 2)^T$. Suponha, no entanto, que somos forçados a usar apenas quatro casas decimais. Então, em vez de (5), temos

$$\begin{aligned} 2{,}000x_1 + 2{,}000x_2 &= 6{,}000 \\ 2{,}000x_1 + 2{,}001x_2 &= 6{,}001 \end{aligned} \tag{6}$$

A solução calculada do sistema (6) vai ser a solução exata $\mathbf{x}' = (2, 1)^T$.

Os sistemas (5) e (6) são iguais exceto pelo coeficiente a_{22}. O erro relativo nesse coeficiente é

$$\frac{a'_{22} - a_{22}}{a_{22}} \approx 0{,}00025$$

No entanto, os erros relativos nas coordenadas das soluções $\mathbf{x}$ e $\mathbf{x}'$ são

$$\frac{x'_1 - x_1}{x_1} = 1{,}0 \qquad \text{e} \qquad \frac{x'_2 - x_2}{x_2} = -0{,}5$$

□

Definição. Uma matriz A é **mal condicionada** se mudanças relativamente pequenas em seus elementos podem causar mudanças relativamente grandes nas soluções de $A\mathbf{x} = \mathbf{b}$. A é **bem condicionada** se mudanças relativamente pequenas em seus elementos resultam em mudanças relativamente pequenas nas soluções de $A\mathbf{x} = \mathbf{b}$.

Se a matriz A for mal condicionada, a solução calculada para $A\mathbf{x} = \mathbf{b}$ em geral não vai ser muito precisa. Mesmo que os elementos de A possam ser representados exatamente como números em ponto flutuante, pequenos erros de arredondamento durante o processo de resolução podem ter efeitos drásticos na solução calculada. Por outro lado, se a matriz for bem condicionada e se a estratégia adequada para a escolha do pivô for usada, devemos ser capazes de calcular a solução com bastante precisão. Em geral, a precisão da solução calculada depende se a matriz é ou não bem condicionada. Se pudéssemos medir o condicionamento de A, essa medida poderia ser usada para se encontrar uma cota superior para o erro relativo na solução calculada.

Seja A uma matriz invertível $n \times n$ e considere o sistema $A\mathbf{x} = \mathbf{b}$. Se $\mathbf{x}$ é a solução exata do sistema e $\mathbf{x}'$ é a solução calculada, então o erro pode ser representado pelo vetor $\mathbf{e} = \mathbf{x} - \mathbf{x}'$. Se $\|\cdot\|$ é uma norma em R^n, então $\|\mathbf{e}\|$ é uma medida do erro absoluto e $\|\mathbf{e}\| / \|\mathbf{x}\|$ é uma medida do erro relativo. Em geral, não temos nenhuma maneira de determinar valores exatos para $\|\mathbf{e}\|$ e $\|\mathbf{e}\| / \|\mathbf{x}\|$. Uma possibilidade para testar a precisão de $\mathbf{x}'$ é colocá-lo no sistema original e ver o quão próximo $\mathbf{b}' = A\mathbf{x}'$ está de $\mathbf{b}$. O vetor

$$\mathbf{r} = \mathbf{b} - \mathbf{b}' = \mathbf{b} - A\mathbf{x}'$$

é chamado de *resíduo* e pode ser calculado facilmente. A razão

$$\frac{\|\mathbf{b} - A\mathbf{x}'\|}{\|\mathbf{b}\|} = \frac{\|\mathbf{r}\|}{\|\mathbf{b}\|}$$

é chamada de *resíduo relativo*. O resíduo relativo é uma boa estimativa para o erro relativo? A resposta depende do condicionamento de A. No Exemplo 2 o resíduo para a solução calculada $\mathbf{x}' = (2, 1)^T$ é

$$\mathbf{r} = \mathbf{b} - A\mathbf{x}' = (0, 0{,}0005)^T$$

O resíduo relativo em relação à norma ∞ é

$$\frac{\|\mathbf{r}\|_\infty}{\|\mathbf{b}\|_\infty} = \frac{0{,}0005}{6{,}0010} \approx 0{,}000083$$

e o erro relativo é dado por

$$\frac{\|\mathbf{e}\|_\infty}{\|\mathbf{x}\|_\infty} = 0{,}5$$

O erro relativo é mais do que 6000 vezes o resíduo relativo. Vamos mostrar que, em geral, se A for mal condicionada, o resíduo relativo pode ser muito menor do que o erro relativo. Por outro lado, para matrizes bem condicionadas, o erro relativo e o resíduo relativo são bastante próximos. Para mostrar isso, precisaremos usar normas matriciais. Lembre-se de que, se $\|\cdot\|$ é uma norma compatível em $R^{n\times n}$, então, quaisquer que sejam a matriz C $n \times n$ e o vetor $\mathbf{y} \in R^n$, temos

$$\|C\mathbf{y}\| \le \|C\|\,\|\mathbf{y}\| \tag{7}$$

Temos

$$\mathbf{r} = \mathbf{b} - A\mathbf{x}' = A\mathbf{x} - A\mathbf{x}' = A\mathbf{e}$$

e, portanto,

$$\mathbf{e} = A^{-1}\mathbf{r}$$

Da propriedade (7),

$$\|\mathbf{e}\| \le \|A^{-1}\|\,\|\mathbf{r}\|$$

e

$$\|\mathbf{r}\| = \|A\mathbf{e}\| \leq \|A\| \, \|\mathbf{e}\|$$

Logo,

$$\frac{\|\mathbf{r}\|}{\|A\|} \leq \|\mathbf{e}\| \leq \|A^{-1}\| \, \|\mathbf{r}\| \tag{8}$$

Como $\mathbf{x}$ é a solução exata de $A\mathbf{x} = \mathbf{b}$, $\mathbf{x} = A^{-1}\mathbf{b}$. Pelo mesmo argumento que usamos para obter (8), temos

$$\frac{\|\mathbf{b}\|}{\|A\|} \leq \|\mathbf{x}\| \leq \|A^{-1}\| \, \|\mathbf{b}\| \tag{9}$$

De (8) e (9), temos que

$$\frac{1}{\|A\| \, \|A^{-1}\|} \frac{\|\mathbf{r}\|}{\|\mathbf{b}\|} \leq \frac{\|\mathbf{e}\|}{\|\mathbf{x}\|} \leq \|A\| \, \|A^{-1}\| \frac{\|\mathbf{r}\|}{\|\mathbf{b}\|}$$

O número $\|A\| \, \|A^{-1}\|$ é chamado de *número condicional* de A e será denotado por cond(A). Temos

$$\frac{1}{\text{cond}(A)} \frac{\|\mathbf{r}\|}{\|\mathbf{b}\|} \leq \frac{\|\mathbf{e}\|}{\|\mathbf{x}\|} \leq \text{cond}(A) \frac{\|\mathbf{r}\|}{\|\mathbf{b}\|} \tag{10}$$

A desigualdade (10) mostra a relação entre o tamanho do erro relativo $\|\mathbf{e}\| / \|\mathbf{x}\|$ e o resíduo relativo $\|\mathbf{r}\| / \|\mathbf{b}\|$. Se o número condicional estiver próximo de 1, o erro relativo e o resíduo relativo estarão próximos. Se o número condicional for grande, o erro relativo pode ser muitas vezes maior do que o resíduo relativo.

EXEMPLO 3. Seja

$$A = \begin{pmatrix} 3 & 3 \\ 4 & 5 \end{pmatrix}$$

Então

$$A^{-1} = \frac{1}{3} \begin{pmatrix} 5 & -3 \\ -4 & 3 \end{pmatrix}$$

$\|A\|_\infty = 9$ e $\|A^{-1}\|_\infty = 8/3$. (Usamos $\|\cdot\|_\infty$ porque é fácil de calcular.) Logo,

$$\text{cond}_\infty(A) = 9 \cdot \tfrac{8}{3} = 24$$

Teoricamente, o erro relativo na solução calculada de um sistema $A\mathbf{x} = \mathbf{b}$ poderia ser até 24 vezes maior do que o resíduo relativo. □

EXEMPLO 4. Suponha que $\mathbf{x}' = (2{,}0,\ 0{,}1)^T$ é a solução calculada para

$$3x_1 + 3x_2 = 6$$
$$4x_1 + 5x_2 = 9$$

Determine o resíduo $\mathbf{r}$ e o resíduo relativo $\|\mathbf{r}\|_\infty / \|\mathbf{b}\|_\infty$.

SOLUÇÃO.

$$\mathbf{r} = \begin{pmatrix} 6 \\ 9 \end{pmatrix} - \begin{pmatrix} 3 & 3 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} 2{,}0 \\ 0{,}1 \end{pmatrix} = \begin{pmatrix} -0{,}3 \\ 0{,}5 \end{pmatrix}$$

$$\frac{\|\mathbf{r}\|_\infty}{\|\mathbf{b}\|_\infty} = \frac{0{,}5}{9} = \frac{1}{18} \qquad \square$$

Podemos ver, analisando o sistema no exemplo anterior, que a solução de fato é $\mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. O erro $\mathbf{e}$ é dado por

$$\mathbf{e} = \mathbf{x} - \mathbf{x}' = \begin{pmatrix} -1{,}0 \\ 0{,}9 \end{pmatrix}$$

e o erro relativo é dado por

$$\frac{\|\mathbf{e}\|_\infty}{\|\mathbf{x}\|_\infty} = \frac{1{,}0}{1} = 1$$

O erro relativo é 18 vezes o resíduo relativo. Isso não surpreende, já que $\operatorname{cond}(A) = 24$. Usando $\|\cdot\|_1$, obtemos resultados análogos. Nesse caso,

$$\frac{\|\mathbf{r}\|_1}{\|\mathbf{b}\|_1} = \frac{0{,}8}{15} = \frac{4}{75} \qquad \text{e} \qquad \frac{\|\mathbf{e}\|_1}{\|\mathbf{x}\|_1} = \frac{1{,}9}{2} = \frac{19}{20}$$

O número condicional de uma matriz invertível nos dá, de fato, informação valiosa sobre o condicionamento de A. Seja A' a nova matriz formada por pequenas alterações nos elementos de A. Seja $E = A' - A$. Então $A' = A + E$, onde os elementos de E são pequenos, se comparados com os elementos de A. A vai ser mal condicionada se para um tal E as soluções de $A'\mathbf{x} = \mathbf{b}$ e de $A\mathbf{x} = \mathbf{b}$ são muito diferentes. Sejam $\mathbf{x}'$ a solução de $A'\mathbf{x} = \mathbf{b}$ e $\mathbf{x}$ a solução de $A\mathbf{x} = \mathbf{b}$. O número condicional nos permite comparar a variação na solução, em relação a $\mathbf{x}'$, com a mudança relativa na matriz A.

$$\mathbf{x} = A^{-1}\mathbf{b} = A^{-1}A'\mathbf{x}' = A^{-1}(A + E)\mathbf{x}' = \mathbf{x}' + A^{-1}E\mathbf{x}'$$

Logo,

$$\mathbf{x} - \mathbf{x}' = A^{-1}E\mathbf{x}'$$

Usando a desigualdade (7), vemos que

$$\|\mathbf{x} - \mathbf{x}'\| \le \|A^{-1}\|\,\|E\|\,\|\mathbf{x}'\|$$

ou

$$\text{(11)} \qquad \frac{\|\mathbf{x} - \mathbf{x}'\|}{\|\mathbf{x}'\|} \le \|A^{-1}\|\,\|E\| = \operatorname{cond}(A)\frac{\|E\|}{\|A\|}$$

Vamos voltar ao Exemplo 2 e ver como usar a desigualdade (11). Sejam A e A' as duas matrizes de coeficientes no Exemplo 2.

$$E = A' - A = \begin{pmatrix} 0 & 0 \\ 0 & 0{,}0005 \end{pmatrix}$$

e

$$A^{-1} = \begin{pmatrix} 2000{,}5 & -2000 \\ -2000 & 2000 \end{pmatrix}$$

Em termos da norma ∞, o erro relativo em A é

$$\frac{\|E\|_\infty}{\|A\|_\infty} = \frac{0{,}0005}{4{,}0005} \approx 0{,}0001$$

e o número condicional é

$$\text{cond}(A) = \|A\|_\infty \, \|A^{-1}\|_\infty = (4{,}0005)(4000{,}5) \approx 16.004$$

A cota superior para o erro relativo dado em (11) é, então,

$$\text{cond}(A)\frac{\|E\|}{\|A\|} = \|A^{-1}\| \, \|E\| = (4000{,}5)(0{,}0005) \approx 2$$

O erro relativo de fato para o sistema no Exemplo 2 é

$$\frac{\|\mathbf{x} - \mathbf{x}'\|_\infty}{\|\mathbf{x}'\|_\infty} = \frac{1}{2}$$

EXERCÍCIOS

1. Determine $\|\cdot\|_F$, $\|\cdot\|_\infty$ e $\|\cdot\|_1$ para cada uma das matrizes a seguir.

$$\text{(a)} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{(b)} \begin{pmatrix} 1 & 4 \\ -2 & 2 \end{pmatrix} \quad \text{(c)} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$\text{(d)} \begin{pmatrix} 0 & 5 & 1 \\ 2 & 3 & 1 \\ 1 & 2 & 2 \end{pmatrix} \quad \text{(e)} \begin{pmatrix} 5 & 0 & 5 \\ 4 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix}$$

2. Sejam

$$A = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \quad \text{e} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Escreva $\|A\mathbf{x}\|_2 / \|\mathbf{x}\|_2$ em termos de x_1 e x_2. Encontre o valor de $\|A\|_2$.

3. Seja

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Mostre que $\|A\|_2 = 1$.

4. Seja I a matriz identidade $n \times n$. Determine os valores de $\|I\|_1$, $\|I\|_\infty$ e $\|I\|_F$.

5. Sejam $\|\cdot\|_M$ uma norma em $R^{n\times n}$, $\|\cdot\|_V$ uma norma em R^n e I a matriz identidade $n \times n$. Mostre que:

(a) Se $\|\cdot\|_M$ e $\|\cdot\|_V$ são compatíveis, então $\|I\|_M \geq 1$;

(b) Se $\|\cdot\|_M$ é subordinada a $\|\cdot\|_V$, então $\|I\|_M = 1$.

6. Seja

$$A = \begin{pmatrix} 3 & -1 & -2 \\ -1 & 2 & -7 \\ 4 & 1 & 4 \end{pmatrix}$$

(a) Determine $\|A\|_\infty$.

(b) Encontre um vetor $\mathbf{x}$ com coordenadas ± 1 tal que $\|A\mathbf{x}\|_\infty = \|A\|_\infty$. (Note que $\|\mathbf{x}\|_\infty = 1$, de modo que $\|A\|_\infty = \|A\mathbf{x}\|_\infty / \|\mathbf{x}\|_\infty$.)

7. O Teorema 7.4.2 diz que

$$\|A\|_\infty = \max_{1\leq i\leq m} \left(\sum_{j=1}^{n} |a_{ij}| \right)$$

Prove isso em duas etapas.

(a) Mostre, primeiro, que

$$\|A\|_\infty \leq \max_{1\leq i\leq m}\left(\sum_{j=1}^{n}|a_{ij}|\right)$$

(b) Construa um vetor $\mathbf{x}$ com coordenadas ± 1 tal que

$$\frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} = \|A\mathbf{x}\|_\infty = \max_{1\leq i\leq m}\left(\sum_{j=1}^{n}|a_{ij}|\right)$$

8. Mostre que $\| A \|_F = \| A^T \|_F$.

9. Seja A uma matriz simétrica $n \times n$. Mostre que $\| A \|_\infty = \| A \|_1$.

10. Seja $\| \cdot \|$ uma família de normas vetoriais e seja $\| \cdot \|_M$ a família de normas matriciais subordinada. Mostre que

$$\|A\|_M = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|$$

11. Seja A uma matriz $n \times n$ e seja $\| \cdot \|_M$ uma norma matricial compatível com alguma norma vetorial em R^n. Se λ é um autovalor de A, mostre que $|\lambda| \leq \| A \|_M$.

12. Sejam A uma matriz $n \times n$ e $\mathbf{x} \in R^n$. Prove que:

(a) $\|A\mathbf{x}\|_\infty \leq n^{1/2}\|A\|_2\|\mathbf{x}\|_\infty$

(b) $\|A\mathbf{x}\|_2 \leq n^{1/2}\|A\|_\infty\|\mathbf{x}\|_2$

(c) $n^{-1/2}\|A\|_2 \leq \|A\|_\infty \leq n^{1/2}\|A\|_2$

13. Seja A uma matriz simétrica $n \times n$ com autovetores ortonormais $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Sejam $\mathbf{x} \in R^n$ e $c_i = \mathbf{u}_i^T\mathbf{x}$ para $i = 1, 2, \ldots, n$. Mostre que

(a) $\|A\mathbf{x}\|_2^2 = \sum_{i=1}^{n}(\lambda_i c_i)^2$

(b) Se $\mathbf{x} \neq \mathbf{0}$, então

$$\min_{1\leq i\leq n}|\lambda_i| \leq \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \leq \max_{1\leq i\leq n}|\lambda_i|$$

(c) $\|A\|_2 = \max_{1\leq i\leq n}|\lambda_i|$

14. Seja

$$A = \begin{pmatrix} 1 & -0{,}99 \\ -1 & 1 \end{pmatrix}$$

Encontre A^{-1} e $\text{cond}_\infty(A)$.

15. Resolva os dois sistemas a seguir e compare suas soluções. As matrizes de coeficientes são bem condicionadas? Mal condicionadas? Explique.

$$\begin{aligned} 1{,}0x_1 + 2{,}0x_2 &= 1{,}12 \qquad & 1{,}000x_1 + 2{,}011x_2 &= 1{,}120 \\ 2{,}0x_1 + 3{,}9x_2 &= 2{,}16 \qquad & 2{,}000x_1 + 3{,}982x_2 &= 2{,}160 \end{aligned}$$

16. Seja

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

Calcule $\text{cond}_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty$.

17. Seja A uma matriz invertível $n \times n$ e seja $\|\cdot\|_M$ uma norma matricial compatível com alguma norma vetorial em R^n. Mostre que

$$\text{cond}_M(A) \geq 1$$

18. Seja

$$A_n = \begin{pmatrix} 1 & 1 \\ 1 & 1 - \frac{1}{n} \end{pmatrix}$$

para cada inteiro positivo n. Calcule:

(a) A_n^{-1}
(b) $\text{cond}_\infty(A_n)$
(c) $\lim\limits_{n\to\infty} \text{cond}_\infty(A_n)$

19. Sejam

$$A = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} \quad \text{e} \quad \mathbf{b} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$$

A solução calculada com aritmética de ponto flutuante com duas casas decimais é $\mathbf{x} = (1{,}1, 0{,}88)^T$.
(a) Determine o vetor resíduo $\mathbf{r}$ e o valor do resíduo relativo $\|\mathbf{r}\|_\infty / \|\mathbf{b}\|_\infty$.
(b) Encontre o valor de $\text{cond}_\infty(A)$.
(c) Sem calcular a solução exata, use os resultados encontrados em (a) e (b) para obter cotas para o erro relativo na solução calculada.
(d) Calcule a solução exata $\mathbf{x}$ e determine o erro relativo. Compare esse valor com as cotas encontradas no item (c).

20. Seja

$$A = \begin{pmatrix} -0{,}50 & 0{,}75 & -0{,}25 \\ -0{,}50 & 0{,}25 & 0{,}25 \\ 1{,}00 & -0{,}50 & 0{,}50 \end{pmatrix}$$

Calcule $\text{cond}_1(A) = \|A\|_1 \|A^{-1}\|_1$.

21. Seja A a matriz do Exercício 20 e seja

$$A' = \begin{pmatrix} -0{,}5 & 0{,}8 & -0{,}3 \\ -0{,}5 & 0{,}3 & 0{,}3 \\ 1{,}0 & -0{,}5 & 0{,}5 \end{pmatrix}$$

Sejam $\mathbf{x}$ e $\mathbf{x}'$ as soluções de $A\mathbf{x} = \mathbf{b}$ e $A'\mathbf{x} = \mathbf{b}$, respectivamente, para algum $\mathbf{b} \in R^n$. Encontre uma cota superior para o erro relativo $(\|\mathbf{x} - \mathbf{x}'\|_1)/\|\mathbf{x}'\|_1$.

22. Sejam

$$A = \begin{pmatrix} 1 & -1 & -1 & -1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 5{,}00 \\ 1{,}02 \\ 1{,}04 \\ 1{,}10 \end{pmatrix}$$

Calcula-se uma solução aproximada para $A\mathbf{x} = \mathbf{b}$ arredondando-se os elementos de $\mathbf{b}$ para o inteiro mais próximo e depois resolvendo o sistema obtido com aritmética inteira. A solução calculada é $\mathbf{x}' = (12, 4, 2, 1)^T$. Denote por $\mathbf{r}$ o vetor resíduo.
(a) Determine os valores de $\|\mathbf{r}\|_\infty$ e $\text{cond}_\infty(A)$.

(b) Use a resposta encontrada em (a) para achar uma cota superior para o erro relativo na solução.

(c) Calcule a solução exata $\mathbf{x}$ e determine o erro relativo

$$\frac{\|\mathbf{x}-\mathbf{x}'\|_\infty}{\|\mathbf{x}\|_\infty}$$

23. Sejam A e B matrizes invertíveis $n \times n$. Mostre que

$$\operatorname{cond}(AB) \leq \operatorname{cond}(A)\operatorname{cond}(B)$$

24. Seja D uma matriz invertível diagonal $n \times n$. Mostre que

$$\operatorname{cond}_1(D) = \operatorname{cond}_\infty(D) = \frac{d_{\max}}{d_{\min}}$$

onde

$$d_{\max} = \max_{1\leq i\leq n} |d_{ii}| \quad \text{e} \quad d_{\min} = \min_{1\leq i\leq n} |d_{ii}|$$

25. Sejam D, $d_{\max}$ e $d_{\min}$ como no Exercício 24. Mostre que

(a) $\|D\|_2 = d_{\max}$

(b) $\operatorname{cond}_2(D) = \dfrac{d_{\max}}{d_{\min}}$

26. Seja Q uma matriz ortogonal $n \times n$. Mostre que:

(a) $\|Q\|_2 = 1$

(b) $\operatorname{cond}_2(Q) = 1$

(c) Para qualquer $\mathbf{b} \in R^n$, o erro relativo na solução de $Q\mathbf{x} = \mathbf{b}$ é igual ao resíduo relativo, isto é,

$$\frac{\|\mathbf{e}\|_2}{\|\mathbf{x}\|_2} = \frac{\|\mathbf{r}\|_2}{\|\mathbf{b}\|_2}$$

27. Sejam A uma matriz $n \times n$ e Q e V matrizes ortogonais $n \times n$. Mostre que:

(a) $\|QA\|_2 = \|A\|_2$

(b) $\|AV\|_2 = \|A\|_2$

(c) $\|QAV\|_2 = \|A\|_2$

28. Sejam A uma matriz invertível $n \times n$ e Q uma matriz ortogonal $n \times n$. Mostre que:

(a) $\operatorname{cond}_2(QA) = \operatorname{cond}_2(AQ) = \operatorname{cond}_2(A)$;

(b) Se $B = Q^TAQ$, então $\operatorname{cond}_2(B) = \operatorname{cond}_2(A)$.

29. Seja A uma matriz invertível simétrica $n \times n$ com autovalores $\lambda_1, \ldots, \lambda_n$. Mostre que

$$\operatorname{cond}_2(A) = \frac{\max_{1\leq i\leq n} |\lambda_i|}{\min_{1\leq i\leq n} |\lambda_i|}$$

5 TRANSFORMAÇÕES ORTOGONAIS

Transformações ortogonais estão entre as ferramentas mais importantes da álgebra linear numérica. Os tipos de transformações ortogonais que estudaremos nesta seção são fáceis de trabalhar e não necessitam de muito espaço de armazenamento. O mais importante é que processos envolvendo transformações ortogonais são inerentemente estáveis. Por exemplo, seja $\mathbf{x} \in R^n$ e seja $\mathbf{x}' = \mathbf{x} + \mathbf{e}$ uma aproximação de $\mathbf{x}$: se Q é uma matriz ortogonal, então

$$Q\mathbf{x}' = Q\mathbf{x} + Q\mathbf{e}$$

O erro em $Q\mathbf{x}'$ é $Q\mathbf{e}$. Em relação à norma 2, o vetor $Q\mathbf{e}$ tem o mesmo tamanho que $\mathbf{e}$.

$$\|Q\mathbf{e}\|_2 = \|\mathbf{e}\|_2$$

Analogamente, se $A' = A + E$, então

$$QA' = QA + QE$$

e

$$\|QE\|_2 = \|E\|_2$$

Quando aplicamos uma transformação ortogonal a um vetor ou uma matriz, o erro não cresce em relação à norma 2.

TRANSFORMAÇÕES ORTOGONAIS ELEMENTARES

Uma *matriz ortogonal elementar* é uma matriz da forma

$$Q = I - 2\mathbf{u}\mathbf{u}^T$$

onde $\mathbf{u} \in R^n$ e $\|\mathbf{u}\|_2 = 1$. Para ver que Q é ortogonal, note que

$$Q^T = (I - 2\mathbf{u}\mathbf{u}^T)^T = I - 2\mathbf{u}\mathbf{u}^T = Q$$

e

$$\begin{aligned} Q^TQ = Q^2 &= (I - 2\mathbf{u}\mathbf{u}^T)(I - 2\mathbf{u}\mathbf{u}^T) \\ &= I - 4\mathbf{u}\mathbf{u}^T + 4\mathbf{u}(\mathbf{u}^T\mathbf{u})\mathbf{u}^T \\ &= I \end{aligned}$$

Logo, se Q é uma matriz ortogonal elementar,

$$Q^T = Q^{-1} = Q$$

A matriz $Q = I - 2\mathbf{u}\mathbf{u}^T$ fica completamente determinada pelo vetor $\mathbf{u}$. Em vez de guardar os n^2 elementos de Q, basta armazenar o vetor $\mathbf{u}$. Para calcular $Q\mathbf{x}$, note que

$$\begin{aligned} Q\mathbf{x} &= (I - 2\mathbf{u}\mathbf{u}^T)\mathbf{x} \\ &= \mathbf{x} - 2\alpha\mathbf{u} \qquad \text{onde} \qquad \alpha = \mathbf{u}^T\mathbf{x} \end{aligned}$$

O produto QA é calculado da seguinte maneira:

$$QA = (Q\mathbf{a}_1, Q\mathbf{a}_2, \ldots, Q\mathbf{a}_n)$$

onde

$$Q\mathbf{a}_i = \mathbf{a}_i - 2\alpha_i\mathbf{u} \qquad \alpha_i = \mathbf{u}^T\mathbf{a}_i$$

Transformações elementares ortogonais podem ser usadas para se obter uma fatoração QR de A, o que, por sua vez, pode ser usado para resolver um sistema linear $A\mathbf{x} = \mathbf{b}$. Como no método de Gauss, as matrizes elementares são escolhidas de modo a se anular elementos na matriz de coeficientes. Para ver como fazer isso, vamos considerar o problema de encontrar um vetor unitário $\mathbf{u}$ tal que

$$(I - 2\mathbf{u}\mathbf{u}^T)\mathbf{x} = (\alpha, 0, \ldots, 0)^T = \alpha\mathbf{e}_1$$

para um dado vetor $\mathbf{x} \in R^n$.

TRANSFORMAÇÕES DE HOUSEHOLDER

Seja $H = I - 2\mathbf{u}\mathbf{u}^T$. Se $H\mathbf{x} = \alpha\mathbf{e}_1$, então, como H é ortogonal, temos

$$|\alpha| = \|\alpha\mathbf{e}_1\|_2 = \|H\mathbf{x}\|_2 = \|\mathbf{x}\|_2$$

Escolhendo $\alpha = \|\mathbf{x}\|_2$ e $H\mathbf{x} = \alpha\mathbf{e}_1$, como H é sua própria inversa, temos que

(1) $$\mathbf{x} = H(\alpha\mathbf{e}_1) = \alpha(\mathbf{e}_1 - (2u_1)\mathbf{u})$$

Logo,

$$\begin{aligned} x_1 &= \alpha(1 - 2u_1^2) \\ x_2 &= -2\alpha u_1 u_2 \\ &\vdots \\ x_n &= -2\alpha u_1 u_n \end{aligned}$$

Resolvendo para os u_i, obtemos

$$\begin{aligned} u_1 &= \pm\left(\frac{\alpha - x_1}{2\alpha}\right)^{1/2} \\ u_i &= \frac{-x_i}{2\alpha u_1} \qquad \text{para} \quad i = 2, \ldots, n \end{aligned}$$

Definindo

$$u_1 = -\left(\frac{\alpha - x_1}{2\alpha}\right)^{1/2} \qquad \text{e} \qquad \beta = \alpha(\alpha - x_1)$$

temos

$$-2\alpha u_1 = [2\alpha(\alpha - x_1)]^{1/2} = (2\beta)^{1/2}$$

Segue que

$$\begin{aligned} \mathbf{u} &= \left(-\frac{1}{2\alpha u_1}\right)(-2\alpha u_1^2, x_2, \ldots, x_n)^T \\ &= \frac{1}{\sqrt{2\beta}}(x_1 - \alpha, x_2, \ldots, x_n)^T \end{aligned}$$

Definindo $\mathbf{v} = (x_1 - \alpha, x_2, \ldots, x_n)^T$, temos

$$\|\mathbf{v}\|_2^2 = (x_1 - \alpha)^2 + \sum_{i=2}^{n} x_i^2 = 2\alpha(\alpha - x_1)$$

e, portanto,

$$\|\mathbf{v}\|_2 = \sqrt{2\beta}$$

Logo,

$$\mathbf{u} = \frac{1}{\sqrt{2\beta}}\mathbf{v} = \frac{1}{\|\mathbf{v}\|_2}\mathbf{v}$$

Resumindo, dado um vetor $\mathbf{x} \in R^n$, se definirmos

$$\alpha = \|\mathbf{x}\|_2, \qquad \beta = \alpha(\alpha - x_1)$$
$$\mathbf{v} = (x_1 - \alpha, x_2, \ldots, x_n)^T$$
$$\mathbf{u} = \frac{1}{\|\mathbf{v}\|_2}\mathbf{v} = \frac{1}{\sqrt{2\beta}}\mathbf{v}$$

e

$$H = I - 2\mathbf{u}\mathbf{u}^T = I - \frac{1}{\beta}\mathbf{v}\mathbf{v}^T$$

então

$$H\mathbf{x} = \alpha\mathbf{e}_1$$

A matriz H formada dessa maneira é chamada de *matriz de Householder*. A matriz H é determinada pelo vetor $\mathbf{v}$ e pelo escalar β. Para qualquer vetor $\mathbf{y} \in R^n$,

$$H\mathbf{y} = \left(I - \frac{1}{\beta}\mathbf{v}\mathbf{v}^T\right)\mathbf{y} = \mathbf{y} - \left(\frac{1}{\beta}\mathbf{v}^T\mathbf{y}\right)\mathbf{v}$$

Em vez de armazenar os n^2 elementos de H, precisamos guardar apenas $\mathbf{v}$ e β.

EXEMPLO 1. Dado o vetor $\mathbf{x} = (4, 4, 2)^T$, encontre uma matriz de Householder que anula os dois últimos elementos de $\mathbf{x}$.

SOLUÇÃO. Defina

$$\alpha = \|\mathbf{x}\| = 6$$
$$\beta = \alpha(\alpha - x_1) = 12$$
$$\mathbf{v} = (x_1 - \alpha, x_2, x_3)^T = (-2, 4, 2)^T$$

A matriz de Householder é dada por

$$H = I - \frac{1}{12}\mathbf{v}\mathbf{v}^T$$
$$= \frac{1}{3}\begin{pmatrix} 2 & 2 & 1 \\ 2 & -1 & -2 \\ 1 & -2 & 2 \end{pmatrix}$$

Deixamos a cargo do leitor a verificação de que

$$H\mathbf{x} = 6\mathbf{e}_1$$

Suponha, agora, que queremos anular apenas as últimas $n - k$ componentes de um vetor $\mathbf{x} = (x_1, \ldots, x_k, x_{k+1}, \ldots, x_n)^T$. Para isso, definimos $\mathbf{x}^{(1)} = (x_1, \ldots, x_{k-1})^T$ e $\mathbf{x}^{(2)} = (x_k, x_{k+1}, \ldots, x_n)^T$. Vamos denotar por $I^{(1)}$ e $I^{(2)}$ as matrizes identidade $(k - 1) \times (k - 1)$ e $(n - k + 1) \times (n - k + 1)$, respectivamente. Pelos métodos que acabamos de descrever, podemos construir uma matriz de Householder $H_k^{(2)} = I^{(2)} - (1/\beta_k)\mathbf{v}_k\mathbf{v}_k^T$ tal que

$$H_k^{(2)}\mathbf{x}^{(2)} = \|\mathbf{x}^{(2)}\|_2\mathbf{e}_1$$

Seja

$$H_k = \begin{pmatrix} I^{(1)} & O \\ O & H_k^{(2)} \end{pmatrix}$$

Temos que

$$\begin{aligned}
H_k\mathbf{x} &= \begin{pmatrix} I^{(1)} & O \\ O & H_k^{(2)} \end{pmatrix}\begin{pmatrix} \mathbf{x}^{(1)} \\ \mathbf{x}^{(2)} \end{pmatrix} \\
&= \begin{pmatrix} I^{(1)}\mathbf{x}^{(1)} \\ H_k^{(2)}\mathbf{x}^{(2)} \end{pmatrix} \\
&= \left(x_1, \dots, x_{k-1}, \left(\sum_{i=k}^{n} x_i^2 \right)^{1/2}, 0, \dots, 0 \right)^T
\end{aligned}$$

Observações

1. A matriz de Householder H_k definida anteriormente é uma matriz ortogonal elementar. Se definirmos

$$\mathbf{v} = \begin{pmatrix} \mathbf{0} \\ \mathbf{v}_k \end{pmatrix} \quad \text{e} \quad \mathbf{u} = (1/\|\mathbf{v}\|)\mathbf{v}$$

   então

$$H_k = I - \frac{1}{\beta_k}\mathbf{v}\mathbf{v}^T = I - 2\mathbf{u}\mathbf{u}^T$$

2. H_k age como a identidade nas $k - 1$ primeiras coordenadas de qualquer vetor $\mathbf{y} \in R^n$. Se $\mathbf{y} = (y_1, \dots, y_{k-1}, y_k, \dots, x_n)^T$, $\mathbf{y}^{(1)} = (y_1, \dots, y_{k-1})^T$ e $\mathbf{y}^{(2)} = (y_k, \dots, y_n)^T$, temos

$$H_k\mathbf{y} = \begin{pmatrix} I^{(1)} & O \\ O & H_k^{(2)} \end{pmatrix}\begin{pmatrix} \mathbf{y}^{(1)} \\ \mathbf{y}^{(2)} \end{pmatrix} = \begin{pmatrix} \mathbf{y}^{(1)} \\ H_k^{(2)}\mathbf{y}^{(2)} \end{pmatrix}$$

   Em particular, se $\mathbf{y}^{(2)} = \mathbf{0}$, então $H_k\mathbf{y} = \mathbf{y}$.
3. Em geral, não é necessário armazenar toda a matriz H_k. Basta guardar os $n - k + 1$ vetores $\mathbf{v}_k$ e os escalares β_k.

EXEMPLO 2. Encontre uma matriz de Householder que anula as duas últimas coordenadas de $\mathbf{y} = (3, 4, 4, 2)^T$ ao mesmo tempo que mantém a primeira sem modificações.

SOLUÇÃO. A matriz de Householder vai mudar apenas as três últimas coordenadas de $\mathbf{y}$. Essas coordenadas correspondem ao vetor $\mathbf{x} = (4, 4, 2)^T$ em R^3. Mas esse é o vetor cujas duas últimas coordenadas foram anuladas no Exemplo 1. A matriz de Householder 3×3 do Exemplo 1 pode ser usada para se formar uma matriz 4×4,

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ 0 & \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ 0 & \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{pmatrix}$$

que vai ter o efeito desejado em $\mathbf{y}$. Deixamos a cargo do leitor verificar que $H\mathbf{y} = (3, 6, 0, 0)^T$. □

Agora estamos prontos para aplicar as transformações de Householder na resolução de sistemas lineares. Se A é uma matriz invertível $n \times n$, podemos usar transformações de Householder para colocar A em forma triangular. Para começar, podemos encontrar uma transformação de Householder $H_1 = I - (1/\beta_1)\mathbf{v}_1\mathbf{v}_1^T$ que, ao ser aplicada à primeira coluna de A, fornece um múltiplo de $\mathbf{e}_1$. Logo, H_1A vai ser da forma

$$\begin{pmatrix} \times & \times & \cdots & \times \\ 0 & \times & \cdots & \times \\ 0 & \times & \cdots & \times \\ \vdots & & & \\ 0 & \times & \cdots & \times \end{pmatrix}$$

Podemos depois encontrar uma transformação de Householder H_2 que vai anular os $n-2$ últimos elementos da segunda coluna de H_1A, deixando fixo o primeiro elemento. Pela observação (2), H_2 não tem nenhum efeito sobre a primeira coluna de H_1A.

$$H_2H_1A = \begin{pmatrix} \times & \times & \times & \cdots & \times \\ 0 & \times & \times & \cdots & \times \\ 0 & 0 & \times & \cdots & \times \\ \vdots & & & & \\ 0 & 0 & \times & \cdots & \times \end{pmatrix}$$

Podemos continuar a aplicar transformações de Householder desse modo até acabar com uma matriz triangular superior, que denotaremos por R. Logo,

$$H_{n-1}\cdots H_2H_1A = R$$

Segue que

$$\begin{aligned} A &= H_1^{-1}H_2^{-1}\cdots H_{n-1}^{-1}R \\ &= H_1H_2\cdots H_{n-1}R \end{aligned}$$

Seja $Q = H_1H_2\ldots H_{n-1}$. A matriz Q é ortogonal e A pode ser fatorada em um produto de uma matriz ortogonal por uma matriz triangular superior:

$$A = QR$$

Uma vez fatorada a matriz A em um produto QR, o sistema $A\mathbf{x} = \mathbf{b}$ pode ser facilmente resolvido.

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} \\ QR\mathbf{x} &= \mathbf{b} \\ R\mathbf{x} &= Q^T\mathbf{b} \end{aligned}$$

(2) $$R\mathbf{x} = H_{n-1}\cdots H_2H_1\mathbf{b}$$

Uma vez calculado $H_{n-1}\ldots H_2H_1\mathbf{b}$, o sistema (2) pode ser resolvido por substituição de baixo para cima.

Armazenamento. O vetor $\mathbf{v}_k$ pode ser guardado na k-ésima coluna de A. Como $\mathbf{v}_k$ tem $n-k+1$ coordenadas diferentes de zero e existem apenas $n-k$ elementos nulos na k-ésima coluna da matriz reduzida, é preciso guardar r_{kk} em outro lugar. Os elementos diagonais de R podem ser armazenados em um vetor de dimensão n ou como uma linha adicional de A. Os de β_k também podem ser armazenados como uma linha adicional de A.

$$\begin{pmatrix} v_{11} & r_{12} & r_{13} & r_{14} \\ v_{12} & v_{22} & r_{23} & r_{24} \\ v_{13} & v_{23} & v_{33} & r_{34} \\ v_{14} & v_{24} & v_{34} & 0 \\ r_{11} & r_{22} & r_{33} & r_{44} \\ \beta_1 & \beta_2 & \beta_3 & 0 \end{pmatrix}$$

Contagem das Operações. Ao se resolver um sistema $n \times n$ usando transformações de Householder, a maior parte do trabalho é feita para se colocar A em forma triangular. O número de operações necessárias é, aproximadamente, $\frac{2}{3}n^3$ multiplicações, $\frac{2}{3}n^3$ somas e $n-1$ raízes quadradas.

ROTAÇÕES E REFLEXÕES

Muitas vezes será desejável ter uma transformação que anula uma única coordenada em um vetor. Nesse caso, é conveniente usar uma rotação ou uma reflexão. Vamos considerar primeiro o caso bidimensional.

Sejam

$$R = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \quad \text{e} \quad G = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{pmatrix}$$

e seja

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} r\cos\alpha \\ r\operatorname{sen}\alpha \end{pmatrix}$$

um vetor em R^2.

$$R\mathbf{x} = \begin{pmatrix} r\cos(\theta+\alpha) \\ r\operatorname{sen}(\theta+\alpha) \end{pmatrix} \quad \text{e} \quad G\mathbf{x} = \begin{pmatrix} r\cos(\theta-\alpha) \\ r\operatorname{sen}(\theta-\alpha) \end{pmatrix}$$

R representa uma rotação no plano* de um ângulo θ. A matriz G reflete $\mathbf{x}$ em torno da reta $x_2 = [\operatorname{tg}(\theta/2)]x_1$ (ver Fig. 7.5.1). Fazendo $\cos\theta = x_1/r$ e $\operatorname{sen}\theta = -x_2/r$, obtemos

$$R\mathbf{x} = \begin{pmatrix} x_1\cos\theta - x_2\operatorname{sen}\theta \\ x_1\operatorname{sen}\theta + x_2\cos\theta \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix}$$

Fazendo $\cos\theta = x_1/r$ e $\operatorname{sen}\theta = x_2/r$, obtemos

$$G\mathbf{x} = \begin{pmatrix} x_1\cos\theta + x_2\operatorname{sen}\theta \\ x_1\operatorname{sen}\theta - x_2\cos\theta \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix}$$

Tanto G quanto R são ortogonais. A matriz G também é simétrica. Na verdade, G é uma matriz ortogonal elementar: se $\mathbf{u} = (\operatorname{sen}\theta/2, \cos\theta/2)^T$, então $G = I - 2\mathbf{u}\mathbf{u}^T$.

EXEMPLO 3. Seja $\mathbf{x} = (-3, 4)^T$. Para encontrar uma matriz de rotação R que anule a segunda coordenada de $\mathbf{x}$, calculamos

$$r = \sqrt{(-3)^2 + 4^2} = 5$$
$$\cos\theta = \frac{x_1}{r} = -\frac{3}{5}$$
$$\operatorname{sen}\theta = -\frac{x_2}{r} = -\frac{4}{5}$$

e

$$R = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} -\frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & -\frac{3}{5} \end{pmatrix}$$

*Em torno da origem. (*N. T.*)

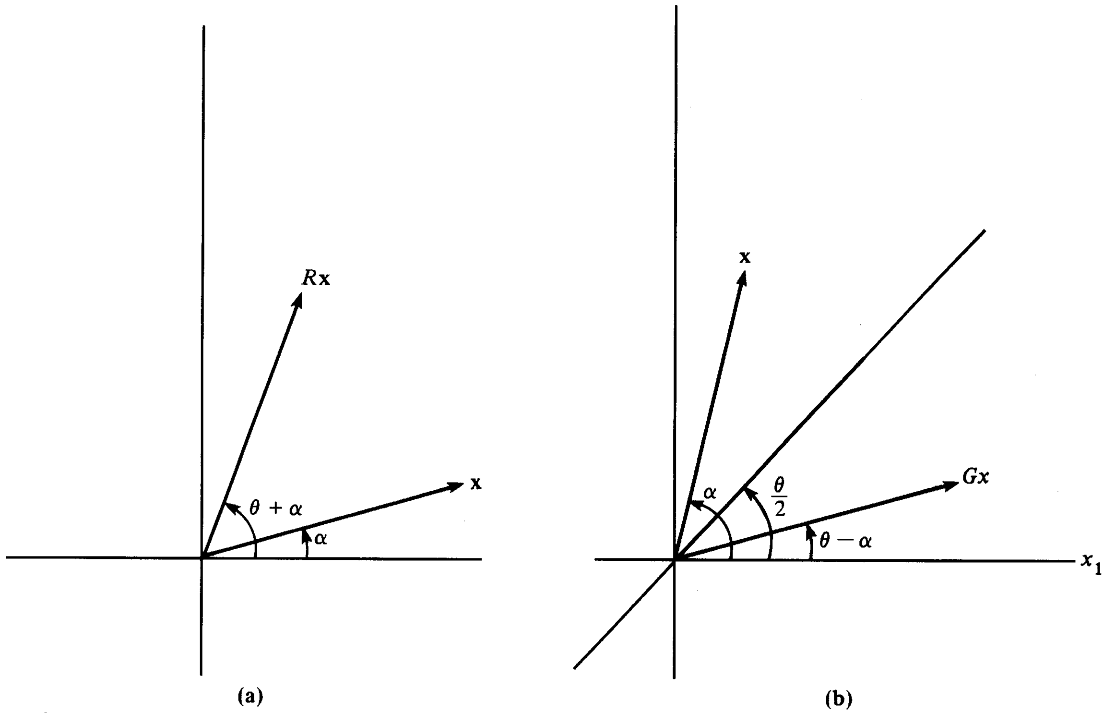


FIG. 7.5.1

O leitor pode verificar que $R\mathbf{x} = 5\mathbf{e}_1$.

Para encontrar uma matriz de reflexão G que anula a segunda coordenada de $\mathbf{x}$, calcule r e $\cos\theta$ da mesma maneira que para a matriz de rotação, mas faça

$$\operatorname{sen}\theta = \frac{x_2}{r} = \frac{4}{5}$$

e

$$G = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{pmatrix} = \begin{pmatrix} -\frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}$$

O leitor pode verificar que $G\mathbf{x} = 5\mathbf{e}_1$. □

Vamos considerar agora o caso de dimensão n. Sejam R e G matrizes $n \times n$ com

$$r_{ii} = r_{jj} = \cos\theta \qquad g_{ii} = \cos\theta,\ g_{jj} = -\cos\theta$$

$$r_{ji} = \operatorname{sen}\theta,\ r_{ij} = -\operatorname{sen}\theta \qquad g_{ji} = g_{ij} = \operatorname{sen}\theta$$

e $r_{st} = g_{st} = \delta_{st}$ para todos os outros elementos de R e de G. Então R e G coincidem com a matriz identidade, exceto pelos elementos nas posições (i,i), (i,j), (j,j) e (j,i). Sejam $c = \cos\theta$ e $s = \operatorname{sen}\theta$. Se $\mathbf{x} \in R^n$, então

$$R\mathbf{x} = (x_1, \ldots, x_{i-1}, x_i c - x_j s, x_{i+1}, \ldots, x_{j-1}, x_i s + x_j c, x_{j+1}, \ldots, x_n)^T$$

e

$$G\mathbf{x} = (x_1, \ldots, x_{i-1}, x_i c + x_j s, x_{i+1}, \ldots, x_{j-1}, x_i s - x_j c, x_{j+1}, \ldots, x_n)^T$$

As transformações R e G alteram apenas as i-ésima e j-ésima componentes de um vetor. Elas não afetam as outras coordenadas. Vamos nos referir a R como uma *rotação plana* e a G como uma *transformação de Givens* ou uma *reflexão de Givens*. Fazendo

$$c = \frac{x_i}{r} \quad \text{e} \quad s = -\frac{x_j}{r} \qquad \left(r = \sqrt{x_i^2 + x_j^2}\right)$$

a j-ésima componente de $R\mathbf{x}$ será 0. Fazendo

$$c = \frac{x_i}{r} \qquad \text{e} \qquad s = \frac{x_j}{r}$$

a j-ésima componente de $G\mathbf{x}$ será 0.

EXEMPLO 4. Seja $\mathbf{x} = (5, 8, 12)^T$. Encontre uma matriz de rotação R que anula a terceira coordenada de $\mathbf{x}$ mas deixa fixa a segunda.

SOLUÇÃO. Como R vai agir apenas em x_1 e x_3, faça

$$r = \sqrt{x_1^2 + x_3^2} = 13$$

$$c = \frac{x_1}{r} = \frac{5}{13}$$

$$s = -\frac{x_3}{r} = -\frac{12}{13}$$

e defina

$$R = \begin{pmatrix} c & 0 & -s \\ 0 & 1 & 0 \\ s & 0 & c \end{pmatrix} = \begin{pmatrix} \frac{5}{13} & 0 & \frac{12}{13} \\ 0 & 1 & 0 \\ -\frac{12}{13} & 0 & \frac{5}{13} \end{pmatrix}$$

O leitor pode verificar que $R\mathbf{x} = (13, 8, 0)^T$. □

Dada uma matriz invertível A $n \times n$, podemos usar rotações planas ou transformações de Givens para obter uma fatoração QR de A. Seja G_{21} a transformação de Givens que age na primeira e segunda coordenadas e que, ao ser aplicada a A, anula o elemento na posição (2, 1). Podemos aplicar outra transformação de Givens, G_{31}, a $G_{21}A$ de modo a obter um zero na posição (3, 1). Esse processo pode ser continuado até os últimos $n - 1$ elementos na primeira coluna ficarem iguais a zero.

$$G_{n1} \cdots G_{31} G_{21} A = \begin{pmatrix} \times & \times & \cdots & \times \\ 0 & \times & \cdots & \times \\ 0 & \times & \cdots & \times \\ \vdots & & & \\ 0 & \times & \cdots & \times \end{pmatrix}$$

Na próxima etapa, as transformações de Givens $G_{32}, G_{42}, \ldots, G_{n2}$ são usadas para se anular os últimos $n - 2$ elementos da segunda coluna. O processo continua até todos os elementos abaixo da diagonal se tornarem nulos.

$$(G_{n,n-1}) \cdots (G_{n2} \cdots G_{32})(G_{n1} \cdots G_{21})A = R \qquad (R \text{ triangular superior})$$

Fazendo $Q^T = (G_{n,n-1}) \ldots (G_{n2} \ldots G_{32})(G_{n1} \ldots G_{21})$, temos $A = QR$ e o sistema $A\mathbf{x} = \mathbf{b}$ é equivalente ao sistema

$$R\mathbf{x} = Q^T\mathbf{b}$$

Esse sistema pode ser resolvido por substituição de baixo para cima.

Contagem das Operações. A fatoração QR de A usando transformações de Givens ou rotações planas necessita de aproximadamente $\frac{4}{3}n^3$ multiplicações, $\frac{2}{3}n^3$ somas e $\frac{1}{2}n^2$ raízes quadradas.

EXERCÍCIOS

1. Para cada um dos vetores $\mathbf{x}$ a seguir, encontre uma matriz de rotação R tal que $R\mathbf{x} = \|\mathbf{x}\|_2\mathbf{e}_1$.

(a) $\mathbf{x} = (1, 1)^T$ (b) $\mathbf{x} = (\sqrt{3}, -1)^T$ (c) $\mathbf{x} = (-4, 3)^T$

2. Dado $\mathbf{x} \in R^3$, defina

$$r_{ij} = \left(x_i^2 + x_j^2\right)^{1/2} \qquad i, j = 1, 2, 3$$

Em cada um dos itens a seguir, encontre uma transformação de Givens G_{ij} tal que as coordenadas i e j de $G_{ij}\mathbf{x}$ são r_{ij} e 0, respectivamente.

(a) $\mathbf{x} = (3, 1, 4)^T$, $i = 1, j = 3$
(b) $\mathbf{x} = (1, -1, 2)^T$, $i = 1, j = 2$
(c) $\mathbf{x} = (4, 1, \sqrt{3})^T$, $i = 2, j = 3$
(d) $\mathbf{x} = (4, 1, \sqrt{3})^T$, $i = 3, j = 2$

3. Para cada um dos vetores $\mathbf{x}$ dados a seguir, encontre uma transformação de Householder tal que $H\mathbf{x} = \alpha\mathbf{e}_1$, onde $\alpha = \|\mathbf{x}\|_2$.

(a) $\mathbf{x} = (8, -1, -4)^T$ (b) $\mathbf{x} = (6, 2, 3)^T$ (c) $\mathbf{x} = (7, 4, -4)^T$

4. Para cada um dos vetores a seguir, encontre uma transformação de Householder que anula suas duas últimas coordenadas.

(a) $\mathbf{x} = (5, 8, 4, 1)^T$ (b) $\mathbf{x} = (4, -3, -2, -1, 2)^T$

5. Seja

$$A = \begin{pmatrix} 3 & 3 & -2 \\ 1 & 1 & 1 \\ 1 & -5 & 1 \\ 5 & -1 & 2 \end{pmatrix}$$

(a) Determine o escalar β e o vetor $\mathbf{v}$ para a matriz de Householder $H = I - (1/\beta)\mathbf{v}\mathbf{v}^T$ que anula os três últimos elementos de $\mathbf{a}_1$.
(b) Sem formar explicitamente a matriz H, calcule o produto HA.

6. Sejam

$$A = \begin{pmatrix} 1 & 2 & -4 \\ 2 & 6 & 7 \\ -2 & 1 & 8 \end{pmatrix} \quad \text{e} \quad \mathbf{b} = \begin{pmatrix} 9 \\ 9 \\ -3 \end{pmatrix}$$

(a) Use transformações de Householder para transformar A em uma matriz triangular superior R. Transforme também o vetor $\mathbf{b}$, isto é, calcule $\mathbf{b}^{(1)} = H_2H_1\mathbf{b}$.
(b) Resolva $R\mathbf{x} = \mathbf{b}^{(1)}$ para $\mathbf{x}$ e verifique sua resposta calculando o resíduo $\mathbf{b} - A\mathbf{x}$.

7. Para cada um dos sistemas a seguir, use reflexões de Givens para transformá-lo em um sistema triangular superior e depois resolva esse sistema triangular.

(a) $\begin{aligned} 3x_1 + 8x_2 &= 5 \\ 4x_1 - x_2 &= -5 \end{aligned}$ (b) $\begin{aligned} x_1 + 4x_2 &= 5 \\ x_1 + 2x_2 &= 1 \end{aligned}$

(c) $$\begin{aligned} 4x_1 - 4x_2 + x_3 &= 2 \\ x_2 + 3x_3 &= 2 \end{aligned}$$

8. Suponha que você quer anular a última coordenada de um vetor $\mathbf{x}$ e deixar as $n - 2$ primeiras inalteradas. Quantas operações são necessárias se isso for feito através de uma transformação de Givens G? E de uma transformação de Householder H? Se A é uma matriz $n \times n$, quantas operações são necessárias para calcular GA e HA?

9. Seja $H_k = I - 2\mathbf{u}\mathbf{u}^T$ uma transformação de Householder com

$$\mathbf{u} = (0, \ldots, 0, u_k, u_{k+1}, \ldots, u_n)^T$$

Sejam $\mathbf{b} \in R^n$ e A uma matriz $n \times n$. Quantas somas e multiplicações são necessárias para se calcular (a) $H_k\mathbf{b}$; (b) H_kA?

10. Seja $Q^T = G_{n-k} \ldots G_2G_1$, onde cada G_i é uma transformação de Givens. Sejam $\mathbf{b} \in R^n$ e A uma matriz $n \times n$. Quantas somas e multiplicações são necessárias para se calcular (a) $Q^T\mathbf{b}$; (b) Q^TA?

11. Sejam R_1 e R_2 duas matrizes de rotação 2×2 e sejam G_1 e G_2 duas transformações 2×2 de Givens. Indique o tipo de transformação em cada um dos itens a seguir.

(a) R_1R_2 (b) G_1G_2 (c) R_1G_1 (d) G_1R_1

12. Sejam $\mathbf{x}$ e $\mathbf{y}$ vetores distintos em R^n com $\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2$. Defina

$$\mathbf{u} = \frac{1}{\|\mathbf{x} - \mathbf{y}\|_2}(\mathbf{x} - \mathbf{y}) \qquad \text{e} \qquad Q = I - 2\mathbf{u}\mathbf{u}^T$$

Mostre que:

(a) $\|\mathbf{x} - \mathbf{y}\|_2^2 = 2(\mathbf{x} - \mathbf{y})^T\mathbf{x}$ (b) $Q\mathbf{x} = \mathbf{y}$

13. Seja $\mathbf{u}$ um vetor unitário em C^n e seja

$$U = I - 2\mathbf{u}\mathbf{u}^T$$

(a) Mostre que $\mathbf{u}$ é um autovetor de U. Qual o autovalor associado?
(b) Seja $\mathbf{z}$ um vetor não-nulo em C^n que é ortogonal a $\mathbf{u}$. Mostre que $\mathbf{z}$ também é um autovetor de U. Qual o autovalor associado?

14. Seja $A = Q_1R_1 = Q_2R_2$, onde Q_1 e Q_2 são ortogonais e R_1 e R_2 são triangulares superiores e invertíveis.
(a) Mostre que $Q_1^TQ_2$ é diagonal.
(b) Qual a relação entre R_1 e R_2? Explique.

6 A DECOMPOSIÇÃO EM VALORES SINGULARES

Em muitas aplicações há necessidade de determinar o posto de uma matriz ou de determinar se a matriz tem ou não posto máximo. Teoricamente, poderíamos usar o método de Gauss para colocar a matriz em forma escada e depois contar o número de linhas não-nulas. No entanto, essa abordagem não é prática ao se trabalhar com aritmética de precisão finita. Se A não tem posto máximo e U é sua forma escada calculada, então, devido a erros de aproximação no processo, dificilmente U vai ter o número correto de linhas nulas. Na prática, a matriz de coeficientes envolve, geralmente, algum erro. Isso pode dever-se a erros nos dados ou ao sistema numérico finito. Portanto, em geral é mais prático perguntar se A está "próxima" de uma matriz que não tem posto máximo. No entanto, pode acontecer que A esteja próxima de uma matriz que não tem posto máximo, mas não sua forma escada calculada U.

Vamos supor, em toda esta seção, que A é uma matriz $m \times n$ com $m \geq n$. (Essa hipótese é formulada apenas por conveniência; todos os resultados continuam válidos no caso $m < n$.) Vamos apresentar um método para determinar quão próxima A está de uma matriz de posto menor. O método envolve a fatoração de A em um produto $U\Sigma V^T$, onde U é uma matriz ortogonal $m \times m$, V é uma matriz ortogonal $n \times$

n e Σ é uma matriz com todos os elementos fora da diagonal iguais a 0 e com os elementos diagonais satisfazendo

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$$

$$\Sigma = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \end{pmatrix}$$

Os σ_i determinados por essa fatoração são únicos e são chamados *valores singulares* de A. A fatoração $U\Sigma V^T$ é chamada de *decomposição em valores singulares*. Vamos mostrar que o posto de A é igual ao número de valores singulares não-nulos e que os módulos dos valores singulares não-nulos medem quão perto está a matriz A de uma matriz de posto menor.

Vamos começar mostrando que uma tal decomposição sempre é possível.

Teorema 7.6.1. *Se A é uma matriz $m \times n$, então A tem uma decomposição em valores singulares.*

Demonstração. A^TA é uma matriz simétrica, logo seus autovalores são todos reais e ela tem uma matriz diagonalizadora ortogonal V. Além disso, nenhum de seus autovalores pode ser negativo. Para ver isso, seja λ um autovalor de A^TA e seja $\mathbf{x}$ um autovetor associado. Então,

$$\|A\mathbf{x}\|^2 = \mathbf{x}^T A^T A\mathbf{x} = \lambda \mathbf{x}^T\mathbf{x} = \lambda\|\mathbf{x}\|^2$$

Logo,

$$\lambda = \frac{\|A\mathbf{x}\|^2}{\|\mathbf{x}\|^2} \geq 0$$

Vamos supor que as colunas de V estão ordenadas de modo que os autovalores correspondentes satisfazem

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$$

Os valores singulares de A são dados por

$$\sigma_j = \sqrt{\lambda_j} \qquad j = 1, \ldots, n$$

Vamos denotar por r o posto de A. A matriz A^TA também tem posto r. Como A^TA é simétrica, seu posto é igual ao número de autovalores não-nulos. Logo,

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0 \qquad \text{e} \qquad \lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_n = 0$$

A mesma relação é válida para os valores singulares:

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 \qquad \text{e} \qquad \sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$$

Vamos definir

$$V_1 = (\mathbf{v}_1, \ldots, \mathbf{v}_r), \qquad V_2 = (\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n)$$

e

$$\Sigma_1 = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{pmatrix}$$

Então, Σ_1 é uma matriz diagonal $r \times r$ cujos elementos diagonais são os valores singulares não-nulos σ_1, ..., σ_r. A matriz $m \times n$ Σ é dada por

$$\Sigma = \begin{pmatrix} \Sigma_1 & O \\ O & O \end{pmatrix}$$

Os vetores colunas de V_2 são os autovetores de A^TA associados a $\lambda = 0$. Logo,

$$A^TA\mathbf{v}_j = \mathbf{0} \qquad j = r+1, \dots, n$$

e, portanto, as colunas de V_2 formam uma base ortonormal para $N(A^TA) = N(A)$. Então,

$$AV_2 = O$$

e, como V é uma matriz ortogonal, segue que

$$I = VV^T = V_1V_1^T + V_2V_2^T$$

(1) $$A = AI = AV_1V_1^T + AV_2V_2^T = AV_1V_1^T$$

Até agora, mostramos como construir as matrizes V e Σ da decomposição singular. Para completar a demonstração, precisamos mostrar como construir uma matriz ortogonal U $m \times m$ tal que

$$A = U\Sigma V^T$$

ou, equivalentemente,

(2) $$AV = U\Sigma$$

Comparando as r primeiras colunas de cada lado de (2), vemos que

$$A\mathbf{v}_j = \sigma_j\mathbf{u}_j \qquad j = 1, \dots, r$$

Logo, se definirmos,

(3) $$\mathbf{u}_j = \frac{1}{\sigma_j}A\mathbf{v}_j \qquad j = 1, \dots, r$$

e

$$U_1 = (\mathbf{u}_1, \dots, \mathbf{u}_r)$$

segue que

(4) $$AV_1 = U_1\Sigma_1$$

As colunas de U formam um conjunto ortonormal, já que

$$\begin{aligned} \mathbf{u}_i^T\mathbf{u}_j &= \left(\frac{1}{\sigma_i}\mathbf{v}_i^TA^T\right)\left(\frac{1}{\sigma_j}A\mathbf{v}_j\right) \qquad 1 \le i \le r, \quad 1 \le j \le r \\ &= \frac{1}{\sigma_i\sigma_j}\mathbf{v}_i^T\left(A^TA\mathbf{v}_j\right) \\ &= \frac{\sigma_j}{\sigma_i}\mathbf{v}_i^T\mathbf{v}_j \\ &= \delta_{ij} \end{aligned}$$

Segue de (3) que cada $\mathbf{u}_j$, $1 \le j \le r$, está no espaço coluna de A. A dimensão do espaço coluna é r, de modo que $\mathbf{u}_1, \dots, \mathbf{u}_r$ forma uma base ortonormal para $R(A)$. O espaço vetorial $I(A)^\perp = N(A^T)$ tem dimensão $m - r$. Seja $\{\mathbf{u}_{r+1}, \mathbf{u}_{r+2}, \dots, \mathbf{u}_m\}$ uma base ortonormal para $N(A^T)$ e defina

$$U_2 = (\mathbf{u}_{r+1}, \mathbf{u}_{r+2}, \ldots, \mathbf{u}_m)$$
$$U = \begin{pmatrix} U_1 & U_2 \end{pmatrix}$$

Pelo Teorema 5.2.2, $\mathbf{u}_1, \ldots, \mathbf{u}_m$ formam uma base ortonormal para R^m. Logo, U é uma matriz ortogonal. Precisamos, ainda, mostrar que $U\Sigma V^T$ é, de fato, igual a A. Isso segue de (4) e (1), pois

$$\begin{aligned} U\Sigma V^T &= \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} \Sigma_1 & O \\ O & O \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix} \\ &= U_1 \Sigma_1 V_1^T \\ &= A V_1 V_1^T \\ &= A \end{aligned}$$

□

Observações. Seja A uma matriz $m \times n$ com uma decomposição em valores singulares $U\Sigma V^T$.

1. Os valores singulares $\sigma_1, \ldots, \sigma_n$ de A são únicos; no entanto, as matrizes U e V não são únicas.
2. Como V diagonaliza A^TA, os $\mathbf{v}_j$ são autovetores de A^TA.
3. Como $AA^T = U\Sigma\Sigma^TU^T$, temos que U diagonaliza AA^T e os $\mathbf{u}_j$ são autovetores de AA^T.
4. Comparando as j-ésimas colunas de cada lado da equação

$$AV = U\Sigma$$

obtemos

$$A\mathbf{v}_j = \sigma_j \mathbf{u}_j \qquad j = 1, \ldots, n$$

Analogamente,

$$A^TU = V\Sigma^T$$

e, portanto,

$$\begin{aligned} A^T\mathbf{u}_j &= \sigma_j \mathbf{v}_j \qquad \text{para} \quad j = 1, \ldots, n \\ A^T\mathbf{u}_j &= \mathbf{0} \qquad \text{para} \quad j = n+1, \ldots, m \end{aligned}$$

Os $\mathbf{v}_j$ são chamados de *valores singulares à direita* de A e os u_j são os *valores singulares à esquerda* de A.

5. Se A tem posto r, então
   (i) $\mathbf{v}_1, \ldots, \mathbf{v}_r$ formam uma base ortonormal para $I(A^T)$.
   (ii) $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$ formam uma base ortonormal para $N(A)$.
   (iii) $\mathbf{u}_1, \ldots, \mathbf{u}_r$ formam uma base ortonormal para $I(A)$.
   (iv) $\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m$ formam uma base ortonormal para $N(A^T)$.
6. O posto da matriz A é igual ao número de valores singulares não-nulos (contados de acordo com a multiplicidade). O leitor deve tomar cuidado para não fazer uma afirmação semelhante sobre autovalores. A matriz

$$M = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

por exemplo, tem posto 3 apesar de todos os seus autovalores serem nulos.

EXEMPLO 1. Seja

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{pmatrix}$$

Calcule os valores singulares e a decomposição em valores singulares de A.

SOLUÇÃO. A matriz

$$A^TA = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$$

tem autovalores $\lambda_1 = 4$ e $\lambda_2 = 0$. Logo, os valores singulares de A são $\sigma_1 = \sqrt{4} = 2$ e $\sigma_2 = 0$. O autovalor λ_1 tem autovetores associados da forma $\alpha(1, 1)^T$ e λ_2 tem autovetores da forma $\beta(1, -1)^T$. Portanto, a matriz ortogonal

$$V = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

diagonaliza A^TA. Pela observação 4, temos que

$$\mathbf{u}_1 = \frac{1}{\sigma_1}A\mathbf{v}_1 = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$$

As colunas restantes de U têm que formar uma base ortonormal para $N(A^T)$. Podemos obter uma base $\{\mathbf{x}_2, \mathbf{x}_3\}$ para $N(A^T)$ da maneira usual.

$$\mathbf{x}_2 = (1, -1, 0)^T \quad \text{e} \quad \mathbf{x}_3 = (0, 0, 1)^T$$

Como esses vetores já são ortogonais, não é preciso usar o processo de Gram-Schmidt para se obter uma base ortonormal. Precisamos apenas definir

$$\mathbf{u}_2 = \frac{1}{\|\mathbf{x}_2\|}\mathbf{x}_2 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right)^T$$
$$\mathbf{u}_3 = \mathbf{x}_3 = (0, 0, 1)^T$$

Temos, então,

$$A = U\Sigma V^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$ □

Se A é uma matriz $m \times n$ de posto r e $0 < k < r$, podemos usar a decomposição em valores singulares para encontrar uma matriz em $R^{m\times n}$ de posto k o mais próximo possível de A em relação à norma de Frobenius. Seja $\mathcal{M}$ o conjunto de todas as matrizes $m \times n$ de posto menor ou igual a k. Pode-se mostrar que existe uma matriz X em $\mathcal{M}$ tal que

$$\|A - X\|_F = \min_{S\in\mathcal{M}} \|A - S\|_F \tag{5}$$

Não vamos provar isso, já que a demonstração está além do escopo deste livro. Supondo que é possível atingir o mínimo, vamos mostrar como obter uma tal matriz X da decomposição em valores singulares de A. O lema a seguir será útil.

Lema 7.6.2. *Se A é uma matriz $m \times n$ e Q é uma matriz $m \times m$ ortogonal, então*

$$\|QA\|_F = \|A\|_F$$

Demonstração

$$\begin{aligned}\|QA\|_F^2 &= \|(Q\mathbf{a}_1, Q\mathbf{a}_2, \ldots, Q\mathbf{a}_n)\|_F^2 \\ &= \sum_{i=1}^{n} \|Q\mathbf{a}_i\|_2^2 \\ &= \sum_{i=1}^{n} \|\mathbf{a}_i\|_2^2 \\ &= \|A\|_F^2\end{aligned}$$

□

Pelo lema, se A tem uma decomposição em valores singulares $U\Sigma V^T$, então

$$\|A\|_F = \|\Sigma V^T\|_F$$

Como

$$\|\Sigma V^T\|_F = \|(\Sigma V^T)^T\|_F = \|V\Sigma^T\|_F = \|\Sigma^T\|_F$$

segue que

$$\|A\|_F = \left(\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2\right)^{1/2}$$

Teorema 7.6.1. *Seja $A = U\Sigma V^T$ uma matriz $m \times n$ e denote por $\mathcal{M}$ o conjunto de todas as matrizes $m \times n$ de posto menor ou igual a k, onde $0 < k <$ posto(A). Se X é uma matriz em $\mathcal{M}$ satisfazendo (5), então*

$$\|A - X\|_F = \left(\sigma_{k+1}^2 + \sigma_{k+2}^2 + \cdots + \sigma_n^2\right)^{1/2}$$

Em particular, se $A' = U\Sigma' V^T$, onde

$$\Sigma' = \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & O \\ & & \sigma_k & \\ \hline & O & & O \end{array}\right) = \begin{pmatrix} \Sigma_k & O \\ O & O \end{pmatrix}$$

então

$$\|A - A'\|_F = \left(\sigma_{k+1}^2 + \cdots + \sigma_n^2\right)^{1/2} = \min_{S \in \mathcal{M}} \|A - S\|_F$$

Demonstração. Seja X uma matriz em $\mathcal{M}$ satisfazendo (5). Como $A' \in \mathcal{M}$, temos

(6) $$\|A - X\|_F \leq \|A - A'\|_F = \left(\sigma_{k+1}^2 + \cdots + \sigma_n^2\right)^{1/2}$$

Vamos mostrar que

$$\|A - X\|_F \geq \left(\sigma_{k+1}^2 + \cdots + \sigma_n^2\right)^{1/2}$$

e, portanto, que temos igualdade em (6). Seja $Q\Omega P^T$ a decomposição em valores singulares de X.

$$\Omega = \left(\begin{array}{cccc|c} \omega_1 & & & & \\ & \omega_2 & & & \\ & & \ddots & & O \\ & & & \omega_k & \\ \hline & & O & & O \end{array}\right) = \begin{pmatrix} \Omega_k & O \\ O & O \end{pmatrix}$$

Definindo $B = Q^TAP$, temos que $A = QBP^T$ e

$$\|A - X\|_F = \|Q(B - \Omega)P^T\|_F = \|B - \Omega\|_F$$

Vamos escrever a matriz B em blocos do mesmo modo que Ω.

$$B = \left(\begin{array}{c|c} \overbrace{B_{11}}^{k\times k} & \overbrace{B_{12}}^{k\times(n-k)} \\ \hline \underbrace{B_{21}}_{(m-k)\times k} & \underbrace{B_{22}}_{(m-k)\times(n-k)} \end{array} \right)$$

Segue que

$$\|A - X\|_F^2 = \|B_{11} - \Omega_k\|_F^2 + \|B_{12}\|_F^2 + \|B_{21}\|_F^2 + \|B_{22}\|_F^2$$

Vamos mostrar por absurdo que $B_{12} = O$. Caso contrário, defina

$$Y = Q\begin{pmatrix} B_{11} & B_{12} \\ O & O \end{pmatrix} P^T$$

A matriz Y pertence a $\mathcal{M}$ e

$$\|A - Y\|_F^2 = \|B_{21}\|_F^2 + \|B_{22}\|_F^2 < \|A - X\|_F^2$$

o que contradiz a definição de X. Logo, $B_{12} = O$. De maneira análoga, podemos mostrar que B_{21} tem que ser igual a O. Definindo

$$Z = Q\begin{pmatrix} B_{11} & O \\ O & O \end{pmatrix} P^T$$

então $Z \in \mathcal{M}$ e

$$\|A - Z\|_F^2 = \|B_{22}\|_F^2 \le \|B_{11} - \Omega_k\|_F^2 + \|B_{22}\|_F^2 = \|A - X\|_F^2$$

Pela definição de X, B_{11} tem que ser igual a Ω_k. Se B_{22} tem decomposição em valores singulares $U_1 \Lambda V_1^T$, então

$$\|A - X\|_F = \|B_{22}\|_F = \|\Lambda\|_F$$

Sejam

$$U_2 = \begin{pmatrix} I_k & O \\ O & U_1 \end{pmatrix} \quad \text{e} \quad V_2 = \begin{pmatrix} I_k & O \\ O & V_1 \end{pmatrix}$$

Temos

$$U_2^T Q^T A P V_2 = \begin{pmatrix} \Omega_k & O \\ O & \Lambda \end{pmatrix}$$

$$A = (QU_2)\begin{pmatrix} \Omega_k & O \\ O & \Lambda \end{pmatrix}(PV_2)^T$$

logo, os elementos diagonais de Λ são valores singulares de A. Então,

$$\|A - X\|_F = \|\Lambda\|_F \ge \left(\sigma_{k+1}^2 + \cdots + \sigma_n^2\right)^{1/2}$$

De (6), temos que

$$\|A - X\|_F = \left(\sigma_{k+1}^2 + \cdots + \sigma_n^2\right)^{1/2} = \|A - A'\|_F \qquad \square$$

Se $A = U\Sigma V^T$ e definimos $E_j = \mathbf{u}_j \mathbf{v}_j^T$ para $j = 1, \ldots, n$, então cada E_j tem posto 1 e

$$A = \sigma_1 E_1 + \sigma_2 E_2 + \cdots + \sigma_n E_n \tag{7}$$

Para provar (7), note que, como V é uma matriz ortogonal,

$$\begin{aligned} I &= VV^T \\ &= (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)\begin{pmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{pmatrix} \\ &= \mathbf{v}_1\mathbf{v}_1^T + \cdots + \mathbf{v}_n\mathbf{v}_n^T \end{aligned}$$

portanto,

$$\begin{aligned} A &= A\left(\mathbf{v}_1\mathbf{v}_1^T + \cdots + \mathbf{v}_n\mathbf{v}_n^T\right) \\ &= (A\mathbf{v}_1)\mathbf{v}_1^T + \cdots + (A\mathbf{v}_n)\mathbf{v}_n^T \\ &= \sigma_1\mathbf{u}_1\mathbf{v}_1^T + \cdots + \sigma_n\mathbf{u}_n\mathbf{v}_n^T \\ &= \sigma_1 E_1 + \cdots + \sigma_n E_n \end{aligned}$$

Se A tem posto n, então

$$\begin{aligned} A' &= U\begin{pmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_{n-1} & \\ & & & & 0 \\ & & & & \end{pmatrix}V^T \\ &= \sigma_1 E_1 + \cdots + \sigma_{n-1}E_{n-1} \end{aligned}$$

vai ser uma matriz de posto $n - 1$ o mais próximo possível de A em relação à norma de Frobenius. Analogamente, $A'' = \sigma_1 E_1 + \ldots + \sigma_{n-2}E_{n-2}$ vai ser a matriz mais próxima de posto $n - 2$, e assim por diante. Em particular, se A é uma matriz invertível $n \times n$, então A' é singular e $\| A - A' \|_F = \sigma_n$. Logo, σ_n pode ser considerado uma medida de quão próxima uma matriz está de ser singular.

O leitor deve tomar cuidado para não usar o valor de det(A) como uma medida de quão próxima A está de ser singular. Se, por exemplo, A é a matriz diagonal 100×100 cujos elementos diagonais são todos iguais a $\frac{1}{2}$, então $\det(A) = 2^{-100}$; no entanto, $\sigma_{100} = \frac{1}{2}$. Por outro lado, a matriz no exemplo a seguir está muito próxima de ser singular, apesar de seu determinante e todos os seus autovalores serem iguais a 1.

EXEMPLO 2. Seja A uma matriz $n \times n$ triangular superior cujos elementos diagonais são todos iguais a 1 e cujos elementos acima da diagonal são todos iguais a -1.

$$A = \begin{pmatrix} 1 & -1 & -1 & \cdots & -1 & -1 \\ 0 & 1 & -1 & \cdots & -1 & -1 \\ 0 & 0 & 1 & \cdots & - & -1 \\ \vdots & & & & & \\ 0 & 0 & 0 & \cdots & 1 & -1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

Note que $\det(A) = \det(A^{-1}) = 1$ e que todos os autovalores de A são iguais a 1. No entanto, se n for grande, A está perto de ser singular. Para ver isso, seja

$$B = \begin{pmatrix} 1 & -1 & -1 & \cdots & -1 & -1 \\ 0 & 1 & -1 & \cdots & -1 & -1 \\ 0 & 0 & 1 & \cdots & -1 & -1 \\ \vdots & & & & & \\ 0 & 0 & 0 & \cdots & 1 & -1 \\ \frac{-1}{2^{n-2}} & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

B é singular, pois o sistema $B\mathbf{x} = \mathbf{0}$ tem uma solução não-trivial $\mathbf{x} = (2^{n-2}, 2^{n-3}, \ldots, 2^0, 1)^T$. As matrizes A e B diferem apenas na posição $(n, 1)$.

$$\|A - B\|_F = \frac{1}{2^{n-2}}$$

Pelo Teorema 7.6.3, temos

$$\sigma_n = \min_{X \text{ singular}} \|A - X\|_F \leq \|A - B\|_F = \frac{1}{2^{n-2}}$$

Logo, se $n = 100$, então $\sigma_n \leq 1/2^{98}$ e, portanto, A está muito próxima de ser singular. □

A inversa da matriz A no Exemplo 2 é dada por

$$A^{-1} = \begin{pmatrix} 1 & 1 & 2 & 4 & \cdots & 2^{n-2} \\ 0 & 1 & 1 & 2 & \cdots & 2^{n-3} \\ \vdots & & & & & \\ 0 & 0 & 0 & 0 & \cdots & 2^0 \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

Logo,

$$\text{cond}_\infty A = \|A\|_\infty \|A^{-1}\|_\infty = n2^{n-1}$$

Se n for grande, A vai ser extremamente mal condicionada. Isso não é de surpreender. Já vimos que A está muito próxima de uma matriz singular. Conseqüentemente, esperamos que uma pequena mudança nos elementos de A tenha efeitos drásticos. Parece razoável existir uma relação entre o número condicional e os valores singulares de uma matriz. Esse é, de fato, o caso quando o número condicional é definido em termos da norma 2 da matriz. Lembre-se de que

$$\|A\|_2 = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$$
$$\text{cond}_2(A) = \|A\|_2 \|A^{-1}\|_2$$

Teorema 7.6.4. *Se A é uma matriz $m \times n$ com decomposição em valores singulares $U\Sigma V^T$, então*

$$\|A\|_2 = \sigma_1 \qquad (\textit{o maior valor singular})$$

Demonstração. Como U e V são ortogonais,

$$\|A\|_2 = \|U\Sigma V^T\|_2 = \|\Sigma\|_2$$

Por outro lado,

$$\|\Sigma\|_2 = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|\Sigma\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$$

$$= \max_{\mathbf{x}\neq\mathbf{0}} \frac{\left(\sum_{i=1}^{n}(\sigma_i x_i)^2\right)^{1/2}}{\left(\sum_{i=1}^{n} x_i^2\right)^{1/2}}$$

$$\leq \sigma_1$$

Entretanto, se escolhermos $\mathbf{x} = \mathbf{e}_1$,

$$\frac{\|\Sigma\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sigma_1$$

e, portanto,

$$\|A\|_2 = \|\Sigma\|_2 = \sigma_1 \qquad \square$$

Corolário 7.6.5. *Se $A = U\Sigma V^T$ é invertível, então*

$$\text{cond}_2(A) = \frac{\sigma_1}{\sigma_n}$$

Demonstração. Os valores singulares de $A^{-1} = V\Sigma^{-1}U^T$, em ordem decrescente, são

$$\frac{1}{\sigma_n} \geq \frac{1}{\sigma_{n-1}} \geq \cdots \geq \frac{1}{\sigma_1}$$

Logo,

$$\|A^{-1}\|_2 = \frac{1}{\sigma_n} \quad \text{e} \quad \text{cond}_2(A) = \frac{\sigma_1}{\sigma_n} \qquad \square$$

Se duas matrizes A e B estão próximas, seus valores singulares também têm que estar próximos. Mais precisamente, se A tem valores singulares $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$ e B tem valores singulares $\omega_1 \geq \omega_2 \geq \ldots \geq \omega_n$, então

$$|\sigma_i - \omega_i| \leq \|A - B\|_2 \qquad i = 1, \ldots, n$$

(ver Cap. 6, Seção 3, Exercício 24). Logo, quando calculamos os valores singulares de uma matriz A, não precisamos nos preocupar com o fato de que pequenas variações nos elementos de A causem mudanças drásticas nos valores singulares calculados. Na Seção 8, vamos aprender um algoritmo para calcular os valores singulares de uma matriz. Vamos, também, ver como a decomposição em valores singulares pode ser usada para resolvermos problemas de mínimos quadráticos.

APLICAÇÃO: PROCESSAMENTO DIGITAL DE IMAGENS

Imagens de vídeo ou fotografias podem ser digitalizadas dividindo-se a imagem em um arranjo retangular de células (ou pixels*) e medindo-se o nível de cinza em cada célula. Essa informação pode ser armazenada e transmitida usando-se uma matriz $n \times n$ A. Os elementos de A são números não-negativos correspondendo às medidas dos níveis de cinza. Como o nível de cinza de uma célula está bem próximo, em geral, do nível de cinza das células vizinhas, é possível reduzir o espaço de armazenagem de n^2 para um múltiplo de n. Geralmente, a matriz A tem muitos valores singulares pequenos. Logo, A pode ser aproximada por uma matriz de posto muito menor.

Se A tem decomposição em valores singulares $U\Sigma V^T$, então A pode ser representada por uma expansão da forma

$$A = \sigma_1\mathbf{u}_1\mathbf{v}_1^T + \sigma_2\mathbf{u}_2\mathbf{v}_2^T + \cdots + \sigma_n\mathbf{u}_n\mathbf{v}_n^T$$

*Usa-se a mesma palavra em português e inglês; pixel é uma abreviação da expressão em inglês *picture elements*, que significa *elementos da imagem* (*N. T.*)

A matriz mais próxima do posto k é obtida truncando-se essa soma depois dos k primeiros termos:

$$A_k = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \cdots + \sigma_k \mathbf{u}_k \mathbf{v}_k^T$$

O espaço total de armazenagem para A_k é $k(2n + 1)$. Podemos escolher k bem menor do que n e ainda ter a imagem digital correspondente a A_k bastante próxima da original. Para escolhas típicas de k, o espaço necessário de armazenagem vai ser menos do que 20% do espaço necessário para toda a matriz A.

Damos, a seguir, uma lista de referências para leitura. As referências 4 e 5 usam a mesma abordagem usada aqui mas diferentes expansões para aproximar A.

REFERÊNCIAS

1. Andrews, H. C. e B. R. Hunt, *Digital Image Processing*, Englewood Cliffs, N.J: Prentice Hall, 1977.
2. Andrews, H. C. e C. L. Patterson, "Outer Product Expansions and Their Uses in Digital Image Processing", *American Mathematical Monthly*, *82*, 1975.
3. Moler, Cleve, "Numerical Linear Algebra", in *Numerical Analysis*, Proceedings of Symposia in Applied Mathematics XXII, American Mathematical Society, Providence, R.I., 1978.
4. Moler, C. B. e G. N. Stewart, "An Efficient Matrix Factorization for Digital Image Processing", manuscrito, Departamento de Ciência da Computação, Universidade de Maryland, 1979.
5. O'Leary, Diane P. e Shmuel Peleg, "Digital Image Compression by Outer Product Expansion", manuscrito, Departamento de Ciência da Computação, Universidade de Maryland, 1981.

EXERCÍCIOS

1. Mostre que A e A^T têm o mesmo número de valores singulares não-nulos. Qual a relação entre seus valores singulares?

2. Use o método do Exemplo 1 para encontrar a decomposição em valores singulares de cada uma das matrizes a seguir.

(a) $\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & -2 \\ 1 & 2 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 3 \\ 3 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$ (d) $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$

3. Para cada uma das matrizes no Exercício 2:
- (i) Determine seu posto;
- (ii) Determine sua norma 2;
- (iii) Encontre a matriz mais próxima (em relação à norma de Frobenius) com posto 1.

4. Seja

$$A = \begin{pmatrix} -2 & 8 & 20 \\ 14 & 19 & 10 \\ 2 & -2 & 1 \end{pmatrix} = \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} & 0 \\ \frac{4}{5} & \frac{3}{5} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 30 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

(a) Determine o valor de $\text{cond}_2(A)$.

(b) Encontre as matrizes mais próximas (em relação à norma de Frobenius) de A com postos 1 e 2.

5. A matriz

$$A = \begin{pmatrix} 2 & 5 & 4 \\ 6 & 3 & 0 \\ 6 & 3 & 0 \\ 2 & 5 & 4 \end{pmatrix}$$

tem decomposição em valores singulares

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 12 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{pmatrix}$$

(a) Use a decomposição em valores singulares para encontrar bases ortonormais para $I(A^T)$ e $N(A)$.

(b) Use a decomposição em valores singulares para encontrar bases ortonormais para $I(A)$ e $N(A^T)$.

6. Se A é uma matriz 5×3 com $\|A\|_2 = 8$, $\text{cond}_2(A) = 2$ e $\|A\|_F = 12$, determine os valores singulares de A.

7. Prove que, se A é uma matriz simétrica com autovalores $\lambda_1, \lambda_2, \ldots, \lambda_n$, então os valores singulares de A são $|\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|$.

8. Seja A uma matriz $m \times n$ com decomposição em valores singulares $U\Sigma V^T$ e suponha que A tem posto r, onde $r < n$. Mostre que $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ é uma base ortonormal para $I(A^T)$.

9. Seja A uma matriz $m \times n$ com decomposição em valores singulares $U\Sigma V^T$. Mostre que

$$\min_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sigma_n$$

10. Seja A uma matriz $m \times n$ com decomposição em valores singulares $U\Sigma V^T$. Mostre que, qualquer que seja o vetor $\mathbf{x} \in R^n$,

$$\sigma_n \|\mathbf{x}\|_2 \leq \|A\mathbf{x}\|_2 \leq \sigma_1 \|\mathbf{x}\|_2$$

11. Mostre que, se σ é um valor singular de A, então existe um vetor não-nulo $\mathbf{x}$ tal que

$$\sigma = \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$$

12. Seja A uma matriz $m \times n$ com decomposição em valores singulares $U\Sigma V^T$. Mostre que, se A tem posto r, $r < n$, então

$$A = U_1 \Sigma_1 V_1^T$$

onde $V_1 = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r)$, $U_1 = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r)$ e Σ_1 é uma matriz diagonal $r \times r$ com elementos diagonais $\sigma_1, \ldots, \sigma_r$.

13. Seja A uma matriz $m \times n$ e sejam $\mathbf{x} \in R^m$ e $\mathbf{y} \in R^n$ vetores não-nulos. Mostre que:

(a) $\dfrac{|\mathbf{x}^T A\mathbf{y}|}{\|\mathbf{x}\|\ \|\mathbf{y}\|} \leq \sigma_1$

(b) $\displaystyle\max_{\mathbf{x} \neq \mathbf{0},\, \mathbf{y} \neq \mathbf{0}} \frac{|\mathbf{x}^T A\mathbf{y}|}{\|\mathbf{x}\|\ \|\mathbf{y}\|} = \sigma_1$

14. Seja A uma matriz $m \times n$ de posto n com decomposição em valores singulares $U\Sigma V^T$. Denote por Σ^+ a matriz $n \times m$

$$\begin{pmatrix} \frac{1}{\sigma_1} & & & \\ & \frac{1}{\sigma_2} & & O \\ & & \ddots & \\ & & & \frac{1}{\sigma_n} \end{pmatrix}$$

e defina $A^+ = V\Sigma^+U^T$. Mostre que $\hat{\mathbf{x}} = A^+\mathbf{b}$ satisfaz as equações normais $A^TA\mathbf{x} = A^T\mathbf{b}$.

15. Seja A^+ como no Exercício 14 e seja $P = AA^+$. Mostre que $P^2 = P$ e que $P^T = P$.

16. Seja $A = \mathbf{x}\mathbf{y}^T$, onde $\mathbf{x} \in R^m$, $\mathbf{y} \in R^n$ e ambos, $\mathbf{x}$ e $\mathbf{y}$, são vetores não-nulos. Mostre que A tem uma decomposição em valores singulares da forma $H_1\Sigma H_2$, onde H_1 e H_2 são matrizes de Householder e

$$\sigma_1 = \|\mathbf{x}\|\,\|\mathbf{y}\|, \qquad \sigma_2 = \sigma_3 = \cdots = \sigma_n = 0$$

7 O PROBLEMA DE AUTOVALORES

Nesta seção, vamos nos preocupar com métodos numéricos para calcular os autovalores e autovetores de uma matriz A $n \times n$. O primeiro método que estudaremos é chamado de método das potências. O método das potências é um método iterativo para encontrar o autovalor dominante de uma matriz e um autovetor associado. O autovalor dominante é um autovalor λ satisfazendo $|\lambda_1| > |\lambda_i|$ para $i = 2, \ldots, n$. Se os autovalores de A satisfazem

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$$

então o método das potências pode ser usado para calcular um autovalor de cada vez. O segundo método é chamado de algoritmo QR. O algoritmo QR é um método iterativo que usa transformações ortogonais de semelhança e tem muitas vantagens sobre o método das potências. Ele vai convergir independentemente de A ter ou não um autovalor dominante e calcula todos os autovalores ao mesmo tempo.

Nos exemplos do Cap. 6, os autovalores foram determinados formando-se o polinômio característico e encontrando-se suas raízes. No entanto, esse procedimento não é recomendado, em geral, para cálculos numéricos. A dificuldade com esse procedimento é que, com freqüência, uma pequena variação em um ou mais coeficientes do polinômio característico pode resultar em uma mudança relativamente grande nos zeros calculados do polinômio. Por exemplo, considere o polinômio $p(x) = x^{10}$. O coeficiente do termo de maior grau é 1 e os outros coeficientes são todos iguais a 0. Se o termo constante é alterado somando-se -10^{-10}, obtemos o polinômio $q(x) = x^{10} - 10^{-10}$. Embora os coeficientes de $p(x)$ e de $q(x)$ difiram por apenas 10^{-10}, as raízes de $q(x)$ têm valor absoluto $\frac{1}{10}$, enquanto as raízes de $p(x)$ são todas iguais a 0. Logo, mesmo quando os coeficientes do polinômio característico foram determinados com bastante precisão, os autovalores calculados podem envolver erros significativos. Por essa razão, os métodos apresentados nesta seção não usam o polinômio característico. Para ver que existe alguma vantagem de se trabalhar diretamente com a matriz A, temos que determinar o efeito de pequenas variações nos elementos de A sobre os autovalores. O próximo teorema trata disso.

Teorema 7.7.1. *Seja A uma matriz $n \times n$ com n autovetores linearmente independentes e seja X uma matriz que diagonaliza A.*

$$X^{-1}AX = D = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

Se $A' = A + E$ e λ' é um autovalor de A', então

(1) $$\min_{1\le i\le n} |\lambda' - \lambda_i| \le \text{cond}_2(X)\|E\|_2$$

Demonstração. Podemos supor que λ' é diferente de todos os outros λ_i (caso contrário, não há nada a provar). Então, definindo $D_1 = D - \lambda' I$, temos que D_1 é uma matriz diagonal invertível. Como λ' é um autovalor de A', ele é, também, um autovalor de $X^{-1}A'X$. Portanto, $X^{-1}A'X - \lambda' I$ é singular e $D_1^{-1}(X^{-1}A'X - \lambda' I)$ também é singular. Por outro lado,

$$\begin{aligned} D_1^{-1}(X^{-1}A'X - \lambda' I) &= D_1^{-1}X^{-1}(A + E - \lambda' I)X \\ &= D_1^{-1}X^{-1}EX + I \end{aligned}$$

Logo, -1 é um autovalor de $D_1^{-1}X^{-1}EX$. Segue que

$$|-1| \le \|D_1^{-1}X^{-1}EX\|_2 \le \|D_1^{-1}\|_2 \,\text{cond}_2(X)\|E\|_2$$

A norma 2 de D_1^{-1} é dada por

$$\|D_1^{-1}\|_2 = \max_{1\le i\le n} |\lambda' - \lambda_i|^{-1}$$

O índice i que maximiza $|\lambda' - \lambda_i|^{-1}$ é o mesmo que minimiza $|\lambda' - \lambda_i|$. Logo,

$$\min_{1\le i\le n} |\lambda' - \lambda_i| \le \text{cond}_2(X)\|E\|_2$$ □

Se a matriz A é simétrica, podemos escolher uma matriz diagonalizadora ortogonal. Em geral, se Q é qualquer matriz ortogonal, temos

$$\text{cond}_2(Q) = \|Q\|_2\|Q^{-1}\|_2 = 1$$

Logo, (1) pode ser simplificada para

$$\min_{1\le i\le n} |\lambda' - \lambda_i| \le \|E\|_2$$

Portanto, se A é simétrica e $\|E\|_2$ é pequena, os autovalores de A' vão estar próximos dos autovalores de A.

Agora estamos prontos para falar de alguns métodos para calcular os autovalores e autovetores de uma matriz A $n \times n$. O primeiro método que apresentaremos calcula um autovetor $\mathbf{x}$ de A aplicando A sucessivamente a um vetor dado em R^n. Para ver a idéia por trás do método, vamos supor que A tem n autovetores linearmente independentes $\mathbf{x}_1, \ldots, \mathbf{x}_n$ e que os autovalores associados satisfazem

(2) $$|\lambda_1| > |\lambda_2| \ge \cdots \ge |\lambda_n|$$

Dado um vetor arbitrário $\mathbf{v}_0$ em R^n, podemos escrever

$$\begin{aligned} \mathbf{v}_0 &= \alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n \\ A\mathbf{v}_0 &= \alpha_1\lambda_1\mathbf{x}_1 + \alpha_2\lambda_2\mathbf{x}_2 + \cdots + \alpha_n\lambda_n\mathbf{x}_n \\ A^2\mathbf{v}_0 &= \alpha_1\lambda_1^2\mathbf{x}_1 + \alpha_2\lambda_2^2\mathbf{x}_2 + \cdots + \alpha_n\lambda_n^2\mathbf{x}_n \end{aligned}$$

e, em geral,

$$A^k\mathbf{v}_0 = \alpha_1\lambda_1^k\mathbf{x}_1 + \alpha_2\lambda_2^k\mathbf{x}_2 + \cdots + \alpha_n\lambda_n^k\mathbf{x}_n$$

Definindo

$$\mathbf{v}_k = A^k\mathbf{v}_0 \qquad k = 1, 2, \ldots$$

temos

$$\text{(3)}\qquad \frac{1}{\lambda_1^k}\mathbf{v}_k = \alpha_1\mathbf{x}_1 + \alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^k \mathbf{x}_2 + \cdots + \alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^k \mathbf{x}_n$$

Como

$$\left|\frac{\lambda_i}{\lambda_1}\right| < 1 \quad \text{para} \quad i = 2, 3, \ldots, n$$

segue que

$$\frac{1}{\lambda_1^k}\mathbf{v}_k \to \alpha_1\mathbf{x}_1 \quad \text{quando} \quad k \to \infty$$

Logo, se $\alpha_1 \neq 0$, a seqüência $\{(1/\lambda_1^k)\mathbf{v}_k\}$ converge a um autovetor $\alpha_1\mathbf{x}_1$ de A. Existem alguns problemas óbvios com o método do modo que foi apresentado até agora. Não podemos calcular $(1/\lambda_1^k)\mathbf{v}_k$, já que λ_1 não é conhecido. Mesmo se conhecêssemos λ_1, ainda teríamos dificuldade, pois λ_1^k pode tender a 0 ou a $\pm\infty$. Afortunadamente, no entanto, não precisamos multiplicar a seqüência $\{\mathbf{v}_k\}$ usando $1/\lambda_1^k$. Se multiplicarmos os $\mathbf{v}_k$ por um escalar de modo a obter um vetor unitário em cada etapa, a seqüência vai convergir para um vetor unitário com mesma direção do que $\mathbf{x}_1$. O autovalor λ_1 pode ser calculado ao mesmo tempo. Esse método para calcular o autovalor de maior módulo e o autovetor associado é chamado de *método das potências*. Ele pode ser resumido da seguinte maneira:

O MÉTODO DAS POTÊNCIAS

Vamos definir duas seqüências $\{\mathbf{v}_k\}$ e $\{\mathbf{u}_k\}$ recursivamente. Para começar, $\mathbf{u}_0$ pode ser qualquer vetor não-nulo em R^n. Uma vez determinado $\mathbf{u}_k$, os vetores $\mathbf{v}_{k+1}$ e $\mathbf{u}_{k+1}$ são calculados da seguinte maneira:

1. $\mathbf{v}_{k+1} = A\mathbf{u}_k$.
2. Encontre a coordenada j_{k+1} de $\mathbf{v}_{k+1}$ de módulo máximo.
3. $\mathbf{u}_{k+1} = (1/v_{j_{k+1}})\mathbf{v}_{k+1}$.

A seqüência $\{\mathbf{u}_k\}$ tem a propriedade de que, para todo $k \geq 1$, $\|\mathbf{u}_k\|_\infty = u_{j_k} = 1$. Se os autovalores de A satisfizerem (2) e $\mathbf{u}_0$ puder ser escrito como uma combinação linear $\alpha_1\mathbf{x}_1 + \ldots \alpha_n\mathbf{x}_n$ dos autovetores com $\alpha_1 \neq 0$, a seqüência $\{\mathbf{u}_k\}$ irá convergir para um autovetor $\mathbf{y}$ associado a λ_1. Se k for grande, então $\mathbf{u}_k$ será uma boa aproximação de $\mathbf{y}$ e $\mathbf{v}_{k+1} = A\mathbf{u}_k$ será uma boa aproximação de $\lambda_1\mathbf{y}$. Como a j_k-ésima coordenada de $\mathbf{u}_k$ é 1, temos que a j_k-ésima coordenada de $\mathbf{v}_{k+1}$ será uma boa aproximação de λ_1.

Em vista de (3), podemos esperar que os $\mathbf{u}_k$ irão convergir para $\mathbf{y}$ à mesma velocidade com que $(\lambda_2/\lambda_1)^k$ converge para 0. Logo, se $|\lambda_2|$ for quase tão grande quanto $|\lambda_1|$, a convergência será lenta.

EXEMPLO 1. Seja

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

É fácil determinar exatamente os autovalores de A. Eles são $\lambda_1 = 3$ e $\lambda_2 = 1$, com autovetores associados $\mathbf{x}_1 = (1, 1)^T$ e $\mathbf{x}_2 = (1, -1)^T$. Para ilustrar a convergência dos vetores gerados pelo método das potências, vamos aplicar o método com $\mathbf{u}_0 = (2, 1)^T$.

$$\mathbf{v}_1 = A\mathbf{u}_0 = \begin{pmatrix} 5 \\ 4 \end{pmatrix}, \qquad \mathbf{u}_1 = \frac{1}{5}\mathbf{v}_1 = \begin{pmatrix} 1{,}0 \\ 0{,}8 \end{pmatrix}$$

$$\mathbf{v}_2 = A\mathbf{u}_1 = \begin{pmatrix} 2{,}8 \\ 2{,}6 \end{pmatrix}, \qquad \mathbf{u}_2 = \frac{1}{2{,}8}\mathbf{v}_2 = \begin{pmatrix} 1 \\ \frac{13}{14} \end{pmatrix} \approx \begin{pmatrix} 1{,}00 \\ 0{,}93 \end{pmatrix}$$

$$\mathbf{v}_3 = A\mathbf{u}_2 = \frac{1}{14}\begin{pmatrix} 41 \\ 40 \end{pmatrix}, \qquad \mathbf{u}_3 = \frac{14}{41}\mathbf{v}_3 = \begin{pmatrix} 1 \\ \frac{40}{41} \end{pmatrix} \approx \begin{pmatrix} 1{,}00 \\ 0{,}98 \end{pmatrix}$$

$$\mathbf{v}_4 = A\mathbf{u}_3 \approx \begin{pmatrix} 2{,}98 \\ 2{,}95 \end{pmatrix}$$

Se escolhemos $\mathbf{u}_3 = (1{,}00,\ 0{,}98)^T$ com autovetor aproximado, então 2,98 é o valor aproximado de λ_1. Logo, com algumas poucas iterações, a aproximação de λ_1 envolve um erro de apenas 0,02. □

O método das potências pode ser usado para calcular o autovalor λ_1 de maior módulo e um autovetor associado $\mathbf{y}_1$. E os autovalores e autovetores restantes? Se pudéssemos reduzir o problema de encontrar os outros autovalores de A ao de encontrar os autovalores de alguma matriz A_1 $(n-1) \times (n-1)$, então o método das potências poderia ser aplicado a A_1. Isso, na verdade, pode ser feito por um processo chamado *deflação*.

DEFLAÇÃO

A idéia básica da deflação é encontrar uma matriz invertível H tal que HAH^{-1} seja uma matriz da forma

$$\left(\begin{array}{c|ccc} \lambda_1 & \times & \cdots & \times \\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array}\right) \tag{4}$$

Como A e HAH^{-1} são semelhantes, elas têm o mesmo polinômio característico. Logo, se HAH^{-1} é da forma (4), temos

$$\det(A - \lambda I) = \det(HAH^{-1} - \lambda I) = (\lambda_1 - \lambda)\det(A_1 - \lambda I)$$

e segue que os $n - 1$ autovalores restantes de A são os autovalores de A_1. O problema continua: como encontrar uma tal matriz H? Note que a forma (4) requer que a primeira coluna de HAH^{-1} seja $\lambda_1\mathbf{e}_1$. Mas a primeira coluna de HAH^{-1} é $HAH^{-1}\mathbf{e}_1$, logo

$$HAH^{-1}\mathbf{e}_1 = \lambda_1\mathbf{e}_1$$

ou, equivalentemente,

$$A(H^{-1}\mathbf{e}_1) = \lambda_1(H^{-1}\mathbf{e}_1)$$

Portanto, $H^{-1}\mathbf{e}_1$ pertence ao auto-espaço associado a λ_1 e, para algum autovetor $\mathbf{x}_1$ associado a λ_1,

$$H^{-1}\mathbf{e}_1 = \mathbf{x}_1 \qquad \text{ou} \qquad H\mathbf{x}_1 = \mathbf{e}_1$$

Precisamos encontrar uma matriz H tal que $H\mathbf{x}_1 = \mathbf{e}_1$ para algum autovetor associado a λ_1. Isso pode ser feito através de uma transformação de Householder. Se $\mathbf{y}_1$ é o autovetor calculado associado a λ_1, defina

$$\mathbf{x}_1 = \frac{1}{\|\mathbf{y}_1\|_2}\mathbf{y}_1$$

Como $\|\mathbf{x}_1\|_2 = 1$, podemos encontrar uma transformação de Householder H tal que

$$H\mathbf{x}_1 = \mathbf{e}_1$$

Como H é uma transformação de Householder, temos que $H^{-1} = H$ e, portanto, HAH é a transformação de semelhança desejada.

REDUÇÃO À FORMA DE HESSENBERG

Os métodos padrões para encontrar autovalores são todos iterativos. A quantidade de trabalho em cada etapa é, muitas vezes, absurdamente alta, a menos que A seja, inicialmente, de uma forma especial que seja fácil de trabalhar. Se não for esse o caso, o procedimento usual é colocar A em uma forma mais simples usando-se transformações de semelhança. Em geral, as matrizes de Householder são usadas para colocar A na forma

$$
\begin{pmatrix}
\times & \times & \cdots & \times & \times & \times \\
\times & \times & \cdots & \times & \times & \times \\
0 & \times & \cdots & \times & \times & \times \\
0 & 0 & \cdots & \times & \times & \times \\
\vdots & & & & & \\
0 & 0 & \cdots & \times & \times & \times \\
0 & 0 & \cdots & 0 & \times & \times
\end{pmatrix}
$$

Uma matriz dessa forma é chamada de uma *matriz de Hessenberg superior*. Logo, B é uma matriz de Hessenberg superior se e somente se $b_{ij} = 0$ sempre que $i \geq j + 2$.

Uma matriz A pode ser colocada em uma forma de Hessenberg superior da maneira descrita a seguir. Escolha, primeiro, uma matriz de Householder H_1 tal que H_1A seja da forma

$$
\begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
\times & \times & \cdots & \times \\
0 & \times & \cdots & \times \\
\vdots & & & \\
0 & \times & \cdots & \times
\end{pmatrix}
$$

A matriz H_1 vai ser da forma

$$
\begin{pmatrix}
1 & 0 & \cdots & 0 \\
0 & \times & \cdots & \times \\
\vdots & & & \\
0 & \times & \cdots & \times
\end{pmatrix}
$$

e, portanto, a multiplicação de H_1A à direita por H_1 vai deixar a primeira coluna sem mudanças. Se $A^{(1)} = H_1AH_1$, então $A^{(1)}$ é da forma

$$
\begin{pmatrix}
a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\
a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\
0 & a_{32}^{(1)} & \cdots & a_{3n}^{(1)} \\
\vdots & & & \\
0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)}
\end{pmatrix}
$$

Como H_1 é uma matriz de Householder, temos que $H_1^{-1} = H_1$ e $A^{(1)}$ é semelhante a A. A seguir, escolhemos uma matriz de Householder H_2 tal que

$$H_2(a_{12}^{(1)}, a_{22}^{(1)}, \ldots, a_{n2}^{(1)})^T = (a_{12}^{(1)}, a_{22}^{(1)}, \times, 0, \ldots, 0)^T$$

A matriz H_2 vai ser da forma

$$
\begin{pmatrix}
1 & 0 & 0 & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & \times & \cdots & \times \\
\vdots & & & & \\
0 & 0 & \times & \cdots & \times
\end{pmatrix}
= \left(\begin{array}{c|c} I_2 & O \\ \hline O & X \end{array}\right)
$$

A multiplicação à esquerda de $A^{(1)}$ por H_2 não vai modificar as duas primeiras linhas e a primeira coluna.

$$H_2A^{(1)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} \\ 0 & \times & \times & \cdots & \times \\ 0 & 0 & \times & \cdots & \times \\ \vdots & & & & \\ 0 & 0 & \times & \cdots & \times \end{pmatrix}$$

A multiplicação de $H_2A^{(1)}$ à direita por H_2 não vai modificar as duas primeiras colunas, logo $A^{(2)} = H_2A^{(1)} H_2$ tem a forma

$$\begin{pmatrix} \times & \times & \times & \cdots & \times \\ \times & \times & \times & \cdots & \times \\ 0 & \times & \times & \cdots & \times \\ 0 & 0 & \times & \cdots & \times \\ \vdots & & & & \\ 0 & 0 & \times & \cdots & \times \end{pmatrix}$$

Esse processo pode ser continuado até se obter uma matriz de Hessenberg superior

$$H = A^{(n-2)} = H_{n-2} \cdots H_2 H_1 A H_1 H_2 \cdots H_{n-2}$$

que é semelhante a A.

Se, em particular, A for simétrica, como

$$\begin{aligned} H^T &= H_{n-2}^T \cdots H_2^T H_1^T A^T H_1^T H_2^T \cdots H_{n-2}^T \\ &= H_{n-2} \cdots H_2 H_1 A H_1 H_2 \cdots H_{n-2} \\ &= H \end{aligned}$$

segue que H é tridiagonal. Logo, qualquer matriz A $n \times n$ pode ser colocada, através de transformações de semelhança, na forma de Hessenberg superior. Se A for simétrica, a matriz obtida será uma matriz simétrica tridiagonal.

Vamos encerrar esta seção esboçando um dos melhores métodos disponíveis para se calcular os autovalores de uma matriz. O método é chamado de algoritmo QR e foi apresentado por K. G. F. Francis em 1961.

ALGORITMO *QR*

Dada uma matriz A $n \times n$, fatore-a em um produto Q_1R_1, onde Q_1 é ortogonal e R_1 é triangular superior. Defina

$$A_1 = A = Q_1 R_1$$

e

$$A_2 = Q_1^T A Q_1 = R_1 Q_1$$

Fatore A_2 em um produto Q_2R_2, onde Q_2 é ortogonal e R_2 é triangular superior. Defina

$$A_3 = Q_2^T A_2 Q_2 = R_2 Q_2$$

Note que $A_2 = Q_1^T AQ_1$ e $A_3 = (Q_1Q_2)^T A(Q_1Q_2)$ são ambas semelhantes a A. Podemos continuar dessa maneira para obter uma seqüência de matrizes semelhantes. Em geral, se

$$A_k = Q_k R_k$$

então A_{k+1} é definida como sendo R_kQ_k. Pode-se mostrar que, sob condições bem gerais, a seqüência de matrizes definida dessa maneira converge para uma matriz da forma

$$\begin{pmatrix} B_1 & \times & \cdots & \times \\ & B_2 & & \times \\ & O & \ddots & \\ & & & B_s \end{pmatrix}$$

onde os B_i são blocos diagonais 1×1 ou 2×2. Cada bloco 2×2 vai corresponder a um par de autovalores de A complexos conjugados. Os autovalores de A serão os autovalores dos B_i. No caso em que A é simétrica, cada uma das A_k também será simétrica e a seqüência vai convergir para uma matriz diagonal.

EXEMPLO 2. Seja A_1 a matriz do Exemplo 1. A fatoração QR de A_1 precisa de apenas uma transformação de Givens.

$$G_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix}$$

Logo,

$$A_2 = G_1AG_1 = \frac{1}{5}\begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} 2{,}8 & -0{,}6 \\ -0{,}6 & 1{,}2 \end{pmatrix}$$

A fatoração QR de A_2 pode ser obtida através da transformação de Givens

$$G_2 = \frac{1}{\sqrt{8{,}2}}\begin{pmatrix} 2{,}8 & -0{,}6 \\ -0{,}6 & -2{,}8 \end{pmatrix}$$

Segue que

$$A_3 = G_2A_2G_2 \approx \begin{pmatrix} 2{,}98 & 0{,}22 \\ 0{,}22 & 1{,}02 \end{pmatrix}$$

Os elementos fora da diagonal ficam cada vez mais próximos de 0 após cada iteração e os elementos diagonais se aproximam dos autovalores $\lambda_1 = 3$ e $\lambda_2 = 1$. □

Observações

1. Devido à quantidade de trabalho em cada iteração do algoritmo QR, é importante que a matriz inicial A seja de Hessenberg ou simétrica tridiagonal. Se não for esse o caso, devemos usar transformações de semelhança em A para obter uma matriz A_1 de um desses tipos.
2. Se A_k é uma matriz de Hessenberg superior, a fatoração QR pode ser feita usando-se $n-1$ transformações de Givens.

$$G_{n,n-1}\cdots G_{32}G_{21}A_k = R_k$$

Definindo

$$Q_k^T = G_{n,n-1}\cdots G_{32}G_{21}$$

temos

$$A_k = Q_kR_k$$

e

$$A_{k+1} = Q_k^T A_k Q_k$$

Para calcular A_{k+1} não é necessário determinar Q_k explicitamente. Precisamos, apenas, guardar as $n - 1$ transformações de Givens. Quando R_k é multiplicada à direita por G_{21}, a matriz resultante vai ter o elemento (2, 1) não-nulo. Os outros elementos abaixo da diagonal ainda serão nulos. A multiplicação à direita de R_kG_{21} por G_{32} tem o efeito de colocar um elemento não-nulo na posição (3, 2). A multiplicação à direita de $R_kG_{21}G_{32}$ por G_{43} vai colocar um elemento não-nulo na posição (4, 3), e assim por diante. Logo, a matriz resultante $A_{k+1} = R_kG_{21}G_{32}\ldots G_{n,n-1}$ será de Hessenberg superior. Se A_1 for uma matriz tridiagonal simétrica, cada uma das sucessivas A_i será de Hessenberg superior e simétrica. Logo, $A_2, A_3, \ldots$ serão todas tridiagonais.

3. Como com o método das potências, a convergência pode ser lenta quando alguns dos autovalores estão próximos. Para acelerar a convergência, costuma-se fazer *deslocamentos da origem*. Na k-ésima etapa, escolhe-se um escalar α_k e fatora-se $A_k - \alpha_k I$ (em vez de A_k) em um produto Q_kR_k. A matriz A_{k+1} é definida por

$$A_{k+1} = R_k Q_k + \alpha_k I$$

Note que

$$Q_k^T A_k Q_k = Q_k^T (Q_k R_k + \alpha_k I) Q_k = R_k Q_k + \alpha_k I = A_{k+1}$$

de modo que A_k e A_{k+1} são semelhantes. Com a escolha apropriada dos escalares α_k, a convergência pode ser bastante acelerada.

4. Apresentamos apenas um esboço do método em nossa discussão rápida. Para uma discussão mais completa e uma demonstração de convergência, ver Wilkinson [27, Cap. 8].

EXERCÍCIOS

1. Seja

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

(a) Faça uma etapa do método das potências com qualquer vetor inicial não-nulo.
(b) Faça uma iteração do algoritmo QR para a matriz A.
(c) Determine os autovalores exatos de A calculando as raízes do polinômio característico e determine o auto-espaço associado ao maior autovalor. Compare suas respostas com as dos itens (a) e (b).

2. Sejam

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 2 \end{pmatrix} \quad \text{e} \quad \mathbf{u}_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

(a) Use o método das potências para calcular $\mathbf{v}_1$, $\mathbf{u}_1$, $\mathbf{v}_2$, $\mathbf{u}_2$ e $\mathbf{v}_3$. (Arredonde para duas casas decimais.)
(b) Determine uma aproximação λ_1' do maior autovalor de A a partir das coordenadas de $\mathbf{v}_3$. Determine o valor exato de λ_1 e compare com λ_1'. Qual o erro relativo?

3. Sejam

$$A = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix} \quad \text{e} \quad \mathbf{u}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

(a) Calcule $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ e $\mathbf{u}_4$ usando o método das potências.

(b) Explique por que o método das potências vai falhar nesse caso.

4. Seja

$$A = A_1 = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}$$

Calcule A_2 e A_3 usando o algoritmo *QR*. Calcule os autovalores exatos de A e compare-os com os elementos diagonais de A_3. Eles são iguais até quantas casas decimais?

5. Seja

$$A = \begin{pmatrix} 5 & 2 & 2 \\ -2 & 1 & -2 \\ -3 & -4 & 2 \end{pmatrix}$$

(a) Verifique que $\lambda_1 = 4$ é um autovalor de A e que $\mathbf{y}_1 = (2, -2, 1)^T$ é um autovetor associado a λ_1.

(b) Encontre uma transformação de Householder H tal que HAH seja da forma

$$\begin{pmatrix} 4 & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{pmatrix}$$

(c) Calcule HAH e encontre os autovalores restantes de A.

6. Seja A uma matriz $n \times n$ com autovalores reais e distintos $\lambda_1, \lambda_2, \ldots, \lambda_n$. Seja λ um escalar que não é um autovalor de A e seja $B = (A - \lambda I)^{-1}$. Mostre que:

(a) Os escalares $\mu_j = 1/(\lambda_j - \lambda), j = 1, \ldots, n$ são os autovalores de B.

(b) Se $\mathbf{x}_j$ é um autovetor de B associado a μ_j, então $\mathbf{x}_j$ é um autovetor de A associado a λ_j.

(c) Se o método das potências for usado em B, então a seqüência de vetores irá convergir para um autovetor de A associado ao autovalor mais próximo de λ. [A convergência será bem rápida se λ estiver muito mais próximo de um dos λ_i do que dos outros. Esse método de calcular autovetores usando as potências de $(A - \lambda I)^{-1}$ é chamado de *método das potências inversas*.]

7. Seja $\mathbf{x} = (x_1, \ldots, x_n)^T$ um autovetor de A associado a λ. Se $|x_i| = \|\mathbf{x}\|_\infty$, mostre que:

(a) $$\sum_{j=1}^{n} a_{ij}x_j = \lambda x_i$$

(b) $$|\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \qquad \text{(Teorema de Gerschgorin)}$$

8. Seja A uma matriz com autovalores $\lambda_1, \ldots, \lambda_n$ e seja λ um autovalor de $A + E$. Seja X uma matriz que diagonaliza A e seja $C = X^{-1}EX$. Prove que:

(a) Para algum i,

$$|\lambda - \lambda_i| \leq \sum_{j=1}^{n} |c_{ij}|$$

[***Sugestão:*** λ é um autovalor de $X^{-1}(A + E)X$. Aplique o teorema de Gerschgorin do Exercício 7.]

(b) $$\min_{1 \leq j \leq n} |\lambda - \lambda_j| \leq \text{cond}_\infty(X)\|E\|_\infty$$

9. Seja $A_k = Q_k R_k$, $k = 1, 2, \ldots$ a seqüência de matrizes obtidas de $A = A_1$ aplicando-se o algoritmo *QR*. Para cada inteiro positivo k, defina

$$P_k = Q_1 Q_2 \cdots Q_k \qquad \text{e} \qquad U_k = R_k \cdots R_2 R_1$$

Mostre que

$$P_k A_{k+1} = A P_k$$

para todo $k \geq 1$.

10. Sejam P_k e U_k como definidos no Exercício 9. Mostre que:
(a) $P_{k+1}U_{k+1} = P_k A_{k+1} U_k = AP_k U_k$;
(b) $P_k U_k = A^k$ e, portanto,

$$(Q_1 Q_2 \cdots Q_k)(R_k \cdots R_2 R_1)$$

é a fatoração QR de A^k.

11. Seja R_k uma matriz triangular superior $k \times k$ e suponha que

$$R_k U_k = U_k D_k$$

onde U_k é uma matriz triangular superior com todos os elementos diagonais iguais a 1 e D_k é uma matriz diagonal. Seja R_{k+1} uma matriz triangular superior da forma

$$\begin{pmatrix} R_k & \mathbf{b}_k \\ \mathbf{0}^T & \beta_k \end{pmatrix}$$

onde β_k não é um autovalor de R_k. Determine matrizes U_{k+1} e D_{k+1} $(k + 1) \times (k + 1)$ da forma

$$U_{k+1} = \begin{pmatrix} U_k & \mathbf{x}_k \\ \mathbf{0}^T & 1 \end{pmatrix}, \qquad D_{k+1} = \begin{pmatrix} D_k & \mathbf{0} \\ \mathbf{0}^T & \beta \end{pmatrix}$$

tais que

$$R_{k+1}U_{k+1} = U_{k+1}D_{k+1}$$

12. Seja R uma matriz triangular superior $n \times n$ cujos elementos diagonais são todos distintos. Denote por R_k a submatriz principal de R de ordem k e defina $U_1 = (1)$.
(a) Use o resultado do Exercício 11 para encontrar um algoritmo que ache os autovetores de R. A matriz U de autovetores deveria ser triangular superior com todos os elementos diagonais iguais a 1.
(b) Mostre que o algoritmo necessita de aproximadamente $n^3/6$ multiplicações/divisões em ponto flutuante.

8 PROBLEMAS DE MÍNIMOS QUADRÁTICOS

Nesta seção, vamos estudar métodos computacionais para encontrar soluções de mínimos quadráticos para sistemas com mais equações do que incógnitas. Seja A uma matriz $m \times n$ com $m \geq n$ e seja $\mathbf{b} \in R^m$. Vamos considerar alguns métodos para calcular um vetor $\hat{\mathbf{x}}$ que minimize $\| \mathbf{b} - A\mathbf{x} \|_2^2$.

EQUAÇÕES NORMAIS

Vimos no Cap. 5 que, se $\hat{\mathbf{x}}$ satisfaz as equações normais

$$A^T A\mathbf{x} = A^T\mathbf{b}$$

então $\hat{\mathbf{x}}$ é uma solução do problema de mínimos quadráticos. Se A tiver posto máximo (posto n), então A^TA é invertível e, portanto, o sistema vai ter uma única solução. Logo, se A^TA for invertível, um método possível para resolver o problema de mínimos quadráticos é formar as equações normais e depois resolvê-las usando o método de Gauss. Um algoritmo que faça isso deveria ter duas partes principais.

1. Calcule $B = A^TA$ e $\mathbf{c} = A^T\mathbf{b}$.
2. Resolva $B\mathbf{x} = \mathbf{c}$.

A formação das equações normais necessita aproximadamente de $mn^2/2$ multiplicações. Assim A^TA é invertível, a matriz B é positiva definida. Para matrizes positiva definidas, existem algoritmos que necessitam de metade do número usual de multiplicações. Assim a solução de $B\mathbf{x} = \mathbf{c}$ necessita aproximadamente de $n^3/6$ multiplicações. A maior parte do trabalho, então, é gasta na formação das equações normais, não na resolução. Entretanto, o maior problema com esse método é que, ao formar as equações normais, podemos perfeitamente terminar transformando o problema em um problema malcondicionado. Lembre da Seção 4 que, se $\mathbf{x}'$ é a solução calculada para o problema $B\mathbf{x} = \mathbf{c}$ e $\mathbf{x}$ é a solução exata, então a desigualdade

$$\frac{1}{\text{cond}(B)}\frac{\|\mathbf{r}\|}{\|\mathbf{c}\|} \leq \frac{\|\mathbf{x}-\mathbf{x}'\|}{\|\mathbf{x}\|} \leq \text{cond}(B)\frac{\|\mathbf{r}\|}{\|\mathbf{c}\|}$$

mostra a relação entre o erro relativo e o resíduo relativo. Se A tem valores singulares $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$, então $\text{cond}_2(A) = \sigma_1/\sigma_n$. Os valores singulares de B são $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$. Logo,

$$\text{cond}_2(B) = \frac{\sigma_1^2}{\sigma_n^2} = [\text{cond}_2(A)]^2$$

Se, por exemplo, $\text{cond}_2(A) = 100$, o erro relativo na solução calculada para as equações normais deve ser 10^4 vezes maior do que o resíduo relativo. Por essa razão, devemos tomar muito cuidado ao usar as equações normais para resolver problemas de mínimos quadráticos.

No Cap. 5, vimos como usar o processo de Gram-Schmidt para obter uma fatoração QR de uma matriz A de posto máximo. Nesse caso, a matriz Q era uma matriz $m \times n$ com colunas ortogonais e R era uma matriz $n \times n$ triangular superior. No método numérico a seguir, vamos usar transformações de Householder para obter uma fatoração QR de A. Nesse caso, Q vai ser uma matriz ortogonal $m \times m$ e R vai ser uma matriz $m \times n$ com todos os elementos abaixo da diagonal iguais a 0.

A FATORAÇÃO QR

Dada uma matriz $m \times n$ de posto máximo, podemos aplicar n transformações de Householder para anular os elementos abaixo da diagonal. Logo,

$$H_nH_{n-1}\cdots H_1A = R$$

onde R é da forma

$$\begin{pmatrix} R_1 \\ O \end{pmatrix} = \begin{pmatrix} \times & \times & \times & \cdots & \times \\ & \times & \times & \cdots & \times \\ & & \times & \cdots & \times \\ & & & \ddots & \vdots \\ & & & & \times \\ & & & & \end{pmatrix}$$

com elementos diagonais não-nulos. Seja

$$Q^T = H_n\cdots H_1 = \begin{pmatrix} Q_1^T \\ Q_2^T \end{pmatrix}$$

onde Q_1^T é uma matriz $n \times m$ consistindo nas n primeiras linhas de Q^T. Como

$$Q^TA = \begin{pmatrix} Q_1^TA \\ Q_2^TA \end{pmatrix} = \begin{pmatrix} R_1 \\ O \end{pmatrix}$$

segue que $A = Q_1R_1$. Seja

$$\mathbf{c} = Q^T\mathbf{b} = \begin{pmatrix} Q_1^T\mathbf{b} \\ Q_2^T\mathbf{b} \end{pmatrix} = \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{pmatrix}$$

As equações normais podem ser colocadas na forma

$$R_1^T Q_1^T Q_1 R_1 \mathbf{x} = R_1^T Q_1^T \mathbf{b}$$

Como $Q_1^T Q_1 = I$ e R_1^T é invertível, a expressão pode ser simplificada para

$$R_1\mathbf{x} = \mathbf{c}_1$$

Esse sistema pode ser resolvido por substituição de baixo para cima. A solução $\mathbf{x} = R_1^{-1}\mathbf{c}_1$ é a única solução do problema de mínimos quadráticos. Para calcular o resíduo, note que

$$Q^T\mathbf{r} = \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{pmatrix} - \begin{pmatrix} R_1 \\ O \end{pmatrix}\mathbf{x} = \begin{pmatrix} \mathbf{0} \\ \mathbf{c}_2 \end{pmatrix}$$

de modo que

$$\mathbf{r} = Q\begin{pmatrix} \mathbf{0} \\ \mathbf{c}_2 \end{pmatrix} \quad \text{e} \quad \|\mathbf{r}\|_2 = \|\mathbf{c}_2\|_2$$

Em resumo, se A é uma matriz $m \times n$ de posto máximo, o problema de mínimos quadráticos pode ser resolvido da seguinte maneira:

1. Use transformações de Householder para calcular

$$R = H_n \cdots H_2 H_1 A \qquad \text{e} \qquad \mathbf{c} = H_n \cdots H_2 H_1 \mathbf{b}$$

2. Coloque R e $\mathbf{c}$ em blocos:

$$R = \begin{pmatrix} R_1 \\ O \end{pmatrix} \qquad \mathbf{c} = \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{pmatrix}$$

onde R_1 e $\mathbf{c}_1$ têm n linhas cada.

3. Use substituição de baixo para cima para resolver $R_1\mathbf{x} = \mathbf{c}_1$.

A PSEUDO-INVERSA

Vamos considerar agora o caso em que a matriz A tem posto $r < n$. A decomposição em valores singulares nos dá a chave para resolver o problema de mínimos quadráticos nesse caso. Ela pode ser usada para se construir uma inversa generalizada de A. No caso em que A é uma matriz invertível $n \times n$ com decomposição em valores singulares $U\Sigma V^T$, a inversa é dada por

$$A^{-1} = V\Sigma^{-1}U^T$$

Mais geralmente, se $A = U\Sigma V^T$ é uma matriz $m \times n$ de posto r, a matriz Σ vai ser uma matriz $m \times n$ da forma

$$\Sigma = \left(\begin{array}{c|c} \Sigma_1 & O \\ \hline O & O \end{array}\right) = \left(\begin{array}{cccc|c} \sigma_1 & & & & \\ & \sigma_2 & & & O \\ & & \ddots & & \\ & & & \sigma_r & \\ \hline & & O & & O \end{array}\right)$$

e podemos definir

$$A^+ = V\Sigma^+U^T \tag{1}$$

onde Σ^+ é a matriz $n \times n$

$$\Sigma^+ = \left(\begin{array}{c|c} \Sigma_1^{-1} & O \\ \hline O & O \end{array}\right) = \left(\begin{array}{ccc|c} \frac{1}{\sigma_1} & & & \\ & \ddots & & O \\ & & \frac{1}{\sigma_r} & \\ \hline & O & & O \end{array}\right)$$

A equação (1) nos fornece uma generalização natural da inversa de uma matriz. A matriz A^+ definida por (1) é chamada de *pseudo-inversa* de A.

É, também, possível definir A^+ por suas propriedades algébricas. Essas propriedades são dadas nas quatro condições a seguir.

As Condições de Penrose

1. $AXA = A$
2. $XAX = X$
3. $(AX)^T = AX$
4. $(XA)^T = XA$

Se A é uma matriz $m \times n$, afirmamos que existe uma única matriz X $n \times m$ que satisfaz essas condições. De fato, escolhendo $X = A^+ = V\Sigma^+U^T$ é fácil verificar que X satisfaz todas as quatro condições. Deixamos a demonstração disso a cargo do leitor. Para provar a unicidade, suponha que Y também satisfaz as condições de Penrose. Aplicando sucessivamente essas condições, podemos argumentar da seguinte maneira:

$$\begin{aligned}
X &= XAX & &(2) & \qquad Y &= YAY & &(2)\\
&= A^TX^TX & &(4) & &= YY^TA^T & &(3)\\
&= (AYA)^TX^TX & &(1) & &= YY^T(AXA)^T & &(1)\\
&= (A^TY^T)(A^TX^T)X & & & &= Y(Y^TA^T)(X^TA^T) & &\\
&= YAXAX & &(4) & &= YAYAX & &(3)\\
&= YAX & &(1) & &= YAX & &(1)
\end{aligned}$$

Logo, $X = Y$ e, portanto, A^+ é a única matriz satisfazendo as quatro condições de Penrose. Essas condições são usadas, muitas vezes, para se definir a pseudo-inversa e A^+ também é conhecida como a *pseudo-inversa de Moore-Penrose*.

Para ver como a pseudo-inversa pode ser usada na resolução de problemas de mínimos quadráticos, vamos considerar, primeiro, o caso em que A é uma matriz $m \times n$ de posto n. Nesse caso, Σ é da forma

$$\Sigma = \begin{pmatrix} \Sigma_1 \\ O \end{pmatrix}$$

onde Σ_1 é uma matriz $n \times n$ diagonal invertível. A matriz A^T
A é invertível e

$$(A^TA)^{-1} = V(\Sigma^T\Sigma)^{-1}V^T$$

A solução das equações normais é dada por

$$
\begin{aligned}
\mathbf{x} &= (A^TA)^{-1}A^T\mathbf{b} \\
&= V(\Sigma^T\Sigma)^{-1}V^TV\Sigma^TU^T\mathbf{b} \\
&= V(\Sigma^T\Sigma)^{-1}\Sigma^TU^T\mathbf{b} \\
&= V\Sigma^+U^T\mathbf{b} \\
&= A^+\mathbf{b}
\end{aligned}
$$

Logo, se A tem posto máximo, $A^+\mathbf{b}$ é a solução do problema de mínimos quadráticos. E no caso em que o posto de A é $r < n$? Nesse caso existem infinitas soluções para o problema de mínimos quadráticos. O próximo teorema mostra que não só A^+ $\mathbf{b}$ é uma solução, mas ela é mínima em relação à norma 2.

Teorema 7.8.1. *Se A é uma matriz $m \times n$ de posto $r < n$ com decomposição em valores singulares $U\Sigma V^T$, então o vetor*

$$\mathbf{x} = A^+\mathbf{b} = V\Sigma^+U^T\mathbf{b}$$

minimiza $\| \mathbf{b} - A\mathbf{x} \|_2^2$. Além disso, se $\mathbf{z}$ é outro vetor qualquer que minimiza $\| \mathbf{b} - A\mathbf{x} \|_2^2$, então $\| \mathbf{z} \|_2 > \| \mathbf{x} \|_2$.

Demonstração. Seja $\mathbf{x}$ um vetor em R^n e defina

$$\mathbf{c} = U^T\mathbf{b} = \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{pmatrix} \qquad \text{e} \qquad \mathbf{y} = V^T\mathbf{x} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}$$

onde $\mathbf{c}_1$ e $\mathbf{y}_1$ são vetores em R^r. Como U^T é ortogonal, temos

$$
\begin{aligned}
\|\mathbf{b} - A\mathbf{x}\|_2^2 &= \|U^T\mathbf{b} - \Sigma(V^T\mathbf{x})\|_2^2 \\
&= \|\mathbf{c} - \Sigma\mathbf{y}\|_2^2 \\
&= \left\| \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{pmatrix} - \begin{pmatrix} \Sigma_1 & O \\ O & O \end{pmatrix} \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} \right\|_2^2 \\
&= \left\| \begin{pmatrix} \mathbf{c}_1 - \Sigma_1\mathbf{y}_1 \\ \mathbf{c}_2 \end{pmatrix} \right\|_2^2 \\
&= \|\mathbf{c}_1 - \Sigma_1\mathbf{y}_1\|_2^2 + \|\mathbf{c}_2\|_2^2
\end{aligned}
$$

Como $\mathbf{c}_2$ é independente de $\mathbf{x}$, segue que $\| \mathbf{b} - A\mathbf{x} \|^2$ vai ser mínima se e somente se

$$\|\mathbf{c}_1 - \Sigma_1\mathbf{y}_1\| = 0$$

Logo, $\mathbf{x}$ é uma solução do problema de mínimos quadráticos se e somente se $\mathbf{x} = V\mathbf{y}$, onde $\mathbf{y}$ é um vetor da forma

$$\begin{pmatrix} \Sigma_1^{-1}\mathbf{c}_1 \\ \mathbf{y}_2 \end{pmatrix}$$

Em particular,

$$
\begin{aligned}
\mathbf{x} &= V\begin{pmatrix} \Sigma_1^{-1}\mathbf{c}_1 \\ \mathbf{0} \end{pmatrix} \\
&= V\begin{pmatrix} \Sigma_1^{-1} & O \\ O & O \end{pmatrix}\begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{pmatrix} \\
&= V\Sigma^+U^T\mathbf{b} \\
&= A^+\mathbf{b}
\end{aligned}
$$

é uma solução. Se $\mathbf{z}$ for outra solução, $\mathbf{z}$ tem que ser da forma

$$V\mathbf{y} = V\begin{pmatrix} \Sigma_1^{-1}\mathbf{c}_1 \\ \mathbf{y}_2 \end{pmatrix}$$

onde $\mathbf{y}_2 \neq \mathbf{0}$. Temos, então, que

$$\|\mathbf{z}\|^2 = \|\mathbf{y}\|^2 = \|\Sigma_1^{-1}\mathbf{c}_1\|^2 + \|\mathbf{y}_2\|^2 > \|\Sigma_1^{-1}\mathbf{c}_1\|^2 = \|\mathbf{x}\|^2 \qquad \square$$

Se a decomposição em valores singulares de A, $U\Sigma V^T$, for conhecida, é fácil calcular a solução do problema de mínimos quadráticos. Se $U = (\mathbf{u}_1, \ldots, \mathbf{u}_m)$ e $V = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$, definindo $\mathbf{y} = \Sigma^+ U^T\mathbf{b}$, temos

$$y_i = \frac{1}{\sigma_i}\mathbf{u}_i^T\mathbf{b} \qquad i = 1, \ldots, r \qquad (r = \text{posto de } A)$$

$$y_i = 0 \qquad i = r+1, \ldots, n$$

e, portanto,

$$A^+\mathbf{b} = V\mathbf{y} = \begin{pmatrix} v_{11}y_1 + v_{12}y_2 + \cdots + v_{1r}y_r \\ v_{21}y_1 + v_{22}y_2 + \cdots + v_{2r}y_r \\ \vdots \\ v_{n1}y_1 + v_{n2}y_2 + \cdots + v_{nr}y_r \end{pmatrix}$$
$$= y_1\mathbf{v}_1 + y_2\mathbf{v}_2 + \cdots + y_r\mathbf{v}_r$$

Logo, a solução $\mathbf{x} = A^+\mathbf{b}$ pode ser calculada em duas etapas:

1. Defina $y_i = (1/\sigma_i)\mathbf{u}_i^T\mathbf{b}$ para $i = 1, \ldots, r$;
2. Seja $\mathbf{x} = y_1\mathbf{v}_1 + \ldots + y_r\mathbf{v}_r$.

Vamos concluir esta seção esboçando um método para calcular os valores singulares de uma matriz. Em primeiro lugar observe que, se A tem decomposição em valores singulares $U\Sigma V^T$ e $B = HAP^T$, onde H é uma matriz ortogonal $m \times m$ e P é uma matriz ortogonal $n \times n$, então B tem decomposição em valores singulares $(HU)\Sigma(PV)^T$. Logo, o problema de encontrar os valores singulares de A pode ser simplificado aplicando-se transformações ortogonais a A para se obter uma matriz mais simples B com os mesmos valores singulares. Golub e Kahan mostraram que A pode ser colocada em uma forma bidiagonal superior usando transformações de Householder.

BIDIAGONALIZAÇÃO

Seja H_1 uma transformação de Householder que anula todos os elementos abaixo da diagonal na primeira coluna de A. Seja P_1 uma transformação de Householder tal que a multiplicação de H_1A por P_1 à direita anula os últimos $n-2$ elementos da primeira linha de H_1A, deixando a primeira coluna sem modificações.

$$H_1AP_1 = \begin{pmatrix} \times & \times & 0 & \cdots & 0 \\ 0 & \times & \times & \cdots & \times \\ \\ \vdots \\ \\ 0 & \times & \times & \cdots & \times \end{pmatrix}$$

O próximo passo é aplicar uma transformação de Householder H_2 que anula os elementos abaixo da diagonal da segunda coluna de H_1AP_1 sem mudar a primeira linha e a primeira coluna.

$$H_2H_1AP_1 = \begin{pmatrix} \times & \times & 0 & \cdots & 0 \\ 0 & \times & \times & \cdots & \times \\ 0 & 0 & \times & \cdots & \times \\ \vdots & & & & \\ 0 & 0 & \times & \cdots & \times \end{pmatrix}$$

A seguir, multiplicamos $H_2H_1AP_1$ à direita por uma transformação de Householder que anula os $n-3$ últimos elementos na segunda linha sem mudar as duas primeiras colunas e a primeira linha.

$$H_2H_1AP_1P_2 = \begin{pmatrix} \times & \times & 0 & 0 & \cdots & 0 \\ 0 & \times & \times & 0 & \cdots & 0 \\ 0 & 0 & \times & \times & \cdots & \times \\ \vdots & & & & & \\ 0 & 0 & \times & \times & \cdots & \times \end{pmatrix}$$

Continuamos dessa maneira até obter uma matriz

$$B = H_n \cdots H_1AP_1 \cdots P_{n-2}$$

da forma

$$\begin{pmatrix} \times & \times & & & \\ & \times & \times & & \\ & & \ddots & \ddots & \\ & & & \times & \times \\ & & & & \times \end{pmatrix}$$

Como $H = H_n \dots H_1$ e $P^T = P_1 \dots P_{n-2}$ são ortogonais, segue que B tem os mesmos valores singulares de A.

Simplificamos então o problema: basta encontrar os valores singulares de uma matriz bidiagonal superior B. Poderíamos agora formar a matriz tridiagonal simétrica B^TB e calcular seus autovalores usando o algoritmo QR. O problema com essa abordagem é que, ao formar B^TB, estaríamos elevando ao quadrado o número condicional e, portanto, nossa solução calculada seria muito menos confiável. O método que indicamos a seguir fornece uma seqüência de matrizes bidiagonais B_1, B_2, ... que converge para uma matriz diagonal Σ. O método envolve a aplicação de uma seqüência de transformações de Givens à direita e à esquerda de B, alternadamente.

O ALGORITMO DE GOLUB-REINSCH

Sejam

$$R_k = \begin{pmatrix} I_{k-1} & O & O \\ O & G(\theta_k) & O \\ O & O & I_{n-k-1} \end{pmatrix}$$

e

$$L_k = \begin{pmatrix} I_{k-1} & O & O \\ O & G(\varphi_k) & O \\ O & O & I_{n-k-1} \end{pmatrix}$$

onde as matrizes 2×2 $G(\theta_k)$ e $G(\varphi_k)$ são dadas por

$$G(\theta_k) = \begin{pmatrix} \cos\theta_k & \operatorname{sen}\theta_k \\ \operatorname{sen}\theta_k & -\cos\theta_k \end{pmatrix} \quad \text{e} \quad G(\varphi_k) = \begin{pmatrix} \cos\varphi_k & \operatorname{sen}\varphi_k \\ \operatorname{sen}\varphi_k & -\cos\varphi_k \end{pmatrix}$$

para ângulos θ_k e φ_k. A matriz $B = B_1$ é multiplicada primeiro à direita por R_1. Isso tem o efeito de colocar um elemento não-nulo na posição (2, 1).

$$B_1R_1 = \begin{pmatrix} \times & \times & & & \\ \times & \times & \times & & \\ & & \times & & \\ & & & \ddots & \times \\ & & & & \times \end{pmatrix}$$

A seguir, L_1 é escolhida de modo a anular o elemento não-nulo colocado por R_1. Essa transformação tem, também, o efeito de colocar um elemento não-nulo na posição (1, 3). Logo,

$$L_1B_1R_1 = \begin{pmatrix} \times & \times & \times & & \\ & \times & \times & & \\ & & & \ddots & \\ & & & & \times \\ & & & & \times \end{pmatrix}$$

R_2 é escolhida de modo a anular o elemento na posição (1, 3). Ele vai colocar um elemento não-nulo na posição (3, 2) da matriz $L_1B_1R_1$. A seguir, L_2 anula o elemento na posição (3, 2) e coloca outro na posição (2, 4), e assim por diante.

$$\underset{L_1B_1R_1R_2}{\begin{pmatrix} \times & \times & & & & \\ & \times & \times & & & \\ & \times & \times & \times & & \\ & & & & \ddots & \\ & & & & & \times \\ & & & & & \times \end{pmatrix}} \quad \underset{L_2L_1B_1R_1R_2}{\begin{pmatrix} \times & \times & & & & \\ & \times & \times & \times & & \\ & & \times & \times & & \\ & & & & \ddots & \\ & & & & & \times \\ & & & & & \times \end{pmatrix}}$$

Continuamos esse processo até obter uma nova matriz bidiagonal

$$B_2 = L_{n-1} \cdots L_1B_1R_1 \cdots R_{n-1}$$

Por que B_2 é melhor do que B_1? Pode-se mostrar que, se a primeira transformação, R_1, for escolhida corretamente, $B_2^TB_2$ será igual à matriz obtida de $B_1^TB_1$ aplicando-se uma iteração do algoritmo QR com deslocamento. O mesmo processo pode ser aplicado agora a B_2 para se obter uma matriz B_3 tal que $B_3^TB_3$ seja igual à matriz obtida a partir de $B_1^TB_1$ aplicando-se duas iterações do algoritmo QR com deslocamento. Embora as matrizes $B_i^TB_i$ nunca sejam calculadas, sabemos que, com uma escolha apropriada de deslocamentos, essas matrizes irão convergir rapidamente para uma matriz diagonal. Então, as matrizes B_i têm que convergir, também, para uma matriz diagonal Σ. Como cada um dos B_i tem os mesmos valores singulares que B, os elementos diagonais de Σ vão ser os valores singulares de B. As matrizes U e V^T podem ser determinadas se guardarmos todas as transformações ortogonais.

Demos apenas um rápido esboço do algoritmo. Incluir mais detalhes ficaria aquém do escopo deste livro. Para uma descrição completa do algoritmo e um programa em ALGOL, o leitor deve consultar o artigo de Golub e Reinsch [28].

EXERCÍCIOS

1. Encontre a solução **x** do problema de mínimos quadráticos em cada um dos itens a seguir, sabendo que $A = QR$.

(a) $Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & 0 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

(b) $Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & 0 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 3 \\ 1 \\ 2 \end{pmatrix}$

(c) $Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ \sqrt{2} \\ -\sqrt{2} \end{pmatrix}$

(d) $Q = \begin{pmatrix} \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{2} \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 2 \\ -2 \\ 0 \\ 2 \end{pmatrix}$

2. Sejam

$$A = \begin{pmatrix} D \\ E \end{pmatrix} = \left(\begin{array}{cccc} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \\ \hline e_1 & & & \\ & e_2 & & \\ & & \ddots & \\ & & & e_n \end{array}\right) \quad \text{e} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{2n} \end{pmatrix}$$

Use as equações normais para encontrar a solução **x** do problema de mínimos quadráticos.

3. Sejam

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 2 \\ 1 & -1 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} -3 \\ 10 \\ 3 \\ 6 \end{pmatrix}$$

(a) Use transformações de Householder para colocar A na forma

$$\begin{pmatrix} R_1 \\ O \end{pmatrix} = \begin{pmatrix} \times & \times \\ 0 & \times \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

e aplique as mesmas transformações a **b**.

(b) Use os resultados do item (a) para encontrar a solução de mínimos quadráticos para $A\mathbf{x} = \mathbf{b}$.

4. Seja

$$A = \begin{pmatrix} 1 & 1 \\ \epsilon & 0 \\ 0 & \epsilon \end{pmatrix}$$

onde ε é um escalar pequeno.

(a) Determine os valores singulares de A exatamente.

(b) Suponha que ε é suficientemente pequeno para que $1 + \varepsilon^2$ seja arredondado para 1 em sua calculadora. Determine os autovalores da matriz calculada A^TA e compare as raízes quadradas desses autovalores com suas respostas no item (a).

5. Mostre que a pseudo-inversa A^+ satisfaz as quatro condições de Penrose.

6. Seja B uma matriz qualquer satisfazendo as condições de Penrose (1) e (3) e seja $\mathbf{x} = B\mathbf{b}$. Mostre que **x** é uma solução das equações normais $A^TA\mathbf{x} = A^T\mathbf{b}$.

7. Se $\mathbf{x} \in R^m$, podemos pensar em **x** como sendo uma matriz $m \times 1$. Defina

$$X = \frac{1}{\|\mathbf{x}\|_2^2}\mathbf{x}^T$$

onde X é uma matriz $1 \times m$. Mostre que X e **x** satisfazem as quatro condições de Penrose e, portanto, que

$$\mathbf{x}^+ = X = \frac{1}{\|\mathbf{x}\|_2^2}\mathbf{x}^T$$

8. Seja A uma matriz $m \times n$ e seja $\mathbf{b} \in R^m$. Mostre que $\mathbf{b} \in I(A)$ se e somente se

$$\mathbf{b} = AA^+\mathbf{b}$$

9. Seja A uma matriz $m \times n$ com decomposição em valores singulares $U\Sigma V^T$ e suponha que A tem posto $r < n$. Seja $\mathbf{b} \in R^m$. Mostre que um vetor $\mathbf{x} \in R^n$ minimiza $\|\mathbf{b} - A\mathbf{x}\|_2$ se e somente se

$$\mathbf{x} = A^+\mathbf{b} + c_{r+1}\mathbf{v}_{r+1} + \cdots + c_n\mathbf{v}_n$$

onde $c_{r+1}, \ldots, c_n$ são escalares.

10. Seja

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{pmatrix}$$

Determine A^+ e verifique que A e A^+ satisfazem as quatro condições de Penrose (ver Exemplo 1 da Seção 6).

11. Sejam

$$A = \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix} \quad \text{e} \quad \mathbf{b} = \begin{pmatrix} 6 \\ -4 \end{pmatrix}$$

(a) Calcule a decomposição em valores singulares de A e use-a para encontrar A^+.
(b) Use A^+ para encontrar uma solução de mínimos quadráticos para o sistema $A\mathbf{x} = \mathbf{b}$.
(c) Encontre todas as soluções do problema de mínimos quadráticos para $A\mathbf{x} = \mathbf{b}$.

12. Mostre cada um das igualdades a seguir.

(a) $(A^+)^+ = A$
(b) $(AA^+)^2 = AA^+$
(c) $(A^+A)^2 = A^+A$

13. Sejam $A_1 = U\Sigma_1V^T$ e $A_2 = U\Sigma_2V^T$, onde

$$\Sigma_1 = \begin{pmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_{r-1} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

e

$$\Sigma_2 = \begin{pmatrix} \sigma_1 & & & & & & \\ & \ddots & & & & & \\ & & \sigma_{r-1} & & & & \\ & & & \sigma_r & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}$$

e $\sigma_r = \varepsilon > 0$. Quais os valores de $\| A_1 - A_2 \|_F$ e de $\| A_1^+ - A_2^+ \|_F$? O que acontece com esses valores quando $\varepsilon \to 0$?

14. Seja $A = XY^T$, onde X é uma matriz $m \times r$, Y^T é uma matriz $r \times n$ e X^TX e Y^TY são, ambas, invertíveis. Mostre que a matriz

$$B = \left[Y(Y^TY)^{-1}(X^TX)^{-1}X^T\right]$$

satisfaz as condições de Penrose e, portanto, tem que ser igual a A^+. Logo, A^+ pode ser determinada por uma fatoração dessa forma.

EXERCÍCIOS COM O MATLAB PARA O CAPÍTULO 7

SENSIBILIDADE DE SISTEMAS LINEARES

Nesses exercícios estaremos preocupados com a solução numérica de sistemas de equações lineares. Os elementos da matriz de coeficientes A e da matriz à direita do sinal de igualdade, **b**, podem, com freqüência, conter erros devido a limitações na precisão dos dados. Mesmo que não existam erros em A nem em **b**, erros de arredondamento irão ocorrer quando os elementos forem convertidos para o sistema numérico de precisão finita do computador. Então, em geral, esperamos que a matriz de coeficientes e os dados à direita do sinal de igualdade contenham erros pequenos. Logo, o sistema que o computador resolve é uma versão ligeiramente perturbada do sistema original. Se o sistema original for muito sensível, sua solução pode ser muito diferente da solução do sistema perturbado.

Em geral, um problema é bem-condicionado se perturbações na solução são da mesma ordem de grandeza que perturbações nos dados. Um problema é malcondicionado se as mudanças nas soluções são muito maiores do que as variações nos dados. Quão bem ou malcondicionado um problema é depende da comparação entre os tamanhos das perturbações na solução e nos dados. Para sistemas lineares, isso depende de quão perto a matriz de coeficientes está de uma matriz de posto menor. O condicionamento de um sistema pode ser medido através do número condicional da matriz. Isso pode ser calculado usando-se a função `cond` do MATLAB. Os cálculos no MATLAB são feitos com 16 dígitos significativos de precisão. Você vai perder dígitos de precisão dependendo da sensibilidade de seu sistema. Quanto maior o número condicional, mais dígitos de precisão serão perdidos.

1. Defina

$$\begin{aligned} A &= \texttt{round}(10 * \texttt{rand}(6)) \\ \mathbf{s} &= \texttt{ones}(6, 1) \\ \mathbf{b} &= A * \mathbf{s} \end{aligned}$$

É claro que a solução do sistema linear $A\mathbf{x} = \mathbf{b}$ é **s**. Resolva o sistema usando o operador \ do MATLAB. Calcule o erro $\mathbf{x} - \mathbf{s}$ (como **s** tem todos os elementos iguais a 1, isso é igual a $\mathbf{x} - \mathbf{1}$). Agora perturbe o sistema ligeiramente, fazendo

$$t = 1.0\text{e}{-12}, \qquad E = \texttt{rand}(6) - 0.5, \qquad \mathbf{r} = \texttt{rand}(6, 1) - 0.5$$

e definindo

$$M = A + t * E, \qquad \mathbf{c} = \mathbf{b} + t * \mathbf{r}$$

Resolva o sistema perturbado $M\mathbf{z} = \mathbf{c}$ para **z**. Compare a solução **z** à solução do sistema original calculando $\mathbf{z} - \mathbf{1}$. Qual a relação entre os tamanhos das perturbações na solução e em A e **b**? Repita a análise de perturbação com $t = 1.0\text{e}{-04}$ e $t = 1.0\text{e}{-02}$. O sistema $A = \mathbf{b}$ é bem-condicionado? Explique. Use o MATLAB para calcular o número condicional de A.

2. Se um vetor $\mathbf{y} \in R^n$ for usado para se construir uma matriz de Vandermonde $V\, n \times n$, então V será invertível desde que $y_1, y_2, \ldots, y_n$ sejam todos distintos.

(a) Construa um sistema de Vandermonde definindo

$$\mathbf{y} = \texttt{rand}(6, 1) \qquad \text{e} \qquad V = \texttt{vander}(\mathbf{y})$$

Gere vetores **b** e **s** em R^6 definindo

$$\mathbf{b} = \texttt{sum}(V')' \qquad \text{e} \qquad \mathbf{s} = \texttt{ones}(6, 1)$$

Se V e **b** tivessem sido calculados em aritmética exata, então a solução exata de $V\mathbf{x} = \mathbf{b}$ seria **s**. Por quê? Explique. Resolva $V\mathbf{x} = \mathbf{b}$ usando o operador \. Compare a solução calculada **x** com a solução exata **s** usando o comando `format long` do MATLAB. Quantos dígitos significativos foram perdidos? Determine o número condicional de V.

(b) As matrizes de Vandermonde tornam-se cada vez mais malcondicionadas à medida que n cresce. Mesmo para valores pequenos de n podemos tornar a matriz malcondicionada tomando dois de seus pontos próximos um do outro. Defina

$$x(2) = x(1) + 1{,}0\text{e}{-}12$$

e recalcule V usando o novo valor $x(2)$. Para a nova matriz V, defina $\mathbf{b} = \texttt{sum}(V')'$ e resolva o sistema $V\mathbf{z} = \mathbf{b}$. Quantos dígitos significativos foram perdidos? Determine o número condicional de V.

3. Construa uma matriz C definindo

$$\begin{aligned} A &= \texttt{round}(100 * \texttt{rand}(4)) \\ L &= \texttt{tril}(A, -1) + \texttt{eye}(4) \\ C &= L * L' \end{aligned}$$

(a) A matriz C é uma boa matriz, no sentido de que ela é simétrica com todos os elementos inteiros e tem determinante igual a 1. Use o MATLAB para verificar essas afirmações. Por que sabíamos imediatamente que o determinante tinha que ser igual a 1? Teoricamente, todos os elementos da inversa exata deveriam ser inteiros. Por quê? Explique. Isso acontece computacionalmente? Calcule $D = \texttt{inv}(C)$ e verifique seus elementos usando `format long`. Calcule $C * D$ e compare com $\texttt{eye}(4)$.

(b) Defina

$$\mathbf{r} = \texttt{ones}(4, 1) \qquad \text{e} \qquad \mathbf{b} = \texttt{sum}(C')'$$

Em aritmética exata a solução do sistema $C\mathbf{x} = \mathbf{b}$ deveria ser $\mathbf{r}$. Calcule a solução usando \ e apresente a resposta usando `format long`. Quantos dígitos significativos foram perdidos? Podemos perturbar ligeiramente o sistema escolhendo um escalar e pequeno, como 1.0e$-$12, e depois trocando a expressão à direita do sinal de igualdade no sistema por

$$\mathbf{b1} = \mathbf{b} + e * [1, -1, 1, -1]'$$

Resolva o sistema perturbado primeiro para o caso $e = 1.0\text{e}{-}12$ e depois para o caso $e = 1.0\text{e}{-}6$. Em cada caso, compare sua solução $\mathbf{x}$ com a solução original mostrando na tela $\mathbf{x} - \mathbf{1}$. Calcule $\texttt{cond}(C)$. C é malcondicionada? Explique.

4. A matriz de Hilbert $n \times n$ H é definida por

$$h(i, j) = 1/(i + j - 1) \qquad i, j = 1, 2, \ldots, n$$

Ela pode ser gerada usando-se o comando `hilb` do MATLAB. A matriz de Hilbert é sabidamente malcondicionada. Ela é usada com freqüência em exemplos para ilustrar os perigos dos cálculos matriciais. A função `invhilb` do MATLAB fornece a inversa exata da matriz de Hilbert. Para os casos $n = 6$, 8, 10 e 12, construa H e $\mathbf{b}$ de modo que $H\mathbf{x} = \mathbf{b}$ seja um sistema de Hilbert cuja solução em aritmética exata é $\texttt{ones}(n, 1)$.Para cada caso, determine a solução $\mathbf{x}$ do sistema usando `invhilb` e examine $\mathbf{x}$ em `format long`. Quantos dígitos significativos foram perdidos em cada caso? Determine o número condicional de cada uma das matrizes de Hilbert. Como o número condicional varia quando n cresce?

SENSIBILIDADE DE AUTOVALORES

Se A é uma matriz $n \times n$ e X é uma matriz que diagonaliza A, então a sensibilidade dos autovalores de A depende do número condicional de X. Se A não for diagonalizável, o número condicional para o problema de autovalores vai ser infinito. Para maiores informações sobre a sensibilidade de autovalores, ver o Cap. 2 do livro de Wilkinson [27].

5. Use o MATLAB para calcular os autovalores e autovetores de uma matriz aleatória B 6×6. Calcule o número condicional da matriz de autovetores. O problema de autovetores é bem-condicionado? Perturbe B ligeiramente, definindo

$$B1 = B + 1.0\text{e} - 04 * \texttt{rand}(6)$$

Calcule os autovalores e compare-os com o valor exato dos autovalores de B.

6. Defina

$$A = \texttt{round}(10 * \texttt{rand}(5)); \qquad A = A + A'$$
$$[X, D] = \texttt{eig}(A)$$

Calcule $\texttt{cond}(X)$ e X^TX. X é que tipo de matriz? O problema de autovalores é bem-condicionado? Explique. Perturbe A definindo

$$A1 = A + 1.0\text{e}{-06} * \texttt{rand}(5)$$

Calcule os autovalores de $A1$ e compare com os de A.

7. Defina $A = \texttt{magic}(4)$ e $t = \texttt{trace}(A)$. O escalar t deveria ser um autovalor de A e os autovalores restantes deveriam ter soma igual a zero. Por quê? Explique. Use o MATLAB para verificar que $A - tI$ é singular. Calcule os autovalores de A e uma matriz X de autovetores. O problema de autovalores é bem-condicionado? Explique. Perturbe A definindo

$$A1 = A + 1.0\text{e}{-04} * \texttt{rand}(4)$$

Qual a relação entre os autovalores de $A1$ e os de A?

8. Defina

$$A = \texttt{diag}(10:-1:1) + 10 * \texttt{diag}(\texttt{ones}(1, 9), 1)$$
$$[X, D] = \texttt{eig}(A)$$

Calcule o número condicional de X. O problema de autovalores é bem-condicionado? Malcondicionado? Explique. Perturbe A definindo

$$A1 = A; \qquad A1(10, 1) = 0{,}1$$

Calcule os autovalores de $A1$ e compare-os com os de A.

9. Construa uma matriz A da seguinte maneira:

$$A = \texttt{diag}(11:-1:1, -1);$$

for $j = 0 : 11$

$$A = A + \texttt{diag}(12 - j : -1 : 1, j);$$

end

(a) Calcule os autovalores e o determinante de A. Use a função `prod` do MATLAB para calcular o produto dos autovalores. Qual a relação entre esse produto e o determinante?

(b) Calcule os autovetores de A e o número condicional para o problema de autovalores. O problema é bem-condicionado? Malcondicionado? Explique.

(c) Defina

$$A1 = A + 1.0\text{e}{-04} * \texttt{rand}(A)$$

Calcule os autovalores de $A1$. Compare-os com os autovalores de A calculando

$$\texttt{sort}(\texttt{eig}(A1)) - \texttt{sort}(\texttt{eig}(A))$$

e apresentando o resultado usando `format long`.

TRANSFORMAÇÕES DE HOUSEHOLDER

Uma matriz de Householder é uma matriz ortogonal $n \times n$ da forma $I - (1/b)\mathbf{v}\mathbf{v}^T$ Dado qualquer vetor não-nulo $\mathbf{x} \in R^n$, é possível escolher b e $\mathbf{v}$ tais que $H\mathbf{x}$ seja múltiplo de $\mathbf{e}_1$.

10. (a) A maneira mais simples de calcular, no MATLAB, uma matriz de Householder que vai anular algumas coordenadas de um vetor **x** é calcular a fatoração QR de **x**. Logo, dado um vetor $\mathbf{x} \in R^n$, o comando

$$[H, R] = \texttt{qr}(\mathbf{x})$$

do MATLAB vai calcular a matriz de Householder H desejada. Calcule uma matriz de Householder que anula as três últimas coordenadas de $\mathbf{e} = \texttt{ones}(4, 1)$. Defina

$$C = [\mathbf{e}, \texttt{rand}(4, 3)]$$

Calcule $H * \mathbf{e}$ e $H * C$.

(b) Podemos, também, calcular o vetor **v** e o escalar b que determinam uma transformação de Householder que anula algumas coordenadas em um vetor. Para fazer isso para um vetor dado **x**, definimos

$$a = \texttt{norm}(\mathbf{x});$$

$$\mathbf{v} = \mathbf{x}; \qquad v(1) = v(1) - a$$

$$b = a * (a - x(1))$$

Construa **v** e b dessa maneira para o vetor **e** no item (a). Se $K = I - (\frac{1}{b})\mathbf{v}\mathbf{v}^T$, então

$$K\mathbf{e} = \mathbf{e} - \left(\frac{\mathbf{v}^T\mathbf{e}}{b}\right)\mathbf{v}$$

Calcule ambas as quantidades no MATLAB e verifique que são iguais. Qual a relação entre $K\mathbf{e}$ e $H\mathbf{e}$ encontrado no item (a)?
Calcule, também, $K{*}C$ e $C - \mathbf{v}{*}((\mathbf{v}'{*}C)/b)$ e verifique que são iguais.

11. Defina

$$\mathbf{x1} = (1:5)'; \qquad \mathbf{x2} = [1, 3, 4, 5, 9]'; \qquad \mathbf{x} = [\mathbf{x1}; \mathbf{x2}]$$

Construa uma matriz de Householder da forma

$$H = \begin{pmatrix} I & O \\ O & K \end{pmatrix}$$

onde K é uma matriz de Householder 5×5 que anula as quatro últimas coordenadas de **x2**. Calcule o produto $H\mathbf{x}$.

ROTAÇÕES E REFLEXÕES

12. Para fazer o gráfico de $y = \text{sen}(x)$, precisamos definir vetores com os valores de x e de y e depois usar o comando `plot`. Isso pode ser feito da seguinte maneira:

$$\mathbf{x} = 0:0.1:6.3; \qquad \mathbf{y} = \sin(\mathbf{x});$$

$$\texttt{plot}(\mathbf{x}, \mathbf{y})$$

(a) Vamos definir uma matriz de rotação e usá-la para rodar o gráfico de $y = \text{sen}(x)$. Defina

$$t = \texttt{pi}/4; \qquad c = \cos(t); \qquad s = \sin(t); \qquad R = [c, -s; s, c]$$

Para encontrar as coordenadas rodadas, sejam

$$Z = R * [\mathbf{x}; \mathbf{y}]; \qquad \mathbf{x1} = Z(1, :); \qquad \mathbf{y1} = Z(2, :);$$

Os vetores **x1** e **y1** contêm as coordenadas para a curva rodada. Defina

$$\mathbf{w} = [0, 5]; \qquad \texttt{axis('square')}$$

e faça o gráfico de **x1** e **y1** usando o comando

$$\texttt{plot}(\mathbf{x1}, \mathbf{y1}, \mathbf{w}, \mathbf{w})$$

do MATLAB. O gráfico foi rodado de que ângulo e em que sentido?

(b) Mantenha todas as suas variáveis do item (a) e defina

$$G = [c, s; s, -c]$$

A matriz G representa uma reflexão de Givens. Para determinar as coordenadas refletidas, defina

$$Z = G * [\mathbf{x}; \mathbf{y}]; \qquad \mathbf{x2} = Z(1, :); \qquad \mathbf{y2} = Z(2, :);$$

Faça o gráfico da curva refletida usando o comando MATLAB

$$\texttt{plot}(\mathbf{x2}, \mathbf{y2}, \mathbf{w}, \mathbf{w})$$

A curva $y = \text{sen}(x)$ foi refletida em torno de uma reta contendo a origem e fazendo um ângulo de $\pi/8$ com o eixo dos x. Para ver isso, defina

$$\mathbf{w1} = [0, 6{,}3 * \cos(t/2)]; \qquad \mathbf{z1} = [0, 6{,}3 * \text{sen}(t/2)];$$

e faça o gráfico dos novos eixos e de ambas as curvas usando o comando

$$\texttt{plot}(\mathbf{x}, \mathbf{y}, \mathbf{x2}, \mathbf{y2}, \mathbf{w1}, \mathbf{z1})$$

(c) Use a matriz de rotação do item (a) para rodar a curva $y = -\text{sen}(x)$. Faça o gráfico da curva rodada. Qual a relação entre esse gráfico e a curva no item (b)? Explique.

DECOMPOSIÇÃO EM VALORES SINGULARES

13. Seja

$$A = \begin{pmatrix} 4 & 5 & 2 \\ 4 & 5 & 2 \\ 0 & 3 & 6 \\ 0 & 3 & 6 \end{pmatrix}$$

Coloque no MATLAB a matriz A e calcule seus valores singulares usando $\mathbf{s} = \texttt{svd}(A)$.

(a) Como podemos usar as coordenadas de **s** para calcular os valores $\| A \|_2$ e $\| A \|_F$? Calcule essas normas definindo $p = \texttt{norm}(A)$ e $q = \texttt{norm}(A, \texttt{'fro'})$ e compare seus resultados com $s(1)$ e $\texttt{norm}(\mathbf{s})$.

(b) Para obter a decomposição em valores singulares completa de A, seja

$$[U, D, V] = \texttt{svd}(A)$$

Calcule a matriz de posto 1 mais próxima de A digitando

$$B = s(1) * U(:, 1) * V(:, 1)'$$

Qual a relação entre as linhas de B e as duas linhas distintas de A?

(c) As matrizes A e B deveriam ter a mesma norma 2. Por quê? Explique. Use o MATLAB para calcular $\| B \|_2$ e $\| B \|_F$. Em geral, para uma matriz de posto 1, a norma 2 e a norma de Frobenius. Por quê? Explique.

14. Defina

$$A = \texttt{round}(10 * \texttt{rand}(10, 5)) \qquad \text{e} \qquad \mathbf{s} = \texttt{svd}(A)$$

(a) Use o MATLAB para calcular $\| A \|_2$, $\| A \|_F$, $\text{cond}_2(A)$ e compare seus resultados com $s(1)$, norm(**s**), $s(1)/s(5)$, respectivamente.
(b) Defina

$$[U,\ D,\ V] = \texttt{svd}(A); \qquad D(5,5) = 0; \qquad B = U * D * V'$$

A matriz B deveria ser a matriz de posto 4 mais próxima de A (onde a distância é medida em termos da norma de Frobenius). Calcule $\| A \|_2$ e $\| B \|_2$. Qual a relação entre esses dois valores? Calcule e compare a norma de Frobenius para as duas matrizes. Calcule, também, $\| A - B \|_F$ e compare o resultado com $s(5)$. Defina $r = \texttt{norm}(s(1:4))$ e compare o resultado com $\| B \|_F$.
(c) Use o MATLAB para construir a matriz C de posto 3 mais próxima de A em relação à norma de Frobenius. Calcule $\| C \|_2$ e $\| C \|_F$. Qual a relação entre esses valores e os valores calculados para $\| A \|_2$ e $\| A \|_F$? Defina

$$p = \texttt{norm}(\mathbf{s}(1:3)) \qquad \text{e} \qquad q = \texttt{norm}(\mathbf{s}(4:5))$$

Calcule $\| C \|_F$ e $\| A - C \|_F$ e compare seus resultados com p e q.

15. Defina

$$A = \texttt{rand}(8,4) * \texttt{rand}(4,6), \qquad [U,\ D,\ V] = \texttt{svd}(A)$$

(a) Qual o posto de A? Use as colunas de V para gerar duas matrizes $V1$ e $V2$ cujas colunas formam bases ortonormais para $I(A^T)$ e $N(A)$, respectivamente. Defina

$$P = V2 * V2', \quad \mathbf{r} = P * \texttt{rand}(6,1), \quad \mathbf{w} = A' * \texttt{rand}(8,1)$$

Se **r** e **w** tivessem sido calculados em aritmética exata, eles deveriam ser ortogonais. Por quê? Explique. Use o MATLAB para calcular $\mathbf{r}^T\mathbf{w}$.
(b) Use as colunas de U para gerar duas matrizes $U1$ e $U2$ cujas colunas formam bases ortonormais para $I(A)$ e $N(A^T)$, respectivamente. Defina

$$Q = U2 * U2', \quad \mathbf{y} = Q * \texttt{rand}(8,1), \quad \mathbf{z} = A * \texttt{rand}(6,1)$$

Explique por que **y** e **z** deveriam ser ortogonais se todos os cálculos tivessem sido feitos em aritmética exata. Use o MATLAB para calcular $\mathbf{y}^T\mathbf{z}$.
(c) Defina $X = \texttt{pinv}(A)$. Use o MATLAB para verificar as quatro condições de Penrose:

(i) $AXA = A$ (ii) $XAX = X$
(iii) $(AX)^T = AX$ (iv) $(XA)^T = XA$

(d) Calcule e compare AX e $U1(U1)^T$. Se todos os cálculos tivessem sido feitos em aritmética exata, essas duas matrizes seriam iguais. Por quê? Explique.

CÍRCULOS DE GERSCHGORIN

16. Podemos associar a cada $A \in R^{n\times n}$ n discos fechados no plano. O i-ésimo disco está centrado em a_{ii} e tem raio $r_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|$. Cada autovalor de A está contido em pelo menos um dos discos (ver Exercício 7 na Seção 7 do Cap. 7).
(a) Defina

$$A = \texttt{round}(10 * \texttt{rand}(5))$$

Calcule os raios dos discos de Gerschgorin de A e guarde-os em um vetor **r**. Para fazer o gráfico dos discos, precisamos parametrizar os círculos. Isso pode ser feito definindo-se

$$t = [0:0,1:6,3]';$$

Podemos, então, gerar duas matrizes X e Y cujas colunas contêm as coordenadas x e y dos círculos. Primeiro anule os elementos de X e Y definindo

$$X = \texttt{zeros}(\texttt{length}(t), 5); \qquad Y = X;$$

As matrizes podem ser geradas, então, pelos seguintes comandos:

for $i = 1:5$

$X(:, i) = r(i) * \cos(t) + \texttt{real}(A(i, i));$

$Y(:, i) = r(i) * \sin(t) + \texttt{imag}(A(i, i));$

end

Faça $\mathbf{e} = \texttt{eig}(A)$ e faça o gráfico dos autovalores e dos discos com o comando

$$\texttt{plot}(X, Y, \texttt{real}(\mathbf{e}), \texttt{imag}(\mathbf{e}), \text{'}x\text{'})$$

Se tudo for feito corretamente, todos os autovalores de A devem estar contidos na união dos discos.

(b) Se k dos discos de Gerschgorin forma um domínio conexo no plano complexo isolado dos outros discos, então exatamente k dos autovalores da matriz estará nesse domínio. Defina

$$B = [3 \quad 0{,}1 \quad 2; \quad 0{,}1 \quad 7 \quad 2; \quad 2 \quad 2 \quad 50]$$

(i) Use o método descrito no item (a) para calcular e fazer o gráfico dos discos de Gerschgorin de B.

(ii) Como B é simétrica, seus autovalores são todos reais, logo pertencem ao eixo real. Sem calcular os autovalores, explique por que B tem que ter exatamente um autovalor no intervalo [46,54]. Multiplique as duas primeiras linhas de B por 0,1 e depois multiplique as duas primeiras colunas por 10. Isso pode ser feito no MATLAB com os comandos

$$D = \texttt{diag}([0.1, 0.1, 1]) \qquad \text{e} \qquad C = D * B/D$$

A nova matriz C deveria ter os mesmos autovalores que B. Por quê? Explique. Use C para encontrar intervalos que contêm os outros dois autovalores. Calcule e faça o gráfico dos discos de Gerschgorin de C.

DISTRIBUIÇÃO DE NÚMEROS CONDICIONAIS E AUTOVALORES DE MATRIZES ALEATÓRIAS

17. Podemos gerar uma matriz simétrica aleatória 10×10 digitando

$$A = \texttt{rand}(10); \; A = (A + A')/2$$

Como A é simétrica, seus autovalores são todos reais. O número de autovalores positivos pode ser calculado digitando-se

$$y = \texttt{sum}(\texttt{eig}(A) > 0)$$

(a) Para $j = 1, 2, \ldots, 100$, gere uma matriz simétrica aleatória 10×10 e determine o número de autovalores positivos. Denote o número de autovalores positivos da j-ésima matriz por $y(j)$. Defina $\mathbf{x} = 0:10$ e determine a distribuição dos dados $\mathbf{y}$ fazendo $\mathbf{n} = \texttt{hist}(\mathbf{y}, \mathbf{x})$. Determine a média dos valores $y(j)$ usando o comando mean(y) do MATLAB. Use o comando $\texttt{hist}(\mathbf{y}, \mathbf{x})$ para gerar um gráfico do histograma.

(b) Podemos gerar uma matriz simétrica aleatória 10×10 com elementos no intervalo $[-1, 1]$ digitando

$$A = 2 * \texttt{rand}(10) - 1; \qquad A = (A + A')/2$$

Repita o item (a) usando matrizes aleatórias geradas desse modo. Qual a relação entre a distribuição dos dados **y** obtidos dessa forma e os obtidos no item (a)?

18. Uma matriz não-simétrica A pode ter autovalores complexos. Podemos determinar o número de autovalores reais positivos usando os comandos

$$\mathbf{e} = \texttt{eig}(A)$$

$$y = \texttt{sum}(\mathbf{e} > 0 \ \& \ \texttt{imag}\,(\mathbf{e}) == 0)$$

Gere 100 matrizes aleatórias não-simétricas 10×10. Para cada uma delas, determine o número de autovalores positivos e guarde esse número como uma coordenada de um vetor **z**. Determine a média dos valores $z(j)$ e compare-a com a média calculada no item (a) do Exercício 17. Determine a distribuição e faça um gráfico do histograma.

19. (a) Gere 100 matrizes aleatórias 5×5 e calcule o número condicional de cada matriz. Determine a média dos números condicionais e faça o gráfico do histograma da distribuição.

(b) Repita o item (a) usando matrizes 10×10. Compare esses resultados com os obtidos no item (a).

APÊNDICE MATLAB

MATLAB é um programa interativo para cálculos matriciais. A versão original do MATLAB, que é uma abreviação de *matrix laboratory**, foi desenvolvida por Cleve Moler para as bibliotecas de programas Linpack e Eispack. Ao longo dos anos, o MATLAB passou por uma série de expansões e revisões. Hoje em dia é o líder entre os programas para cálculos científicos. A versão profissional do MATLAB é distribuída por Math Works, Inc., de Natick, Massachusetts.

Além do amplo uso nos contextos industrial e de engenharia, o MATLAB tornou-se, também, uma ferramenta-padrão para o ensino de álgebra linear em cursos de graduação. Para satisfazer a demanda crescente, foi lançada em 1992 uma edição mais barata para estudantes (*Student Edition*). Essa edição está disponível para computadores pessoais do tipo PC e para MacIntosh (ver [22]).

ELEMENTOS BÁSICOS DE DADOS

Os elementos básicos utilizados pelo MATLAB são matrizes. Uma vez colocadas ou geradas, o usuário pode fazer rapidamente cálculos sofisticados com uma quantidade mínima de programação.

Colocar matrizes no MATLAB é fácil. Para colocar a matriz

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix}$$

poderíamos digitar

$$A = [1\quad 2\quad 3\quad 4;\quad 5\quad 6\quad 7\quad 8;\quad 9\quad 10\quad 11\quad 12;\quad 13\quad 14\quad 15\quad 16]$$

ou poderíamos digitar uma linha de cada vez:

$$\begin{array}{rl} A = & [\,1\quad 2\quad 3\quad 4 \\ & \;\;5\quad 6\quad 7\quad 8 \\ & \;\;9\quad 10\quad 11\quad 12 \\ & 13\quad 14\quad 15\quad 16\,] \end{array}$$

* Laboratório de matrizes em português. (*N.T.*)

Vetores-linhas com coordenadas espaçadas igualmente podem ser gerados usando-se o operador : do MATLAB. O comando $\mathbf{x} = 2{:}6$ gera um vetor-linha com coordenadas inteiras variando de 2 a 6.

$$\mathbf{x} =$$

$$2 \quad 3 \quad 4 \quad 5 \quad 6$$

Não é necessário usar inteiros ou ter uma variação de 1. Por exemplo, o comando $\mathbf{x} = 1.2{:}0.2{:}2$ gera

$$\mathbf{x} =$$

$$1{,}2000 \quad 1{,}4000 \quad 1{,}6000 \quad 1{,}8000 \quad 2{,}0000$$

SUBMATRIZES

Para nos referirmos a uma submatriz de A, precisamos usar : para especificar as linhas e colunas. Por exemplo, a submatriz que consiste nos elementos das segunda e terceira linhas pertencentes às colunas de 2 a 4 é dada por $A(2{:}3, 2{:}4)$. Portanto, o comando

$$C = A(2:3, 2:4)$$

gera

$$C =$$

$$\begin{matrix} 6 & 7 & 8 \\ 10 & 11 & 12 \end{matrix}$$

Se os dois pontos são usados como um dos argumentos, todas as linhas ou todas as colunas da matriz serão incluídas. Por exemplo, $A(:, 2{:}3)$ representa a submatriz que contém todos os elementos nas segunda e terceira colunas de A e $A(4, :)$ denota a quarta linha de A.

GERANDO MATRIZES

Podemos, também, gerar matrizes usando funções próprias do MATLAB. Por exemplo, o comando

$$B = \texttt{rand}(4)$$

gera uma matriz 4×4 cujos elementos são números aleatórios entre 0 e 1. Outras funções que podem ser usadas para gerar matrizes são `eyes`, `zeros`, `ones`, `magic`, `hilb`, `pascal`, `toeplitz`, `compan` e `vander`. Para construir matrizes triangulares ou diagonais, podemos usar as funções `triu`, `tril` e `diag` do MATLAB.

Os comandos de construção de matrizes podem ser usados para gerar matrizes em blocos. Por exemplo, o comando

$$E = [\,\texttt{eye}(2),\ \texttt{ones}(2,3);\ \texttt{zeros}(2),\ [1:3;\quad 3:-1:1]$$

do MATLAB vai gerar a matriz

$$E =$$

$$\begin{matrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 3 & 2 & 1 \end{matrix}$$

ARITMÉTICA MATRICIAL

A aritmética matricial é direta no MATLAB. Podemos multiplicar nossa matriz original A por B digitando, simplesmente, $A{*}B$. A soma e a diferença de A e B são dadas por $A + B$ e $A - B$, respectivamente. A transposta de A é dada por A'. Se $\mathbf{c}$ representa um vetor em R^4, a solução do sistema linear $A\mathbf{x} = \mathbf{c}$ pode ser calculada por

$$\mathbf{x} = A\backslash\mathbf{c}$$

Potências de matrizes são geradas com facilidade. A matriz A^5 é calculada no MATLAB digitando-se A^5. Podemos, também, efetuar operações nos elementos colocando um ponto antes do operador. Por exemplo, se $W = [1\ \ 2;\ 3\ \ 4]$, então, digitando W^2, obtemos

```
ans =
      7   10
     15   22
```

enquanto, se digitarmos W.^2, vai aparecer

```
ans =
      1    4
      9   16
```

FUNÇÕES DO MATLAB

Para calcular os autovalores de uma matriz quadrada A, basta digitar `eig`(A). Os autovetores e autovalores podem ser obtidos por

$$[X \quad D] = \texttt{eig}(A)$$

Analogamente, podemos calcular o determinante, a inversa, o número condicional, a norma e o posto de uma matriz com comandos simples de uma palavra. Fatorações matriciais, como as fatorações de LU, QR, Cholesky, e as decomposições de Schur e em valores singulares podem ser calculadas com um único comando. Por exemplo, o comando

$$[Q \quad R] = \texttt{qr}(A)$$

produz uma matriz ortogonal (ou unitária) Q e uma matriz triangular superior R, com o mesmo tamanho de A, tais que $A = QR$.

CARACTERÍSTICAS DE PROGRAMAÇÃO

O MATLAB tem todas as estruturas de controle de fluxo que você esperaria de uma linguagem de alto nível, como ciclos de para (`for` loops), enquanto (`while` loops) e sentenças condicionais (com o comando `if`). Isso permite ao usuário escrever seus próprios programas e criar funções adicionais no MATLAB. Devemos observar que o MATLAB mostra na tela, automaticamente, o resultado de cada comando, a menos que a linha de comando termine com um ponto e vírgula. Ao usar ciclos fechados (*loops*), recomendamos terminar cada comando com um ponto e vírgula para evitar que apareçam na tela os resultados de todos os cálculos intermediários.

OPERADORES RELACIONAIS E LÓGICOS

O MATLAB tem seis operadores relacionais que são usados para comparação de escalares ou de elementos de vetores (*arrays*). Esses operadores são:

Operadores Relacionais	
$<$	menor do que
$<=$	menor ou igual a
$>$	maior do que
$>=$	maior ou igual a
$==$	igual a
$\sim=$	diferente de

Dadas duas matrizes $m \times n$ A e B, o comando

$$C = A < B$$

vai gerar uma matriz com todos os elementos iguais a zero ou um. O elemento (i, j) vai ser igual a 1 se e somente se $a_{ij} < b_{ij}$. Por exemplo, suponha que

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

O comando $A >= 0$ vai gerar

```
ans =
      0 1 1
      1 1 0
      0 0 1
```

Existem três operadores lógicos, como ilustrado na tabela a seguir.

Operadores Lógicos	
&	E
|	OU
~	NEGAÇÃO

Esses operadores lógicos consideram qualquer escalar diferente de zero como sendo VERDADEIRO e 0 corresponde a FALSO. O operador & corresponde ao operador lógico E. Se a e b são escalares, a expressão $a\&b$ vai ser igual a 1 se a e b forem, ambos, diferentes de zero (VERDADEIRO) e vai ser 0 em caso contrário. O operador | corresponde ao operador lógico OU. A expressão $a|b$ vai ter valor 0 se ambos, a e b, forem iguais a 0; caso contrário, vai ser igual a 1. O operador ~ corresponde ao operador lógico de NEGAÇÃO. Para um escalar a, ele assume o valor 1 (VERDADEIRO) se $a = 0$ (FALSO) e o valor 0 (FALSO) se $a \neq 0$ (VERDADEIRO).

Em matrizes, esses operadores são aplicados aos elementos. Se A e B são matrizes $m \times n$, então $A\&B$ é uma matriz contendo apenas elementos iguais a zero ou um; o elemento (i, j) é igual a $a(i, j)\&b(i, j)$. Por exemplo, se

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} -1 & 2 & 0 \\ 1 & 0 & 3 \\ 0 & 1 & 2 \end{pmatrix}$$

então

$$A\&B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad A|B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad \tilde{}A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$

Os operadores relacionais e lógicos são usados, freqüentemente, em sentenças condicionais.

OPERADORES QUE AGEM EM COLUNAS OU LINHAS

O MATLAB tem uma série de funções que, ao serem aplicadas a um vetor linha ou coluna **x**, dão como resposta um único número. Por exemplo, o comando `max`(**x**) dá o valor da maior coordenada de **x**, e o comando `sum`(**x**) fornece a soma de todas as coordenadas de **x**. Outras funções dessa forma são `min`, `prod`, `mean`, `all` e `any`.* Ao serem usados com um argumento matricial, esses operadores são aplicados a cada vetor-coluna e os resultados aparecem como um vetor-linha. Por exemplo, se

$$A = \begin{pmatrix} -3 & 2 & 5 & 4 \\ 1 & 3 & 8 & 0 \\ -6 & 3 & 1 & 3 \end{pmatrix}$$

* Mínimo, produto, média, todo e qualquer. (*N.T.*)

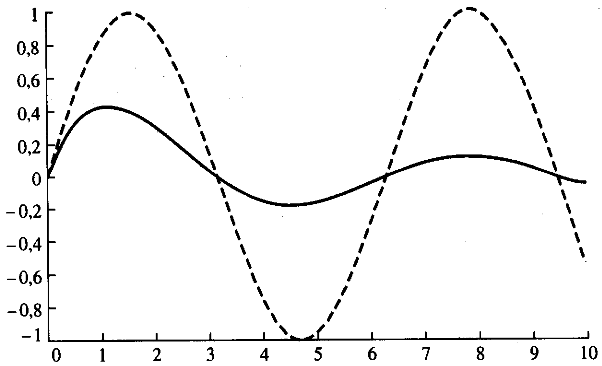


Fig. 1

então

$$\begin{aligned}\texttt{min}(A) &= (-6, 2, 1, 0)\\ \texttt{max}(A) &= (1, 3, 8, 4)\\ \texttt{sum}(A) &= (-8, 8, 14, 7)\\ \texttt{prod}(A) &= (18, 18, 40, 0)\end{aligned}$$

GRÁFICOS

Se **x** e **y** são vetores de mesmo comprimento, o comando `plot`(**x**, **y**) gera um gráfico de todos os pares (x_i, y_i) e cada ponto estará ligado ao próximo por um segmento de reta. Se as coordenadas x estão suficientemente próximas umas das outras, o gráfico deveria parecer uma curva suave. O comando `plot`(**x**, **y**, 'x') mostra o gráfico dos pares ordenados com os x, mas não conecta os pontos.

Por exemplo, para fazer o gráfico da função $f(x) = \text{sen}(x)/(x + 1)$ no intervalo [0, 10] defina

$$\mathbf{x} = 0{:}0.2{:}10 \quad \text{e} \quad \mathbf{y} = \text{sen}(\mathbf{x})./(\mathbf{x} + 1)$$

O comando `plot`(**x**, **y**) gera o gráfico da função. Para comparar esse gráfico ao de sen(x), definimos **z** = sin(**x**) e usamos o comando

$$\texttt{plot}(\mathbf{x}, \mathbf{y}, \mathbf{x}, \mathbf{z})$$

para desenhar ambas as curvas ao mesmo tempo, como na Fig. 1.

É possível também fazer gráficos mais sofisticados usando o MATLAB, incluindo coordenadas polares, gráficos tridimensionais e curvas de nível.

AJUDA

O MATLAB inclui um arquivo de ajuda (HELP) que lista e descreve todas as funções, operações e comandos do programa. Para obter informações sobre qualquer comando do MATLAB, basta digitar `help` seguido pelo nome do comando.

CONCLUSÕES

O MATLAB é uma ferramenta poderosa para cálculos matriciais que é, também, amigável. Os comandos básicos podem ser aprendidos facilmente e, por isso, os alunos são capazes de começar experimentos numéricos com um mínimo de preparação. De fato, o material neste apêndice, junto com a facilidade do arquivo de ajuda *on-line*, deve ser suficiente para se começar.

Os exercícios a serem feitos com o MATLAB ao final de cada capítulo são projetados para aumentar a compreensão de álgebra linear. Os exercícios não pressupõem familiaridade com o programa. Muitas vezes são dados comandos específicos para guiar o leitor nas construções mais complicadas. Conseqüentemente, deveria ser possível fazer todos os exercícios sem livros ou manuais adicionais do MATLAB.

Embora este apêndice resuma as características do MATLAB que são relevantes para uma disciplina de graduação de álgebra linear, existem muitos outros recursos mais avançados que não foram discutidos. Para maiores detalhes sobre esses recursos, o leitor deve consultar *The Student Edition of* MATLAB [22]. As referências [13] e [18] também descrevem o programa com maiores detalhes.

BIBLIOGRAFIA

ÁLGEBRA LINEAR

1. Brualdi, Richard A., and Herbert J. Ryser, *Combinatorial Matrix Theory.* New York: Cambridge University Press, 1991.
2. Gantmacher, F. R., *The Theory of Matrices,* 2 vols. New York: Chelsea Publishing Company, Inc., 1960.
3. Horn, Roger A., and Charles R. Johnson, *Matrix Analysis.* New York: Cambridge University Press, 1985.
4. Horn, Roger A., and Charles R. Johnson, *Topics in Matrix Analysis.* New York: Cambridge University Press, 1991.
5. Lancaster, Peter, and M. Tismenetsky, *The Theory of Matrices with Applications,* 2nd ed. New York: Academic Press, Inc., 1985.
6. Ortega, James M., *Matrix Theory: A Second Course.* New York: Plenum Press, 1987.
7. Pullman, N. J., *Matrix Theory and Its Applications: Selected Topics.* New York: Marcel Dekker, Inc., 1976.

ÁLGEBRA LINEAR APLICADA

8. Bellman, Richard, *Introduction to Matrix Analysis,* 2nd ed. New York: McGraw-Hill Book Company, 1970.
9. Fletcher, T. J., *Linear Algebra Through Its Applications.* New York: Van Nostrand Reinhold Company, 1972.
10. Noble, Ben, and James W. Daniel, *Applied Linear Algebra,* 3rd ed. Englewood Cliffs, N.J.: Prentice Hall, 1988.
11. Strang, Gilbert, *Linear Algebra and Its Applications,* 3rd ed. San Diego: Harcourt Brace Jovanovich, 1988.

ÁLGEBRA LINEAR NUMÉRICA

12. Anderson, E., Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorenson, *LAPACK Users' Guide.* Philadelphia: SIAM, 1992.

13. Coleman, Thomas F., and Charles Van Loan, *Handbook for Matrix Computations.* Philadelphia: SIAM, 1988.
14. Conte, S. D., and C. deBoor, *Elementary Numerical Analysis: An Algorithmic Approach,* 3rd ed. New York: McGraw-Hill Book Company, 1980.
15. Dahlquist, G., and A. Bjorck, *Numerical Methods.* Englewood Cliffs, N.J.: Prentice Hall, 1974.
16. Forsythe, G. E., and C. B. Moler, *Computer Solutions of Linear Algebraic Systems.* Englewood Cliffs, N.J.: Prentice Hall, 1967.
17. Golub, Gene H., and Charles F. Van Loan, *Matrix Computations,* 2nd ed. Baltimore: Johns Hopkins University Press, 1989.
18. Hill, David R., *Experiments in Computational Matrix Algebra.* New York: Random House, Inc., 1988.
19. Jepsen, Charles H., and Eugene A. Herman, *MAX: MAtriX Algebra Calculator: Linear Algebra Problems for Computer Solution.* Pacific Grove, Calif.: Brooks/Cole Publishing Company, 1988.
20. Kahaner, D., C. B. Moler, and S. Nash, *Numerical Methods and Software.* Englewood Cliffs, N.J.: Prentice Hall, 1988.
21. Lawson, Charles L., and Richard J. Hanson, *Solving Least Squares Problems.* Englewood Cliffs, N.J.: Prentice Hall, 1974.
22. Moler, Cleve, John Little, and Steve Bangert, *The Student Edition of MATLAB.* Englewood Cliffs, N.J.: Prentice Hall, 1992.
23. Parlett, B. N., *The Symmetric Eigenvalue Problem.* Englewood Cliffs, N.J.: Prentice Hall, 1980.
24. Stewart, G. W., *Introduction to Matrix Computations.* New York: Academic Press, Inc., 1973.
25. Watkins, David S., *Fundamentals of Matrix Computation.* New York: Wiley, 1991.
26. Wilkinson, J. H., *Rounding Errors in Algebraic Processes.* Englewood Cliffs, N.J.: Prentice Hall, 1963.
27. Wilkinson, J. H., *The Algebraic Eigenvalue Problem.* Oxford: Clarendon Press, 1965.
28. Wilkinson, J. H., and C. Reinsch, *Handbook for Automatic Computation,* Vol. II: *Linear Algebra.* New York: Springer-Verlag, 1971.

LIVROS DE INTERESSE

29. Cheney, L. W., *Introduction to Approximation Theory.* New York: McGraw-Hill Book Company, 1966.
30. Chiang, Alpha C., *Fundamental Methods of Mathematical Economics.* New York: McGraw-Hill Book Company, 1967.
31. Courant, R., and D. Hilbert, *Methods of Mathematical Physics,* Vol. I. New York: Wiley-Interscience, 1953.
32. Kreider, D. L., R. G. Kuller, and D. R. Ostberg, *Elementary Differential Equations.* Reading, Mass.: Addison-Wesley Publishing Company, Inc., 1968.
33. Rivlin, T. J., *The Chebyshev Polynomials.* New York: Wiley-Interscience, 1974.

As referências 13, 18 e 22 contêm informações sobre o MATLAB. As seguintes referências contêm bibliografias mais completas: 2, 3, 10, 15, 17, 21, 23 e 27.

Respostas de Exercícios Selecionados

CAPÍTULO 1

SEÇÃO 1

1. (a) $(11, 3)$; (b) $(4, 1, 3)$; (c) $(-2, 0, 3, 1)$; (d) $(-2, 3, 0, 3, 1)$

2. (a) $\begin{pmatrix} 1 & -3 \\ 0 & 2 \end{pmatrix}$; (b) $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$; (c) $\begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & 3 & 1 & -2 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 4 \end{pmatrix}$

3. (a) Uma solução. As duas retas se interceptam no ponto (3, 1).
(b) Não tem solução. As retas são paralelas.
(c) Uma infinidade de soluções. Ambas as equações representam a mesma reta.
(d) Não tem solução. Cada par de retas se intercepta em um ponto; no entanto, não existe nenhum ponto que pertença às três retas.

4. (a) $\left(\begin{array}{rr|r} 1 & 1 & 4 \\ 1 & -1 & 2 \end{array}\right)$; (b) $\left(\begin{array}{rr|r} 1 & 2 & 4 \\ -2 & -4 & 4 \end{array}\right)$;
(c) $\left(\begin{array}{rr|r} 2 & -1 & 3 \\ -4 & 2 & -6 \end{array}\right)$; (d) $\left(\begin{array}{rr|r} 1 & 1 & 1 \\ 1 & -1 & 1 \\ -1 & 3 & 3 \end{array}\right)$

6. (a) $(1, -2)$; (b) $(3, 2)$; (c) $(\frac{1}{2}, \frac{2}{3})$;
(d) $(1, 1, 2)$; (e) $(-3, 1, 2)$; (f) $(-1, 1, 1)$;
(g) $(1, 1, -1)$; (h) $(4, -3, 1, 2)$

7. (a) $(2, -1)$; (b) $(-2, 3)$

8. (a) $(-1, 2, 1)$; (b) $(3, 1, -2)$

SEÇÃO 2

1. Em forma escada: (a), (c), (d), (g) e (h); em forma escada reduzida por linhas: (c), (d) e (g).
2. (a) Incompatível; (b) compatível, $(4, -1)$; (c) compatível, uma infinidade de soluções; (d) compatível, $(4, 5, 2)$; (e) incompatível; (f) compatível, $(5, 3, 2)$.
3. (a) $(-2, 5, 3)$; (b) $\emptyset$; (c) $\{(2 + 3\alpha, \alpha, -2) \mid \alpha \text{ real}\}$;
(d) $\{(5 - 2\alpha - \beta, \alpha, 4 - 3\beta, \beta) \mid \alpha, \beta \text{ reais}\}$;
(e) $\{(3 - 5\alpha + 2\beta, \alpha, \beta, 6) \mid \alpha, \beta \text{ reais}\}$;
(f) $\{(\alpha, 2, -1) \mid \alpha \text{ real}\}$
4. (a) $(5, 1)$; (b) incompatível; (c) $(0, 0)$;
(d) $\left\{\left(\frac{5-\alpha}{4}, \frac{1+7\alpha}{8}, \alpha\right) \mid \alpha \text{ real}\right\}$;
(e) $\{(8 - 2\alpha, \alpha - 5, \alpha)\}$;
(f) incompatível; (g) incompatível; (h) incompatível;
(i) $(0, \frac{3}{2}, 1)$; (j) $\{(2 - 6\alpha, 4 + \alpha, 3 - \alpha, \alpha)\}$;
(k) $\{(\frac{15}{4} - \frac{5}{8}\alpha - \beta, -\frac{1}{4} - \frac{1}{8}\alpha, \alpha, \beta)\}$;
(l) $\{(1 + \frac{2}{7}\alpha, \frac{3}{7}\alpha, \alpha)\}$
5. (a) $(0, -1)$; (b) $\{(\frac{3}{4} - \frac{5}{8}\alpha, -\frac{1}{4} - \frac{1}{8}\alpha, \alpha, 3) \mid \alpha \text{ real}\}$;
(c) $\{(0, \alpha, -\alpha)\}$; (d) $\{\alpha(-\frac{4}{3}, 0, \frac{1}{3}, 1)\}$
6. $a \neq -2$
7. $\beta = 2$
8. (a) $a = 5, b = 4$; (b) $a = 5, b \neq 4$
9. (a) $(-2, 2)$; (b) $(-7, 4)$
10. (a) $(-3, 2, 1)$; (b) $(2, -2, 1)$
12. $x_1 = 280, x_2 = 230, x_3 = 350, x_4 = 590$
14. (a) $(5, 3, -2)$; (b) $(2, 4, 2)$; (c) $(2, 0, -2, -2, 0, 2)$

SEÇÃO 3

1. (a) $\begin{pmatrix} 6 & 2 & 8 \\ -4 & 0 & 2 \\ 2 & 4 & 4 \end{pmatrix}$; (b) $\begin{pmatrix} 4 & 1 & 6 \\ -5 & 1 & 2 \\ 3 & -2 & 3 \end{pmatrix}$; (c) $\begin{pmatrix} 3 & 2 & 2 \\ 5 & -3 & -1 \\ -4 & 16 & 1 \end{pmatrix}$;
(d) $\begin{pmatrix} 3 & 5 & -4 \\ 2 & -3 & 16 \\ 2 & -1 & 1 \end{pmatrix}$; (f) $\begin{pmatrix} 5 & 5 & 8 \\ -10 & -1 & -9 \\ 15 & 4 & 6 \end{pmatrix}$; (h) $\begin{pmatrix} 5 & -10 & 15 \\ 5 & -1 & 4 \\ 8 & -9 & 6 \end{pmatrix}$
2. (a) $\begin{pmatrix} 15 & 19 \\ 4 & 0 \end{pmatrix}$; (c) $\begin{pmatrix} 19 & 21 \\ 17 & 21 \\ 8 & 10 \end{pmatrix}$; (d) $\begin{pmatrix} 36 & 10 & 56 \\ 10 & 3 & 16 \end{pmatrix}$
(b) e (e) não são possíveis.
3. (a) 3×3; (b) 1×2
4. (a) $\begin{pmatrix} 3 & 2 \\ 2 & -3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$;
(b) $\begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & -1 \\ 3 & -2 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix}$; (c) $\begin{pmatrix} 2 & 1 & 1 \\ 1 & -1 & 2 \\ 3 & -2 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix}$

10. $A = A^2 = A^3 = A^n$

11. $A^{2n} = I$, $A^{2n+1} = A$

19. Segunda-feira, 575; terça-feira, 936; quarta-feira, 457, 8; quinta-feira, 1105; sexta-feira, 457, 8.

20. 4500 casadas, 5500 solteiras.

21. (b) 0 caminhos de comprimento 2 de V_2 para V_3 e 3 caminhos de comprimento 2 de V_2 para V_5;
(c) 6 caminhos de comprimento 3 de V_2 para V_3 e 2 caminhos de comprimento 3 de V_2 para V_5.

22. (a) $A = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$;

(c) 5 caminhos de comprimento 3 de V_2 para V_4 e 7 caminhos de comprimento menor ou igual a 3.

23. $b = a_{22} - \dfrac{a_{12}a_{21}}{a_{11}}$

SEÇÃO 4

1. (a) Tipo I; (b) não é uma matriz elementar; (c) tipo III; (d) tipo II.

3. (a) $\begin{pmatrix} -2 & 0 \\ 0 & 1 \end{pmatrix}$; (b) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$; (c) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$

4. (a) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$; (b) $\begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix}$; (c) $\begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

5. (a) $E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$; (b) $F = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$

6. (a) $E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$, $E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$

7. (b) (i) $(0, -1, 1)^T$, (ii) $(-4, -2, 5)^T$, (iii) $(0, 3, -2)^T$

8. (a) $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$; (b) $\begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$; (c) $\begin{pmatrix} -4 & 3 \\ \frac{3}{2} & -1 \end{pmatrix}$;

(d) $\begin{pmatrix} \frac{1}{3} & 0 \\ -1 & \frac{1}{3} \end{pmatrix}$; (e) $\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$; (f) $\begin{pmatrix} 3 & 0 & -5 \\ 0 & \frac{1}{3} & 0 \\ -1 & 0 & 2 \end{pmatrix}$;

(g) $\begin{pmatrix} 2 & -3 & 3 \\ -\frac{3}{5} & \frac{6}{5} & -1 \\ -\frac{2}{5} & -\frac{1}{5} & 0 \end{pmatrix}$; (h) $\begin{pmatrix} -\frac{1}{2} & -1 & -\frac{1}{2} \\ -2 & -1 & -1 \\ \frac{3}{2} & 1 & \frac{1}{2} \end{pmatrix}$

9. (a) $\begin{pmatrix} -1 & 0 \\ 4 & 2 \end{pmatrix}$; (b) $\begin{pmatrix} -8 & 5 \\ -14 & 9 \end{pmatrix}$

10. (a) $\begin{pmatrix} 20 & -5 \\ -34 & 7 \end{pmatrix}$; (c) $\begin{pmatrix} 0 & -2 \\ -2 & 2 \end{pmatrix}$

SEÇÃO 5

1. (a) $\begin{pmatrix} I & A^{-1} \end{pmatrix}$; (b) $\begin{pmatrix} I \\ A^{-1} \end{pmatrix}$; (c) $\begin{pmatrix} A^TA & A^T \\ A & I \end{pmatrix}$;

(d) $AA^T + I$; (e) $\begin{pmatrix} I & A^{-1} \\ A & I \end{pmatrix}$

3. (a) $A\mathbf{b}_1 = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$, $A\mathbf{b}_2 = \begin{pmatrix} 4 \\ -1 \end{pmatrix}$;

(b) $\begin{pmatrix} 1 & 1 \end{pmatrix} B = \begin{pmatrix} 3 & 4 \end{pmatrix}$, $\begin{pmatrix} 2 & -1 \end{pmatrix} B = \begin{pmatrix} 3 & -1 \end{pmatrix}$;

(c) $AB = \begin{pmatrix} 3 & 4 \\ 3 & -1 \end{pmatrix}$

4. (a) $\left(\begin{array}{cc|cc} 3 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 \\ \hline 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \end{array}\right)$; (b) $\left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ \hline 3 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{array}\right)$;

(c) $\left(\begin{array}{cc|cc} 2 & 2 & 2 & 2 \\ 2 & 4 & 2 & 2 \\ \hline 3 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 \end{array}\right)$; (d) $\left(\begin{array}{cc|cc} 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ \hline 3 & 2 & 1 & 2 \\ 3 & 1 & 1 & 1 \end{array}\right)$

5. (b) $\left(\begin{array}{ccc|c} 0 & 2 & 0 & -2 \\ 8 & 5 & 8 & -5 \\ 3 & 2 & 3 & -2 \\ \hline 5 & 3 & 5 & -3 \end{array}\right)$; (d) $\left(\begin{array}{cc} 3 & -3 \\ 2 & -2 \\ 1 & -1 \\ \hline 5 & -5 \\ 4 & -4 \end{array}\right)$

9. $A^2 = \begin{pmatrix} B & O \\ O & B \end{pmatrix}$, $A^4 = \begin{pmatrix} B^2 & O \\ O & B^2 \end{pmatrix}$

10. (a) $\begin{pmatrix} O & I \\ I & O \end{pmatrix}$; (b) $\begin{pmatrix} I & O \\ -B & I \end{pmatrix}$

CAPÍTULO 2

SEÇÃO 1

1. (a) $\det(M_{21}) = -8$, $\det(M_{22}) = -2$, $\det(M_{23}) = 5$;
(b) $A_{21} = 8$, $A_{22} = -2$, $A_{23} = -5$

2. (a) e (c) são invertíveis.

3. (a) 1; (b) 4; (c) 0; (d) 58; (e) -39;
(f) 0; (g) 8; (h) 20

4. (a) 2; (b) -4; (c) 0; (d) 0

5. $-x^3 + ax^2 + bx + c$

6. $\lambda = 6$ ou -1

SEÇÃO 2

1. (a) -24; (b) 30; (c) -1

2. (a) 10; (b) 20

3. (a), (e) e (f) são singulares, enquanto (b), (c) e (d) são invertíveis.
4. $c = 5$ ou -3
7. (a) 20; (b) 108; (c) 160; (d) $\frac{5}{4}$
8. (a) -6; (c) 6; (e) 1
11. $\det(A) = u_{11}u_{22}u_{33}$

SEÇÃO 3

1. (a) $\det(A) = -7$, $\operatorname{adj} A = \begin{pmatrix} -1 & -2 \\ -3 & 1 \end{pmatrix}$, $A^{-1} = \begin{pmatrix} \frac{1}{7} & \frac{2}{7} \\ \frac{3}{7} & -\frac{1}{7} \end{pmatrix}$;

   (c) $\det(A) = 3$, $\operatorname{adj} A = \begin{pmatrix} -3 & 5 & 2 \\ 0 & 1 & 1 \\ 6 & -8 & -5 \end{pmatrix}$, $A^{-1} = \frac{1}{3} \operatorname{adj} A$
2. (a) $(\frac{5}{7}, \frac{8}{7})$; (b) $(\frac{11}{5}, -\frac{4}{5})$; (c) $(4, -2, 2)$; (d) $(2, -1, 2)$;
   (e) $(-\frac{2}{3}, \frac{2}{3}, \frac{1}{3}, 0)$
3. $-\frac{3}{4}$
4. $(\frac{1}{2}, -\frac{3}{4}, 1)^T$
5. (a) $\det(A) = 0$, logo A é singular.

   (b) $\operatorname{adj} A = \begin{pmatrix} -1 & 2 & -1 \\ 2 & -4 & 2 \\ -1 & 2 & -1 \end{pmatrix}$ e $A \operatorname{adj} A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$
9. (a) $\det(\operatorname{adj}(A)) = 8$ e $\det(A) = 2$;

   (b) $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & -1 & 1 \\ 0 & -6 & 2 & -2 \\ 0 & 1 & 0 & 1 \end{pmatrix}$
14. RESOLVA OS EXERCÍCIOS PARA CASA.

CAPÍTULO 3

SEÇÃO 1

1. (a) $\|\mathbf{x}_1\| = 10$, $\|\mathbf{x}_2\| = \sqrt{17}$;
   (b) $\|\mathbf{x}_3\| = 13 < \|\mathbf{x}_1\| + \|\mathbf{x}_2\|$
2. (a) $\|\mathbf{x}_1\| = \sqrt{5}$, $\|\mathbf{x}_2\| = 3\sqrt{5}$;
   (b) $\|\mathbf{x}_3\| = 4\sqrt{5} = \|\mathbf{x}_1\| + \|\mathbf{x}_2\|$
7. Se $\mathbf{x} + \mathbf{y} = \mathbf{x}$ para todo $\mathbf{x}$ no espaço vetorial, então $\mathbf{0} = \mathbf{0} + \mathbf{y} = \mathbf{y}$.
8. Se $\mathbf{x} + \mathbf{y} = \mathbf{x} + \mathbf{z}$, então $-\mathbf{x} + (\mathbf{x} + \mathbf{y}) = -\mathbf{x} + (\mathbf{x} + \mathbf{z})$ e o resultado segue usando os axiomas 1, 2, 3 e 4.
11. V não é um espaço vetorial. O axioma 6 não é válido.

SEÇÃO 2

1. (a) e (c) são subespaços; (b) e (d) não são.
2. (b) e (c) são subespaços; (a) e (d) não são.

3. (a), (b), (d) e (e) são subespaços; (c) e (f) não são.
4. (a) $\{(0, 0)^T\}$; (b) $[(-2, 1, 0, 0)^T, (3, 0, 1, 0)^T]$; (c) $[(1, 1, 1)^T]$; (d) $[(-5, 0, -3, 1)^T, (-1, 1, 0, 0)^T]$.
5. Apenas o conjunto no item (c) é um subespaço de P_4.
6. (a), (b) e (d) são subespaços.
9. (a), (c) e (e) são conjuntos geradores.
10. (a) e (b) são conjuntos geradores.
12. (b) e (c).

SEÇÃO 3

1. (a) e (e) são linearmente independentes; (b), (c) e (d) são linearmente dependentes.
2. (a) e (e) são linearmente independentes; (b), (c) e (d) não são.
3. (a) e (b) geram todo o espaço tridimensional; (c) um plano contendo (0, 0, 0); (d) uma reta contendo (0, 0, 0); (e) um plano contendo (0, 0, 0).
4. (a) Linearmente independentes; (b) linearmente independentes; (c) linearmente dependentes.
5. (a) e (b) são linearmente dependentes, enquanto (c) e (d) são linearmente independentes.
8. Quando α for um múltiplo ímpar de $\pi/2$. Se o gráfico de $y = \cos x$ for deslocado para a direita ou para a esquerda por um múltiplo ímpar de $\pi/2$, obtemos o gráfico de sen x ou de $-$sen x.

SEÇÃO 4

1. Apenas os dos itens (a) e (e) formam uma base.
2. Apenas os do item (a) formam uma base.
3. (c) 2.
4. 1.
5. (c) 2; (d) um plano contendo (0, 0, 0) no espaço tridimensional.
6. (b) $\{(1, 1, 1)^T\}$, dimensão 1; (c) $\{(1, 0, 1)^T, (0, 1, 1)^T\}$, dimensão 2.
7. Base $\{(1, 1, 0, 0)^T, (1, -1, 1, 0)^T, (0, 2, 0, 1)^T\}$
10. $\{x^2 + 2, x + 3\}$
11. (a) $\{E_{11}, E_{22}\}$; (b) $\{E_{11}, E_{21}, E_{22}\}$; (d) $\{E_{12}, E_{21}, E_{22}\}$; (e) $\{E_{11}, E_{22}, E_{21} + E_{12}\}$
12. 2
13. (a) 3; (b) 3; (c) 2; (d) 2
14. (a) $\{x, x^2\}$; (b) $\{x - 1, (x - 1)^2\}$; (c) $\{x(x - 1)\}$

SEÇÃO 5

1. (a) $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$; (b) $\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$; (c) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
2. (a) $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$; (b) $\begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix}$; (c) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

3. (a) $\begin{pmatrix} \frac{5}{2} & \frac{7}{2} \\ -\frac{1}{2} & -\frac{1}{2} \end{pmatrix}$; (b) $\begin{pmatrix} 11 & 14 \\ -4 & -5 \end{pmatrix}$; (c) $\begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix}$

4. $[\mathbf{x}]_E = (-1, 2)^T$, $[\mathbf{y}]_E = (5, -8)^T$, $[\mathbf{z}]_E = (-1, 5)^T$

5. (a) $\begin{pmatrix} 2 & 0 & -1 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$;
(b) (i) $(1, -4, 3)^T$, (ii) $(0, -1, 1)^T$, (iii) $(2, 2, -1)^T$

6. (a) $\begin{pmatrix} 1 & -1 & -2 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$; (b) $\begin{pmatrix} 7 \\ 5 \\ -2 \end{pmatrix}$

7. $\mathbf{w}_1 = (5, 9)^T$ e $\mathbf{w}_2 = (1, 4)^T$

8. $\mathbf{u}_1 = (0, -1)^T$ e $\mathbf{u}_2 = (1, 5)^T$

9. (a) $\begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix}$; (b) $\begin{pmatrix} \frac{1}{4} & -\frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} \end{pmatrix}$

10. $\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$

SEÇÃO 6

2. (a) 3; (b) 3; (c) 2.
3. (a) $\mathbf{u}_2$, $\mathbf{u}_4$ e $\mathbf{u}_5$ são os vetores-colunas de U que correspondem às variáveis livres. $\mathbf{u}_2 = 2\mathbf{u}_1$, $\mathbf{u}_4 = 5\mathbf{u}_1 - \mathbf{u}_3$, $\mathbf{u}_5 = -3\mathbf{u}_1 + 2\mathbf{u}_3$.
4. (a) Compatível; (b) incompatível; (e) compatível.
5. (a) Uma infinidade de soluções; (c) uma única solução.
8. Posto de $A = 3$; $\dim N(B) = 1$.
16. (b) $n - 1$.
21. Se $\mathbf{x}_j$ for uma solução de $A\mathbf{x} = \mathbf{e}_j$ para $j = 1, ..., m$ e $X = (\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m)$, então $AX = I_m$.

CAPÍTULO 4

SEÇÃO 1

1. (a) Reflexão em relação ao eixo dos x_2; (b) reflexão em relação à origem; (c) reflexão em relação à reta $x_2 = x_1$; (d) o comprimento do vetor é dividido por 2; (e) projeção sobre o eixo dos x_2.
4. Todas, exceto o item (c), são transformações lineares de R^3 em R^2.
5. (b) e (c) são transformações lineares de R^2 em R^3.
6. (a), (b) e (d) são transformações lineares.
7. (a) e (c) são transformações lineares de P_2 em P_3.
8. $L(e^x) = e^x - 1$ e $L(x^2) = x^3/3$.
9. (a) e (c) são transformações lineares de $C[0, 1]$ em R^1.
15. (a) $\ker(L) = \{\mathbf{0}\}$, $L(R^3) = R^3$; (c) $\ker(L) = \text{Span}[\mathbf{e}_2, \mathbf{e}_3]$, $L(R^3) = \text{Span}[(1, 1, 1)^T]$.

16. (a) $L(S) = \text{Span}[\mathbf{e}_2, \mathbf{e}_3]$; (b) $L(S) = \text{Span}[\mathbf{e}_1, \mathbf{e}_2]$.

17. (a) $\ker(L) = P_1$, $L(P_3) = \text{Span}[x^2, x]$;
(c) $\ker(L) = \text{Span}[x^2 - x]$, $L(P_3) = P_2$.

21. O operador no item (a) é injetor e sobrejetor.

SEÇÃO 2

1. (a) $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$; (b) $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$; (c) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$;
(d) $\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$; (e) $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

2. (a) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$; (b) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$; (c) $\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$

3. (a) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$; (b) $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$; (c) $\begin{pmatrix} 0 & 0 & 2 \\ 3 & 1 & 0 \\ 2 & 0 & -1 \end{pmatrix}$

4. (a) $(0, 0, 0)^T$; (b) $(2, -1, -1)^T$; (c) $(-15, 9, 6)^T$

5. (a) $\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$; (b) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$; (c) $\begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix}$;
(d) $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$

6. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$

7. (b) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}$

8. (a) $\begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 0 & -2 & -1 \end{pmatrix}$; (b) (i) $7\mathbf{y}_1 + 6\mathbf{y}_2 - 8\mathbf{y}_3$,
(ii) $3\mathbf{y}_1 + 3\mathbf{y}_2 - 3\mathbf{y}_3$, (iii) $\mathbf{y}_1 + 5\mathbf{y}_2 + 3\mathbf{y}_3$

9. $\begin{pmatrix} 1 & \frac{1}{2} \\ 1 & 0 \end{pmatrix}$

10. $\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ -2 & 0 & 0 \end{pmatrix}$; (a) $\begin{pmatrix} \frac{1}{2} \\ -2 \end{pmatrix}$; (d) $\begin{pmatrix} 5 \\ -8 \end{pmatrix}$

11. $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$

14. (a) $\begin{pmatrix} -1 & -3 & 1 \\ 0 & 2 & 0 \end{pmatrix}$ (c) $\begin{pmatrix} 2 & -2 & -4 \\ -1 & 3 & 3 \end{pmatrix}$

SEÇÃO 3

1. Para a matriz A, veja as respostas do Exercício 1 da Seção 2.

(a) $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$; (b) $B = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$; (c) $B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$; (d) $B = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$; (e) $B = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$

2. (a) $\begin{pmatrix} 1 & 1 \\ -1 & -3 \end{pmatrix}$; (b) $\begin{pmatrix} 1 & 0 \\ -4 & -1 \end{pmatrix}$

3. (a) $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$; (b) $\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$

4. $V = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & -2 \\ 1 & 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

5. (a) $\begin{pmatrix} 0 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$; (b) $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$; (c) $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$; (d) $a_1x + a_2 2^n(1 + x^2)$

6. (a) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$; (b) $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$; (c) $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$

CAPÍTULO 5

SEÇÃO 1

1. (a) 0°; (b) 90°

2. (a) $\sqrt{14}$ (projeção escalar), $(2, 1, 3)^T$ (projeção vetorial); (b) $0, \mathbf{0}$;

(c) $\dfrac{14\sqrt{13}}{13}, (\frac{42}{13}, \frac{28}{13})^T$; (d) $\dfrac{8\sqrt{21}}{21}, (\frac{8}{21}, \frac{16}{21}, \frac{32}{21})^T$

3. (a) $\mathbf{p} = (3, 0)^T, \mathbf{x} - \mathbf{p} = (0, 4)^T, \mathbf{p}^T(\mathbf{x} - \mathbf{p}) = 3 \cdot 0 + 0 \cdot 4 = 0$;

(c) $\mathbf{p} = (3, 3, 3)^T, \mathbf{x} - \mathbf{p} = (-1, 1, 0)^T, \mathbf{p}^T(\mathbf{x} - \mathbf{p}) = -1 \cdot 3 + 1 \cdot 3 + 0 \cdot 3 = 0$

4. (1,8, 3,6)

5. (1,4, 3,8)

6. 0,4

7. (a) $2x + 4y + 3z = 0$; (c) $z - 4 = 0$

8. $\frac{5}{3}$

9. $\frac{8}{7}$

SEÇÃO 2

1. (a) $\{(3, 4)^T\}$ é uma base para $I(A^T)$, $\{(-4, 3)^T\}$ é uma base para $N(A)$, $\{(1, 2)^T\}$ é uma base para $I(A)$, $\{(-2, 1)^T\}$ é uma base para $N(A^T)$;

(d) base para $I(A^T)$: $\{(1, 0, 0, 0)^T, (0, 1, 0, 0)^T, (0, 0, 1, 1)^T\}$, base para $N(A)$: $\{(0, 0, -1, 1)^T\}$, base para $I(A)$: $\{(1, 0, 0, 1)^T, (0, 1, 0, 1)^T, (0, 0, 1, 1)^T\}$, base para $N(A^T)$: $\{(1, 1, 1, -1)^T\}$.

2. (a) $\{(1, 1, 0)^T, (-1, 0, 1)^T\}$.

3. (b) O complemento ortogonal é gerado por $(-5, 3, 1)^T$.

4. $\{(-1, 2, 0, 1)^T, (2, -3, 1, 0)^T\}$ é uma base para $S^\perp$.

5. (a) $\mathbf{N} = (8, -2, 1)^T$; (b) $8x - 2y + z = 7$.

9. $\dim N(A) = n - r$, $\dim N(A^T) = m - r$.

SEÇÃO 3

1. $\|\mathbf{x}\|_2 = 2$, $\|\mathbf{y}\|_2 = 6$, $\|\mathbf{x}+\mathbf{y}\|_2 = 2\sqrt{10}$

2. (a) $\theta = \dfrac{\pi}{4}$; $\mathbf{p} = (\frac{4}{3}, \frac{1}{3}, \frac{1}{3}, 0)^T$

3. (b) $\|\mathbf{x}\| = 1$, $\|\mathbf{y}\| = 3$

4. (a) 0; (b) 5; (c) 7; (d) $\sqrt{74}$

7. (a) 1; (b) $\dfrac{1}{\pi}$; (c) $\frac{1}{6}$

8. (a) $\dfrac{\pi}{6}$; (b) $\mathbf{p} = \frac{3}{2}x$

11. (a) $\dfrac{\sqrt{10}}{2}$; (b) $\dfrac{\sqrt{34}}{4}$

15. (a) $\|\mathbf{x}\|_1 = 7$, $\|\mathbf{x}\|_2 = 5$, $\|\mathbf{x}\|_\infty = 4$; (b) $\|\mathbf{x}\|_1 = 4$, $\|\mathbf{x}\|_2 = \sqrt{6}$, $\|\mathbf{x}\|_\infty = 2$;
(c) $\|\mathbf{x}\|_1 = 3$, $\|\mathbf{x}\|_2 = \sqrt{3}$, $\|\mathbf{x}\|_\infty = 1$

16. $\|\mathbf{x}-\mathbf{y}\|_1 = 5$, $\|\mathbf{x}-\mathbf{y}\|_2 = 3$, $\|\mathbf{x}-\mathbf{y}\|_\infty = 2$

26. (a) Não é uma norma; (b) é norma; (c) é norma.

SEÇÃO 4

1. (a) $(2, 1)^T$; (c) $(1{,}6, 0{,}6, 1{,}2)^T$

2. (1a) $\mathbf{p} = (3, 1, 0)^T$, $\mathbf{r} = (0, 0, 2)^T$ (1c)
$\mathbf{p} = (3{,}4, 0{,}2, 0{,}6, 2{,}8)^T$, $\mathbf{r} = (0{,}6, -0{,}2, 0{,}4, -0{,}8)^T$

3. (a) $\{(1 - 2\alpha, \alpha)^T \mid \alpha \text{ real}\}$; (b) $\{(2 - 2\alpha, 1 - \alpha, \alpha)^T \mid \alpha \text{ real}\}$

4. (a) $\mathbf{p} = (1, 2, -1)^T$, $\mathbf{b} - \mathbf{p} = (2, 0, 2)^T$;
(b) $\mathbf{p} = (3, 1, 4)^T$, $\mathbf{p} - \mathbf{b} = (-5, -1, 4)^T$

5. (a) $y = 1{,}8 + 2{,}9x$

6. $0{,}55 + 1{,}65x + 1{,}25x^2$

SEÇÃO 5

1. (a) e (d)

2. (b) $\mathbf{x} = -\dfrac{\sqrt{2}}{3}\mathbf{x}_1 + \dfrac{5}{3}\mathbf{x}_2$, $\|\mathbf{x}\| = \left[\left(-\dfrac{\sqrt{2}}{3}\right)^2 + \left(\dfrac{5}{3}\right)^2\right]^{1/2} =$
$= \sqrt{3}$

3. $\mathbf{p} = (\frac{23}{18}, \frac{41}{18}, \frac{8}{9})^T$, $\mathbf{p} - \mathbf{x} = (\frac{5}{18}, \frac{5}{18}, -\frac{10}{9})^T$

4. (b) $c_1 = y_1 \cos\theta + y_2 \operatorname{sen}\theta$, $c_2 = -y_1 \operatorname{sen}\theta + y_2 \cos\theta$

6. (a) 15; (b) $\|\mathbf{u}\| = 3$, $\|\mathbf{v}\| = 5\sqrt{2}$; (c) $\dfrac{\pi}{4}$

8. (b) (i) 0, (ii) $-\dfrac{\pi}{2}$, (iii) 0, (iv) $\dfrac{\pi}{8}$

17. (b) (i) $(2, -2)^T$, (ii) $(5, 2)^T$, (iii) $(3, 1)^T$

18. (a) $P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$

19. (b) $Q = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$

23. (b) $\|1\| = \sqrt{2}$, $\|x\| = \dfrac{\sqrt{6}}{3}$; (c) $l(x) = \dfrac{9}{7}x$

SEÇÃO 6

1. (a) $\left\{\left(-\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}\right)^T, \left(\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}\right)^T\right\}$; (b) $\left\{\left(\dfrac{2}{\sqrt{5}}, \dfrac{1}{\sqrt{5}}\right)^T, \left(-\dfrac{1}{\sqrt{5}}, \dfrac{2}{\sqrt{5}}\right)^T\right\}$

2. (a) $\begin{pmatrix} -\dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 0 & 4\sqrt{2} \end{pmatrix}$;

(b) $\begin{pmatrix} \dfrac{2}{\sqrt{5}} & -\dfrac{1}{\sqrt{5}} \\ \dfrac{1}{\sqrt{5}} & \dfrac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \sqrt{5} & 4\sqrt{5} \\ 0 & 3\sqrt{5} \end{pmatrix}$

3. $\{(\frac{1}{3}, \frac{2}{3}, -\frac{2}{3})^T, (\frac{2}{3}, \frac{1}{3}, \frac{2}{3})^T, (-\frac{2}{3}, \frac{2}{3}, \frac{1}{3})^T\}$

4. $u_1(x) = \dfrac{1}{\sqrt{2}}$, $u_2(x) = \dfrac{\sqrt{6}}{2}x$, $u_3(x) = \dfrac{3\sqrt{10}}{4}\left(x^2 - \dfrac{1}{3}\right)$

5. (a) $\left\{\dfrac{1}{3}(2, 1, 2)^T, \dfrac{\sqrt{2}}{6}(-1, 4, -1)^T\right\}$;

(b) $Q = \begin{pmatrix} \dfrac{2}{3} & \dfrac{-\sqrt{2}}{6} \\ \dfrac{1}{3} & \dfrac{2\sqrt{2}}{3} \\ \dfrac{2}{3} & \dfrac{-\sqrt{2}}{6} \end{pmatrix}$; $R = \begin{pmatrix} 3 & \dfrac{5}{3} \\ 0 & \dfrac{\sqrt{2}}{3} \end{pmatrix}$; (c) $\mathbf{x} = \begin{pmatrix} 9 \\ -3 \end{pmatrix}$

6. (b) $\begin{pmatrix} \frac{3}{5} & -\frac{4}{5\sqrt{2}} \\ \frac{4}{5} & \frac{3}{5\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 5 & 1 \\ 0 & 2\sqrt{2} \end{pmatrix}$; (c) $(2{,}1, 5{,}5)^T$

7. $\left\{ \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right)^T, \left(\frac{\sqrt{2}}{3}, \frac{\sqrt{2}}{3}, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{6}\right)^T \right\}$

9. $\left\{ \left(\frac{4}{5}, \frac{2}{5}, \frac{2}{5}, \frac{1}{5}\right)^T, \left(\frac{1}{5}, -\frac{2}{5}, -\frac{2}{5}, \frac{4}{5}\right)^T, \left(0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right)^T \right\}$

SEÇÃO 7

1. (a) $T_4 = 8x^4 - 8x^2 + 1$, $T_5 = 16x^5 - 20x^3 + 5x$;
(b) $H_4 = 16x^4 - 48x^2 + 12$, $H_5 = 32x^5 - 160x^3 + 120x$

2. $p_1(x) = x$, $p_2(x) = x^2 - \frac{4}{\pi} + 1$

4. $p(x) = (\operatorname{senh} 1)P_0(x) + \frac{3}{e}P_1(x), +5\left(\operatorname{senh} 1 - \frac{3}{e}\right)P_2(x)$,
$p(x) \approx 0{,}9963 + 1{,}1036x + 0{,}5367x^2$

6. (a) $U_0 = 1$, $U_1 = 2x$, $U_2 = 4x^2 - 1$

11. $p(x) = (x-2)(x-3) + (x-1)(x-3) + 2(x-1)(x-2)$

13. $1 \cdot f\left(-\frac{1}{\sqrt{3}}\right) + 1 \cdot f\left(\frac{1}{\sqrt{3}}\right)$

14. (a) Grau menor ou igual a 3; (b) a fórmula dá a resposta exata para a primeira integral. O valor aproximado da segunda integral é 1,5, enquanto a resposta exata é $\pi/2$.

CAPÍTULO 6

SEÇÃO 1

1. (a) $\lambda_1 = 5$, o auto-espaço é gerado por $(1, 1)^T$,
$\lambda_2 = -1$, o auto-espaço é gerado por $(1, -2)^T$;
(b) $\lambda_1 = 3$, o auto-espaço é gerado por $(4, 3)^T$,
$\lambda_2 = 2$, o auto-espaço é gerado por $(1, 1)^T$;
(c) $\lambda_1 = \lambda_2 = 2$, o auto-espaço é gerado por $(1, 1)^T$;
(d) $\lambda_1 = 3 + 4i$, o auto-espaço é gerado por $(2i, 1)^T$,
$\lambda_2 = 3 - 4i$, o auto-espaço é gerado por $(-2i, 1)^T$;
(e) $\lambda_1 = 2 + i$, o auto-espaço é gerado por $(1, 1 + i)^T$,
$\lambda_2 = 2 - i$, o auto-espaço é gerado por $(1, 1 - i)^T$;
(f) $\lambda_1 = \lambda_2 = \lambda_3 = 0$, o auto-espaço é gerado por $(1, 0, 0)^T$;
(g) $\lambda_1 = 2$, o auto-espaço é gerado por $(1, 1, 0)^T$,
$\lambda_2 = 1$, o auto-espaço é gerado por $(1, 0, 0)^T$, $(0, 1, -1)^T$;
(h) $\lambda_1 = 1$, o auto-espaço é gerado por $(1, 0, 0)^T$,
$\lambda_2 = 4$, o auto-espaço é gerado por $(1, 1, 1)^T$,
$\lambda_3 = -2$, o auto-espaço é gerado por $(-1, -1, 5)^T$;

(i) $\lambda_1 = 2$, o auto-espaço é gerado por $(7, 3, 1)^T$,
$\lambda_2 = 1$, o auto-espaço é gerado por $(3, 2, 1)^T$,
$\lambda_3 = 0$, o auto-espaço é gerado por $(1, 1, 1)^T$;
(j) $\lambda_1 = \lambda_2 = \lambda_3 = -1$, o auto-espaço é gerado por $(1, 0, 1)^T$,
(k) $\lambda_1 = \lambda_2 = 2$, o auto-espaço é gerado por $\mathbf{e}_1$ e $\mathbf{e}_2$,
$\lambda_3 = 3$, o auto-espaço é gerado por $\mathbf{e}_3$,
$\lambda_4 = 4$, o auto-espaço é gerado por $\mathbf{e}_4$;
(l) $\lambda_1 = 3$, o auto-espaço é gerado por $(1, 2, 0, 0)^T$,
$\lambda_2 = 1$, o auto-espaço é gerado por $(0, 1, 0, 0)^T$,
$\lambda_3 = \lambda_4 = 2$, o auto-espaço é gerado por $(0, 0, 1, 0)^T$.

8. β é um autovalor de B se e somente se $\beta = \lambda - \alpha$ para algum autovalor λ de A.

11. $\lambda_1 = 6, \lambda_2 = 2$

24. $\lambda_1\mathbf{x}^T\mathbf{y} = (A\mathbf{x})^T\mathbf{y} = \mathbf{x}^T A^T\mathbf{y} = \lambda_2\mathbf{x}^T\mathbf{y}$

SEÇÃO 2

1. (a) $\begin{pmatrix} c_1e^{2t} + c_2e^{3t} \\ c_1e^{2t} + 2c_2e^{3t} \end{pmatrix}$; (b) $\begin{pmatrix} -c_1e^{-2t} - 4c_2e^{t} \\ c_1e^{-2t} + c_2e^{t} \end{pmatrix}$;

(c) $\begin{pmatrix} 2c_1 + c_2e^{5t} \\ c_1 - 2c_2e^{5t} \end{pmatrix}$; (d) $\begin{pmatrix} -c_1e^{t}\operatorname{sen} t + c_2e^{t}\cos t \\ c_1e^{t}\cos t + c_2e^{t}\operatorname{sen} t \end{pmatrix}$;

(e) $\begin{pmatrix} -c_1e^{3t}\operatorname{sen} 2t + c_2e^{3t}\cos 2t \\ c_1e^{3t}\cos 2t + c_2e^{3t}\operatorname{sen} 2t \end{pmatrix}$; (f) $\begin{pmatrix} -c_1 + c_2e^{5t} + c_3e^{t} \\ -3c_1 + 8c_2e^{5t} \\ c_1 + 4c_2e^{5t} \end{pmatrix}$

2. (a) $\begin{pmatrix} e^{-3t} + 2e^{t} \\ -e^{-3t} + 2e^{t} \end{pmatrix}$; (b) $\begin{pmatrix} e^{t}\cos 2t + 2e^{t}\operatorname{sen} 2t \\ e^{t}\operatorname{sen} 2t - 2e^{t}\cos 2t \end{pmatrix}$;

(c) $\begin{pmatrix} -6e^{t} + 2e^{-t} + 6 \\ -3e^{t} + e^{-t} + 4 \\ -e^{t} + e^{-t} + 2 \end{pmatrix}$; (d) $\begin{pmatrix} -2 - 3e^{t} + 6e^{2t} \\ 1 + 3e^{t} - 3e^{2t} \\ 1 + 3e^{2t} \end{pmatrix}$

4. $y_1(t) = 15e^{-0{,}24t} + 25e^{-0{,}08t}$, $y_2(t) = -30e^{-0{,}24t} + 50e^{-0{,}08t}$

5. (a) $\begin{pmatrix} -2c_1e^{t} - 2c_2e^{-t} + c_3e^{\sqrt{2}t} + c_4e^{-\sqrt{2}t} \\ c_1e^{t} + c_2e^{-t} - c_3e^{\sqrt{2}t} - c_4e^{-\sqrt{2}t} \end{pmatrix}$;

(b) $\begin{pmatrix} c_1e^{2t} + c_2e^{-2t} - c_3e^{t} - c_4e^{-t} \\ c_1e^{2t} - c_2e^{-2t} + c_3e^{t} - c_4e^{-t} \end{pmatrix}$

6. $y_1(t) = -e^{2t} + e^{-2t} + e^{t}$; $y_2(t) = -e^{2t} - e^{-2t} + 2e^{t}$

8. $x_1(t) = \cos t + 3\operatorname{sen} t + \dfrac{1}{\sqrt{3}}\operatorname{sen}\sqrt{3}t$,

$x_2(t) = \cos t + 3\operatorname{sen} t - \dfrac{1}{\sqrt{3}}\operatorname{sen}\sqrt{3}t$

10. (a) $m_1x_1''(t) = -kx_1 + k(x_2 - x_1),$
$m_2x_2''(t) = -k(x_2 - x_1) + k(x_3 - x_2),$
$m_3x_3''(t) = -k(x_3 - x_2) - kx_3;$

(b) $\begin{pmatrix} 0{,}1\cos 2\sqrt{3}t + 0{,}9\cos\sqrt{2}t \\ -0{,}2\cos 2\sqrt{3}t + 1{,}2\cos\sqrt{2}t \\ 0{,}1\cos 2\sqrt{3}t + 0{,}9\cos\sqrt{2}t \end{pmatrix}$

11. $p(\lambda) = (-1)^n(\lambda^n - a_{n-1}\lambda^{n-1} - \cdots - a_1\lambda - a_0)$

SEÇÃO 3

8. (b) $\alpha = 2$; (c) $\alpha = 3$ ou $\alpha = -1$; (d) $\alpha = 1$

19. (a) $A = \begin{pmatrix} 0{,}70 & 0{,}20 & 0{,}10 \\ 0{,}20 & 0{,}70 & 0{,}10 \\ 0{,}10 & 0{,}10 & 0{,}80 \end{pmatrix}$

(c) A quantidade de pessoas em cada um dos três grupos vai se aproximar de 100.000 quando n torna-se muito grande.

21. (b) $\begin{pmatrix} e & e \\ 0 & e \end{pmatrix}$

22. (a) $\begin{pmatrix} 3 - 2e & 1 - e \\ -6 + 6e & -2 + 3e \end{pmatrix}$;

(c) $\begin{pmatrix} e & -1 + e & -1 + e \\ 1 - e & 2 - e & 1 - e \\ -1 + e & -1 + e & e \end{pmatrix}$

23. (a) $\begin{pmatrix} e^{-t} \\ e^{-t} \end{pmatrix}$; (b) $\begin{pmatrix} -3e^t - e^{-t} \\ e^t + e^{-t} \end{pmatrix}$; (c) $\begin{pmatrix} 3e^t - 2 \\ 2 - e^{-t} \\ e^{-t} \end{pmatrix}$

SEÇÃO 4

1. (a) $\|\mathbf{z}\| = 6, \|\mathbf{w}\| = 3, \langle \mathbf{z}, \mathbf{w}\rangle = -4 + 4i, \langle \mathbf{w}, \mathbf{z}\rangle =$
$= -4 - 4i;$

(b) $\|\mathbf{z}\| = 4, \|\mathbf{w}\| = 7, \langle \mathbf{z}, \mathbf{w}\rangle = -4 + 10i, \langle \mathbf{w}, \mathbf{z}\rangle =$
$= -4 - 10i$

2. (b) $\mathbf{z} = 4\mathbf{z}_1 + 2\sqrt{2}\mathbf{z}_2$

3. (a) $\mathbf{u}_1^H\mathbf{z} = 4 + 2i, \mathbf{z}^H\mathbf{u}_1 = 4 - 2i, \mathbf{u}_2^H\mathbf{z} = 6 - 5i, \mathbf{z}^H\mathbf{u}_2 =$
$= 6 + 5i;$

(b) $\|\mathbf{z}\| = 9$

4. (b) e (f) são auto-adjuntas, enquanto (b), (c), (e) e (f) são normais.

11. (b) $\|U\mathbf{x}\|^2 = (U\mathbf{x})^HU\mathbf{x} = \mathbf{x}^HU^HU\mathbf{x} = \mathbf{x}^H\mathbf{x} = \|\mathbf{x}\|^2$.

12. U é unitária, já que $U^HU = (I - 2\mathbf{u}\mathbf{u}^H)^2 = I - 4\mathbf{u}\mathbf{u}^H + 4\mathbf{u}(\mathbf{u}^H\mathbf{u})\mathbf{u}^H = I$.

20. $\lambda_1 = 1, \lambda_2 = -1, \mathbf{u}_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T, \mathbf{u}_2 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T,$

$A = 1\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} + (-1)\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$

SEÇÃO 5

1. (a) $\begin{pmatrix} 3 & -\frac{5}{2} \\ -\frac{5}{2} & 1 \end{pmatrix}$; (b) $\begin{pmatrix} 2 & \frac{1}{2} & -1 \\ \frac{1}{2} & 3 & \frac{3}{2} \\ -1 & \frac{3}{2} & 1 \end{pmatrix}$

3. (a) $Q = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, $\frac{(x')^2}{4} + \frac{(y')^2}{12} = 1$, elipse;

(d) $Q = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, $\left(y' + \frac{\sqrt{2}}{2}\right)^2 =$

$= -\frac{\sqrt{2}}{2}(x' - \sqrt{2})$ ou $(y'')^2 = -\frac{\sqrt{2}}{2}x''$, parábola

6. (a) Positiva definida; (b) indefinida; (d) negativa definida; (e) indefinida.

7. (a) Mínimo; (b) ponto de sela; (c) ponto de sela; (f) máximo local.

SEÇÃO 6

1. (a) $\det(A_1) = 2$, $\det(A_2) = 3$, positiva definida;
(b) $\det(A_1) = 3$, $\det(A_2) = -10$, não é positiva definida;
(c) $\det(A_1) = 6$, $\det(A_2) = 14$, $\det(A_3) = -38$, não é positiva definida;
(d) $\det(A_1) = 4$, $\det(A_2) = 8$, $\det(A_3) = 13$, positiva definida;

2. $a_{11} = 3$, $a_{22}^{(1)} = 2$, $a_{33}^{(2)} = \frac{4}{3}$

3. (a) $\begin{pmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix}\begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix}$;

(b) $\begin{pmatrix} 1 & 0 \\ -\frac{1}{3} & 1 \end{pmatrix}\begin{pmatrix} 9 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -\frac{1}{3} \\ 0 & 1 \end{pmatrix}$;

(c) $\begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{4} & -1 & 1 \end{pmatrix}\begin{pmatrix} 16 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{4} \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$;

(d) $\begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ -\frac{2}{3} & 1 & 1 \end{pmatrix}\begin{pmatrix} 9 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & \frac{1}{3} & -\frac{2}{3} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

4. (a) $\begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$; (b) $\begin{pmatrix} 3 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ 0 & 1 \end{pmatrix}$;

(c) $\begin{pmatrix} 4 & 0 & 0 \\ 2 & \sqrt{2} & 0 \\ 1 & -\sqrt{2} & 2 \end{pmatrix}\begin{pmatrix} 4 & 2 & 1 \\ 0 & \sqrt{2} & -\sqrt{2} \\ 0 & 0 & 2 \end{pmatrix}$;

(d) $\begin{pmatrix} 3 & 0 & 0 \\ 1 & \sqrt{3} & 0 \\ -2 & \sqrt{3} & \sqrt{2} \end{pmatrix}\begin{pmatrix} 3 & 1 & -2 \\ 0 & \sqrt{3} & \sqrt{3} \\ 0 & 0 & \sqrt{2} \end{pmatrix}$

SEÇÃO 7

1. (a) $\lambda_1 = 4, \lambda_2 = -1, \mathbf{x}_1 = (3, 2)^T$;
(b) $\lambda_1 = 8, \lambda_2 = 3, \mathbf{x}_1 = (1, 2)^T$;
(c) $\lambda_1 = 7, \lambda_2 = 2, \lambda_3 = 0, \mathbf{x}_1 = (1, 1, 1)^T$

2. (a) $\lambda_1 = 3, \lambda_2 = -1, \mathbf{x}_1 = (3, 1)^T$;
(b) $\lambda_1 = 2 = 2\exp(0), \lambda_2 = -2 = 2\exp(\pi i), \mathbf{x}_1 = (1, 1)^T$;
(c) $\lambda_1 = 2 = 2\exp(0), \lambda_2 = -1 + \sqrt{3}i = 2\exp\left(\frac{2\pi i}{3}\right)$,
$\lambda_3 = -1 - \sqrt{3}i = 2\exp\left(\frac{4\pi i}{3}\right), \mathbf{x}_1 = (4, 2, 1)^T$

3. $x_1 = 70.000, x_2 = 56.000, x_3 = 44.000$

4. $x_1 = x_2 = x_3$

5. $(I - A)^{-1} = I + A + \cdots + A^{m-1}$

6. (a) $(I - A)^{-1} = \begin{pmatrix} 1 & -1 & 3 \\ 0 & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix}$;

(b) $A^2 = \begin{pmatrix} 0 & -2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

7. (b) e (c) são redutíveis.

CAPÍTULO 7

SEÇÃO 1

1. (a) $0{,}231 \times 10^4$; (b) $0{,}326 \times 10^2$; (c) $0{,}128 \times 10^{-1}$;
(d) $0{,}824 \times 10^5$

2. (a) $\epsilon = -2$; $\delta \approx -8{,}7 \times 10^{-4}$;
(b) $\epsilon = 0{,}04$; $\delta \approx 1{,}2 \times 10^{-3}$;
(c) $\epsilon = 3{,}0 \times 10^{-5}$; $\delta \approx 2{,}3 \times 10^{-3}$;
(d) $\epsilon = -31$; $\delta \approx -3{,}8 \times 10^{-4}$

3. (a) $0{,}10101 \times 2^5$; (b) $0{,}10100 \times 2^{-1}$; (c) $0{,}10111 \times 2^4$;
(d) $-0{,}11010 \times 2^{-3}$

4. (a) $10.420, \epsilon = -0{,}0018, \delta \approx -1{,}7 \times 10^{-7}$;
(b) $0, \epsilon = -8, \delta = -1$;
(c) $1 \times 10^{-4}, \epsilon = 5 \times 10^{-5}, \delta = 1$;
(d) $82{,}190, \epsilon = 25{,}7504, \delta \approx 3{,}1 \times 10^{-4}$

5. (a) $0{,}1043 \times 10^6$; (b) $0{,}1045 \times 10^6$; (c) $0{,}1045 \times 10^6$

6. $\epsilon = (0{,}00001)_2 = \frac{1}{32}$

7. 23

SEÇÃO 2

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

2. (a) $(2, -1, 3)^T$; (b) $(1, -1, 3)^T$; (c) $(1, 5, 1)^T$
3. (a) n^2 multiplicações e $n(n-1)$ somas;
   (b) n^3 multiplicações e $n^2(n-1)$ somas;
   (c) $(AB)\mathbf{x}$ precisa de $n^3 + n^2$ multiplicações e $n^3 - n$ somas; $A(B\mathbf{x})$ precisa de $2n^2$ multiplicações e $2n(n-1)$ somas.
4. (b) (i) 156 multiplicações e 105 somas, (ii) 47 multiplicações e 24 somas, (iii) 100 multiplicações e 60 somas.
8. $5n - 4$ multiplicações/divisões, $3n - 3$ somas/subtrações.
9. (a) $[(n-j)(n-j+1)]/2$ multiplicações, $[(n-j-1)(n-j)]/2$ somas; (c) dada a fatoração LU, são necessárias da ordem de $(2/3)\,n^3$ de multiplicações/divisões adicionais para calcular A^{-1}.

SEÇÃO 3

1. (a) $(1, 1, -2)$;

   (b) $\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & -2 \\ 0 & 1 & 8 \\ 0 & 0 & -23 \end{pmatrix}$
2. (a) $(1, 2, 2)$; (b) $(4, -3, 0)$; (c) $(1, 1, 1)$
3. $P = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, $L = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & -\frac{1}{3} & 1 \end{pmatrix}$,

   $U = \begin{pmatrix} 2 & 4 & -6 \\ 0 & 6 & 9 \\ 0 & 0 & 5 \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} 6 \\ -\frac{1}{2} \\ 1 \end{pmatrix}$
4. $P = Q = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $PAQ = LU = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{pmatrix}\begin{pmatrix} 4 & 2 \\ 0 & 2 \end{pmatrix}$
5. (a) $\hat{\mathbf{c}} = P\mathbf{c} = (-4, 6)^T$, $\mathbf{y} = L^{-1}\hat{\mathbf{c}} = (-4, 8)^T$, $\mathbf{z} = U^{-1}\mathbf{y} = (-3, 4)^T$; (b) $\mathbf{x} = Q\mathbf{z} = (4, -3)^T$
6. (b) $P = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$, $Q = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$,

   $L = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{2} & \frac{2}{3} & 1 \end{pmatrix}$, $U = \begin{pmatrix} 8 & 6 & 2 \\ 0 & 6 & 3 \\ 0 & 0 & 2 \end{pmatrix}$
7. Erro $\dfrac{-2000\varepsilon}{0,6} \approx -3333\varepsilon$. Se $\varepsilon = 0{,}001$, então $\delta = -\dfrac{2}{3}$.
8. (1,667, 1,001)
9. (5,002, 1,000)
10. (5,001, 1,001)

SEÇÃO 4

1. (a) $\|A\|_F = \sqrt{2}$, $\|A\|_\infty = 1$, $\|A\|_1 = 1$;
   (b) $\|A\|_F = 5$, $\|A\|_\infty = 5$, $\|A\|_1 = 6$;

(c) $\|A\|_F = \|A\|_\infty = \|A\|_1 = 1$;
(d) $\|A\|_F = 7$, $\|A\|_\infty = 6$, $\|A\|_1 = 10$;
(e) $\|A\|_F = 9$, $\|A\|_\infty = 10$, $\|A\|_1 = 12$

2. 2

4. $\|I\|_1 = \|I\|_\infty = 1$, $\|I\|_F = \sqrt{n}$

6. (a) 10; (b) $(-1, 1, -1)^T$

12. (a) $\|A\mathbf{x}\|_\infty \leq \|A\mathbf{x}\|_2 \leq \|A\|_2\|\mathbf{x}\|_2 \leq \sqrt{n}\|A\|_2\|\mathbf{x}\|_\infty$

14. $\text{cond}_\infty A = 400$

15. (a) $(-0{,}48, 0{,}8)$; (b) $(-2{,}902. 2{,}0)$

16. $\text{cond}_\infty(A) = 28$

18. (a) $A_n^{-1} = \begin{pmatrix} 1-n & n \\ n & -n \end{pmatrix}$; (b) $\text{cond}_\infty A_n = 4n$;

(c) $\lim_{n\to\infty} \text{cond}_\infty A_n = \infty$

19. (a) $\mathbf{r} = (-0{,}06, 0{,}02)^T$ e o resíduo relativo é 0,012; (b) 20; (d) $\mathbf{x} = (1, 1)^T$, $\|\mathbf{x} - \mathbf{x}'\| = 0{,}12$.

20. $\text{cond}_1(A) = 6$

21. 0,3

22. (a) $\|\mathbf{r}\|_\infty = 0{,}10$, $\text{cond}_\infty(A) = 32$; (b) 0,64;
(c) $\mathbf{x} = (12{,}50, 4{,}26, 2{,}14, 1{,}10)^T$, $\delta = 0{,}04$

SEÇÃO 5

1. (a) $\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$; (b) $\begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$; (c) $\begin{pmatrix} -\frac{4}{5} & \frac{3}{5} \\ -\frac{3}{5} & -\frac{4}{5} \end{pmatrix}$

2. (a) $\begin{pmatrix} \frac{3}{5} & 0 & \frac{4}{5} \\ 0 & 1 & 0 \\ \frac{4}{5} & 0 & -\frac{3}{5} \end{pmatrix}$; (b) $\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$;

(c) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$; (d) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$

3. $H = I - (1/\beta)\mathbf{v}\mathbf{v}^T$ para β e $\mathbf{v}$ dados.
(a) $\beta = 9$, $\mathbf{v} = (-1, -1, -4)^T$; (b) $\beta = 7$, $\mathbf{v} = (-1, 2, 3)^T$;
(c) $\beta = 18$, $\mathbf{v} = (-2, 4, -4)^T$;

4. (a) $\beta = 9$, $\mathbf{v} = (0, -1, 4, 1)^T$;
(b) $\beta = 15$, $\mathbf{v} = (0, 0, -5, -1, 2)^T$

5. (a) $\beta = 18$, $\mathbf{v} = (-3, 1, 1, 5)^T$; (b) $HA = \begin{pmatrix} 6 & 0 & 6 \\ 0 & 2 & 0 \\ 0 & -4 & -6 \\ 0 & 4 & -6 \end{pmatrix}$

6. (a) $H_2H_1A = R$, onde $H_i = I - \dfrac{1}{\beta_i}\mathbf{v}_i\mathbf{v}_i^T$, $i = 1, 2$, e $\beta_1 = 6, \beta_2 = 5$.

$$\mathbf{v}_1 = \begin{pmatrix} -2 \\ 2 \\ -2 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ -1 \\ 3 \end{pmatrix}, \quad R = \begin{pmatrix} 3 & 4 & -2 \\ 0 & 5 & 10 \\ 0 & 0 & -5 \end{pmatrix},$$

$\mathbf{b}^{(1)} = H_2H_1\mathbf{b} = (11, 5, 5)^T$;

(b) $\mathbf{x} = (-1, 3, -1)^T$

7. (a) $G = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{pmatrix}$, $\quad \mathbf{x} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

8. Para determinar H, são necessárias três multiplicações, duas somas e uma raiz quadrada. Para determinar G, são necessárias quatro multiplicações/divisões, uma soma e uma raiz quadrada. Para calcular GA são necessárias $4n$ multiplicações e $2n$ adições, enquanto o cálculo de HA necessita de $3n$ multiplicações/divisões e $3n$ somas.

9. (a) $n - k + 1$ multiplicações/divisões, $2n - 2k + 1$ somas;
(b) $n(n - k + 1)$ multiplicações/divisões, $n(2n - 2k + 1)$ somas.

10. (a) $4(n - k)$ multiplicações/divisões, $2(n - k)$ somas;
(b) $4n(n - k)$ multiplicações/divisões, $2n(n - k)$ somas.

11. (a) Rotação; (b) rotação; (c) transformação de Givens; (d) transformação de Givens.

SEÇÃO 6

2. (a) $\sigma_1 = \sqrt{10}$, $\sigma_2 = 0$; (b) $\sigma_1 = 3$, $\sigma_2 = 2$; (c) $\sigma_1 = 4$, $\sigma_2 = 2$; (d) $\sigma_1 = 3$, $\sigma_2 = 2$, $\sigma_3 = 1$. As matrizes U e V não são únicas. O leitor pode verificar suas respostas multiplicando as matrizes $U\Sigma V^T$.

3. (b) Posto de $A = 2$, $\|A\|_2 = 3$, $A' = \begin{pmatrix} 1,2 & -2,4 \\ -0,6 & 1,2 \end{pmatrix}$

4. (a) $\text{cond}_2(A) = 10$;

(b) a matriz mais próxima de posto 2 é $\begin{pmatrix} -2 & 8 & 20 \\ 14 & 19 & 10 \\ 0 & 0 & 0 \end{pmatrix}$,

a matriz mais próxima de posto 1 é $\begin{pmatrix} 6 & 12 & 12 \\ 8 & 16 & 16 \\ 0 & 0 & 0 \end{pmatrix}$

5. (a) Base para $I(A^T)$: $\{\mathbf{v}_1 = (2/3, 2/3, 1/3)^T, \mathbf{v}_2 = (-2/3, 1/3, 2/3)^T\}$, base para $N(A)$: $\{\mathbf{v}_3 = (1/3, -2/3, 2/3)^T\}$.

6. $\sigma_1 = 8$, $\sigma_2 = 8$, $\sigma_3 = 4$.

SEÇÃO 7

1. (a) $\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$; (b) $A_2 = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$; (c) $\lambda_1 = 2$, $\lambda_2 = 0$; o auto-espaço associado a λ_1 é gerado por $\mathbf{u}_1$.

2. (a) $\mathbf{v}_1 = \begin{pmatrix} 3 \\ 5 \\ 3 \end{pmatrix}$, $\mathbf{u}_1 = \begin{pmatrix} 0{,}6 \\ 1{,}0 \\ 0{,}6 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 2{,}2 \\ 4{,}2 \\ 2{,}2 \end{pmatrix}$,

$\mathbf{u}_2 = \begin{pmatrix} 0{,}52 \\ 1{,}00 \\ 0{,}52 \end{pmatrix}$,

$\mathbf{v}_3 = \begin{pmatrix} 2.05 \\ 4{,}05 \\ 2{,}05 \end{pmatrix}$; (b) $\lambda_1' = 4{,}05$;

(c) $\lambda_1 = 4$, $\delta = 0{,}0125$

3. (b) A não tem autovalor dominante.

4. $A_2 = \begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}$, $A_3 = \begin{pmatrix} 3{,}4 & 0{,}2 \\ 0{,}2 & 0{,}6 \end{pmatrix}$,
$\lambda_1 = 2 + \sqrt{2} \approx 3{,}414$, $\lambda_2 = 2 - \sqrt{2} \approx 0{,}586$

5. (b) $H = I - \dfrac{1}{\beta}\mathbf{v}\mathbf{v}^T$, onde $\beta = \frac{1}{3}$ e $\mathbf{v} = (-\frac{1}{3}, -\frac{2}{3}, \frac{1}{3})^T$;

(c) $\lambda_2 = 3$, $\lambda_3 = 1$, $HAH = \begin{pmatrix} 4 & 0 & 3 \\ 0 & 5 & -4 \\ 0 & 2 & -1 \end{pmatrix}$

SEÇÃO 8

1. (a) $(\sqrt{2}, 0)^T$; (b) $(1 - 3\sqrt{2}, 3\sqrt{2}, -\sqrt{2})^T$; (c) $(1, 0)^T$;
(d) $(1 - \sqrt{2}, \sqrt{2}, -\sqrt{2})^T$

2. $x_i = \dfrac{d_i b_i + e_i b_{n+i}}{d_i^2 + e_i^2}$, $i = 1, \ldots, n$

3. (a) $\sigma_1 = \sqrt{2 + \epsilon^2}$, $\sigma_2 = \epsilon$;
(b) $\lambda_1' = 2$, $\lambda_2' = 0$, $\sigma_1' = \sqrt{2}$, $\sigma_2' = 0$

10. $A^+ = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 \end{pmatrix}$

11. (a) $A^+ = \begin{pmatrix} \frac{1}{10} & -\frac{1}{10} \\ \frac{2}{10} & -\frac{2}{10} \end{pmatrix}$; (b) $A^+\mathbf{b} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$;
(c) $\left\{ \mathbf{y} \mid \mathbf{y} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \alpha \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\}$

13. $\|A_1 - A_2\|_F = \epsilon$, $\|A_1^+ - A_2^+\|_F = 1/\epsilon$. As $\epsilon \to 0$,
$\|A_1 - A_2\|_F \to 0$ e $\|A_1^+ - A_2^+\|_F \to \infty$.

ÍNDICE

Mat 65 5.2

13 271 43 236 64 668 65 249 01271