# 3

# Matrizes

> Nós [Halmos e Kaplansky] compartilhamos uma filosofia sobre álgebra linear: pensamos em base-livre, escrevemos em base-livre, mas, quando as dificuldades surgem, fechamos a porta de nossos escritórios e calculamos com matrizes ferozmente.
>
> – Irving Kaplansky
> Em *Paul Halmos: Celebrating 50 years of mathematics*
> J. H. Ewing e F. W. Gehring, Eds.
> Springer-Verlag, 1991, p. 88

## 3.1 Introdução: Matrizes em Ação

Neste capítulo, estudaremos as matrizes por si mesmas. Já usamos matrizes — na forma de matrizes completas — para gravar informações sobre sistemas de equações lineares e para facilitar a resolução deles. Agora, veremos que as matrizes têm propriedades algébricas próprias, que nos capacitam a calcular seguindo as regras da álgebra de matrizes. Além disso, observaremos que matrizes não são objetos estáticos, que gravam informações e dados; mais do que isso, elas representam certos tipos de funções e "agem" em vetores, transformando-os em outros vetores. Essas "transformações por meio de matrizes" começarão a representar um papel-chave em nosso estudo da álgebra linear e darão uma nova visão do que aprendemos sobre vetores e sobre sistemas de equações lineares. Mais ainda, matrizes surgem em muitas formas além da forma de matrizes completas; investigaremos algumas das muitas aplicações das matrizes no final deste capítulo.

Nesta seção, apresentaremos alguns exemplos simples para ilustrar como matrizes podem transformar vetores. No processo, você terá uma primeira noção da "aritmética matricial".

Considere as equações

$$\begin{aligned} y_1 &= x_1 + 2x_2 \\ y_2 &= \phantom{x_1 + {}} 3x_2 \end{aligned} \qquad (1)$$

Podemos ver que essas equações descrevem a transformação do vetor $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ no vetor $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. Se denotarmos a matriz de coeficientes do lado direito da equação por $F$, então $F = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$, e poderemos reescrever a transformação como

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

ou, mais sucintamente, $\mathbf{y} = F\mathbf{x}$. [Pense nessa expressão como análoga à notação de função $y = f(x)$, com a qual você está familiarizado: $\mathbf{x}$ é aqui a "variável" independente, $\mathbf{y}$ é a "variável" dependente e $F$ é o nome da "função".]

Assim, se $\mathbf{x} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, as equações (1) resultam em

$$\begin{aligned} y_1 &= -2 + 2 \cdot 1 = 0 \\ y_2 &= \phantom{-2+{}} 3 \cdot 1 = 3 \end{aligned} \quad \text{ou} \quad \mathbf{y} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

Podemos escrever essa expressão como $\begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

**Problema 1:** Calcule $F\mathbf{x}$ para os seguintes vetores $\mathbf{x}$:

(a) $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ (b) $\mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ (c) $\mathbf{x} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ (d) $\mathbf{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

**Problema 2:** Os quatro vetores $\mathbf{x}$ do Problema 1 representam os quatro vértices de um quadrado no plano $x_1x_2$. Desenhe esse quadrado e nomeie os vértices de $A'$, $B'$, $C'$ e $D'$, correspondendo às partes (a), (b), (c), e (d) do Problema 1.

Em um outro sistema de coordenadas (chamando os eixos de $y_1$ e $y_2$), marque os quatro pontos determinados por $F\mathbf{x}$ no Problema 1. Nomeie esses pontos como $A'$, $B'$, $C'$ e $D'$. Façamos a suposição (razoável) de que o segmento de reta $\overline{AB}$ seja transformado no segmento de reta $\overline{A'B'}$, o mesmo acontecendo com os outros três lados do quadrado $ABCD$. Que figura geométrica é representada por $A'B'C'D'$?

**Problema 3:** O centro do quadrado $ABCD$ é a origem $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Qual é o centro de $A'$, $B'$, $C'$ e $D'$? Que cálculo algébrico confirma isso?

Agora, considere as equações

$$\begin{aligned} z_1 &= y_1 - y_2 \\ z_2 &= -2y_1 \end{aligned} \tag{2}$$

que transformam um vetor $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ no vetor $\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$. Podemos abreviar essa transformação como $\mathbf{z} = G\mathbf{y}$, em que

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

**Problema 4:** Descobriremos como $G$ transforma a figura $A'\ B'\ C'\ D'$. Calcule $G\mathbf{y}$ para cada um dos quatro vetores $\mathbf{y}$ que você calculou no Problema 1 [isto é, calcule $\mathbf{z} = G(F\mathbf{x})$. Você pode reconhecer essa expressão como análoga à composição de funções que já lhe é familiar]. Chame os novos pontos de $A''$, $B''$, $C''$ e $D''$ e desenhe a figura $A''B''C''D''$ nos eixos coordenados $z_1z_2$.

**Problema 5:** Substituindo as equações (1) nas equações (2), obtemos equações para $z_1$ e $z_2$ em termos de $x_1$ e $x_2$. Se denotarmos a matriz dessas equações como $H$, teremos $\mathbf{z} = H\mathbf{x}$. Como também temos que $\mathbf{z} = GF\mathbf{x}$, é razoável escrever

$$H = GF$$

Você consegue perceber como os elementos de $H$ estão relacionados com os elementos de $F$ e de $G$?

**Problema 6:** Vamos resolver de trás para a frente o processo que acabamos de ver. Primeiro, transforme o quadrado $ABCD$, usando $G$, para obter a figura $A^*B^*C^*D^*$. Depois, transforme a figura resultante usando $F$ para obter $A^{**}B^{**}C^{**}D^{**}$. [*Observação*: não se preocupe aqui com as "variáveis" **x**, **y** e **z**. Simplesmente substitua as coordenadas de $A$, $B$, $C$ e $D$ nas equações (2) e então substitua os resultados nas equações (1).] $A^{**}B^{**}C^{**}D^{**}$ e $A''B''C''D''$ são iguais? O que isso lhe mostra sobre a ordem em que aplicamos as transformações $F$ e $G$?

**Problema 7:** Repita o Problema 5 com as matrizes genéricas

$$F = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}, \quad G = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \quad \text{e} \quad H = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}$$

Se as equações (1) e (2) tiverem coeficientes como especificado por $F$ e $G$, ache os elementos de $H$ em termos dos elementos de $F$ e $G$. O resultado será uma fórmula para o "produto" $H = GF$.

**Problema 8:** Repita os Problemas de 1 a 6 com as matrizes a seguir. (Sua fórmula obtida no Problema 7 pode ajudar a agilizar os cálculos algébricos.) Note as semelhanças e diferenças que você considerar significantes.

$$\text{(a)}\ F = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, G = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \qquad \text{(b)}\ F = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, G = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\text{(c)}\ F = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, G = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \qquad \text{(d)}\ F = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}, G = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

## 3.2 Operações com Matrizes

Apesar de já termos encontrado matrizes, começamos dando uma definição formal e registrando alguns fatos para referência futura.

**Definição** Uma ***matriz*** é uma tabela retangular de números chamados de ***elementos*** ou ***termos*** da matriz.[1]

Estes são exemplos de matrizes:

$$\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} \sqrt{5} & -1 & 0 \\ 2 & \pi & \frac{1}{2} \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 4 \\ 17 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 5{,}1 & 1{,}2 & -1 \\ 6{,}9 & 0 & 4{,}4 \\ -7{,}3 & 9 & 8{,}5 \end{bmatrix}, \quad [7]$$

A ***ordem*** de uma matriz descreve o número de linhas e colunas que ela tem. Uma matriz é chamada $m \times n$ (pronuncia-se "$m$ por $n$") quando tem $m$ linhas e $n$ colunas. Conseqüentemente, os exemplos que acabamos de ver são matrizes de ordem $2\times2$, $2\times3$, $3\times1$, $1\times4$, $3\times3$ e $1\times1$, respectivamente. Uma matriz $1\times m$ é chamada de ***matriz linha*** (ou ***vetor linha***), e uma matriz $n\times1$ é chamada de ***matriz coluna*** (ou ***vetor coluna***).[2]

---

[1] Embora esses números sejam normalmente escolhidos do conjunto dos números reais, $\mathbb{R}$, eles também podem ser tomados do conjunto dos números complexos, $\mathbb{C}$, ou dos inteiros módulo $p$, $\mathbb{Z}_p$, onde $p$ é primo.

[2] Tecnicamente, existe diferença entre matrizes linha/coluna e vetores, mas não focaremos essa distinção. Entretanto, distinguiremos matrizes linha/vetores de matrizes coluna/vetores. Essa distinção é importante — pelo menos — para os cálculos algébricos, como demonstraremos.

Usamos a notação de *índice subscrito duplo* para nos referirmos aos elementos de uma matriz $A$: o elemento de $A$ na linha $i$ e coluna $j$ é denotado por $a_{ij}$. Assim, se

$$A = \begin{bmatrix} 3 & 9 & -1 \\ 0 & 5 & 4 \end{bmatrix}$$

$a_{13} = -1$ e $a_{22} = 5$ (a notação $A_{ij}$ é usada algumas vezes no lugar de $a_{ij}$). Podemos denotar uma matriz $A$ compactamente por $[a_{ij}]$ (ou $[a_{ij}]_{m\times n}$, se for importante especificar a ordem de $A$, embora a ordem usualmente esteja clara a partir do contexto).

Com essa notação, uma matriz genérica $m\times n$ $A$ tem a forma

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

Se as colunas de $A$ são os vetores $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, podemos representar $A$ por

$$A = [\mathbf{a}_1 \ \mathbf{a}_2 \ \ldots \ \mathbf{a}_n]$$

Se as linhas de $A$ são $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_m$, podemos representar $A$ por

$$A = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_m \end{bmatrix}$$

Os ***elementos da diagonal*** de $A$ são $a_{11}, a_{22}, a_{33}, \ldots$, e, se $m = n$ (isto é, se $A$ tem o mesmo número de linhas e de colunas), $A$ é chamada de ***matriz quadrada***. Uma matriz quadrada cujos elementos fora da diagonal são todos zero é chamada de ***matriz diagonal***. Uma matriz diagonal em que todos os elementos da diagonal são iguais é chamada de ***matriz escalar***. Se o escalar na diagonal for 1, a matriz escalar é chamada de ***matriz identidade***.

Por exemplo, sejam

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

Os elementos da diagonal de $A$ são 2 e 4, mas $A$ não é quadrada; $B$ é uma matriz quadrada de ordem $2\times2$ cujos elementos da diagonal são 3 e 5; $C$ é uma matriz diagonal; $D$ é a matriz identidade $3\times3$. A matriz identidade $n\times n$ é denotada por $I_n$ (ou simplesmente por $I$, se sua ordem estiver subentendida).

Por podermos ver matrizes como generalizações de vetores (de fato, matrizes podem e devem ser imaginadas como inventadas a partir de ambos os vetores linha e coluna), muitas das convenções e operações para vetores são levadas (de modo natural) para matrizes.

Duas matrizes são ***iguais*** quando têm a mesma ordem e os elementos correspondentes são iguais. Assim, se $A = [a_{ij}]_{m\times n}$ e $B = [b_{ij}]_{r\times s}$, então $A = B$ se, e somente se, $m = r$, $n = s$ e $a_{ij} = b_{ij}$ para todos $i$ e $j$.

**EXEMPLO 1** Considere as matrizes

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 \\ 5 & 3 \end{bmatrix} \quad \text{e} \quad C = \begin{bmatrix} 2 & 0 & x \\ 5 & 3 & y \end{bmatrix}$$

Nem $A$ nem $B$ podem ser iguais a $C$ (independentemente dos valores de $x$ e $y$), já que $A$ e $B$ são matrizes $2\times 2$ e $C$ é $2\times 3$. Entretanto, $A = B$ se, e somente se, $a = 2$, $b = 0$, $c = 5$ e $d = 3$.

**EXEMPLO 2** Considere as matrizes

$$R = \begin{bmatrix} 1 & 4 & 3 \end{bmatrix} \quad \text{e} \quad C = \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}$$

Apesar de $R$ e C terem os mesmos elementos e estes aparecerem na mesma ordem, $R \neq C$, pois $R$ é $1\times 3$ e $C$ é $3\times 1$. (Se lermos $R$ e $C$ em voz alta, elas parecerão iguais: "um, quatro, três".) Portanto, a distinção entre matrizes linha/vetores e matrizes coluna/vetores é importante.

## *Adição e Multiplicação por Escalar de Matrizes*

Generalizando da adição de vetores, definimos a adição de matrizes *elemento a elemento*. Se $A = [a_{ij}]$ e $B = [b_{ij}]$ são matrizes $m\times n$, sua ***soma*** $A + B$ é a matriz $m\times n$ obtida adicionando-se os elementos correspondentes. Assim,

$$A + B = [a_{ij} + b_{ij}]$$

[Poderíamos igualmente ter definido $A + B$ em termos de adição de vetores, especificando que cada coluna (ou linha) de $A + B$ fosse a soma das colunas (ou linhas) correspondentes de $A$ e $B$.] Se $A$ e $B$ não tiverem a mesma ordem, $A + B$ não estará definida.

**EXEMPLO 3** Sejam

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

Então

$$A + B = \begin{bmatrix} -2 & 5 & -1 \\ 1 & 6 & 7 \end{bmatrix}$$

Entretanto, nem $A + C$ nem $B + C$ estão definidas.

A definição de multiplicação de matriz por escalar elemento a elemento não trará surpresa. Se $A$ é uma matriz $m\times n$ e $c$ é um escalar, o ***múltiplo escalar*** $cA$ é a matriz $m\times n$ obtida pela multiplicação de cada elemento de $A$ por $c$. Mais formalmente, temos

$$cA = c[a_{ij}] = [ca_{ij}]$$

[Em termos de vetores, poderíamos, equivalentemente, estipular que cada coluna (ou linha) de $cA$ fosse $c$ vezes a coluna (ou linha) correspondente de $A$.]

**EXEMPLO 4** Para a matriz $A$ do Exemplo 3,

$$2A = \begin{bmatrix} 2 & 8 & 0 \\ -4 & 12 & 10 \end{bmatrix}, \quad \tfrac{1}{2}A = \begin{bmatrix} \tfrac{1}{2} & 2 & 0 \\ -1 & 3 & \tfrac{5}{2} \end{bmatrix} \quad \text{e} \quad (-1)A = \begin{bmatrix} -1 & -4 & 0 \\ 2 & -6 & -5 \end{bmatrix}$$

A matriz $(-1)A$ é escrita como $-A$ e chamada de ***oposta*** de $A$. Como com vetores, podemos usar esse fato para definir a ***diferença*** de duas matrizes: se $A$ e $B$ têm a mesma ordem, então

$$A - B = A + (-B)$$

**EXEMPLO 5** Para as matrizes $A$ e $B$ do Exemplo 3,

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

Uma matriz $A$ com todos os elementos zero é chamada de ***matriz nula*** e denotada por $O$ (ou $O_{m\times n}$, se for importante especificar sua ordem). Deve ficar claro que, se $A$ for uma matriz qualquer e $O$ for a matriz nula de mesma ordem, então

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

e

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

## *Multiplicação de Matrizes*

A Introdução na Seção 3.1 deu a entender que existe um "produto" de matrizes que é análogo à composição de funções. Agora tornaremos essa noção mais precisa. A definição que daremos generaliza o que você deve ter descoberto nos Problemas 5 e 7 da Seção 3.1. Ao contrário das definições de adição de matriz e de multiplicação de matriz por escalar, a definição do produto de duas matrizes *não* é uma definição elemento a elemento. É claro que não há o que nos impeça de definir um produto de matrizes elemento a elemento;[3] lamentavelmente, tal definição tem poucas aplicações e não é tão "natural" como a que damos agora.

**Definição** Se $A$ é uma matriz $m\times n$ e $B$ é uma matriz $n\times r$, então o ***produto*** $C = AB$ é uma matriz $m\times r$. O elemento $(i, j)$ do produto é calculado da seguinte forma:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$$

Obsevações:

- Note que $A$ e $B$ não precisam ter a mesma ordem. Entretanto, o número de *colunas* de $A$ deve ser igual ao número de *linhas* de $B$. Se escrevermos as ordens de $A$, $B$ e $AB$ nesta mesma ordem, poderemos ver rapidamente se esse requisito está satisfeito. Mais ainda, poderemos dizer a ordem do produto antes mesmo de fazer contas, já que o número de *linhas* de $AB$ será o mesmo número de linhas de $A$, enquanto o número de *colunas* de $AB$ será o mesmo número de colunas de $B$, como se mostra a seguir:

[3] A esse respeito, matemáticos são algumas vezes como Humpty Dumpty, de Lewis Carroll: "Quando *eu* uso uma palavra", disse Humpty Dumpty, "ela significa apenas aquilo que eu quis dizer — nem mais, nem menos" (tirado de *Through the looking glass*).

$$\underset{m \times n}{A} \quad \underset{n \times r}{B} = \underset{m \times r}{AB}$$

Mesmo

Ordem de $AB$

◆ A fórmula para os elementos do produto lembra um produto escalar, e na verdade ela é. Essa fórmula diz que o elemento $(i, j)$ da matriz $AB$ é o produto escalar da $i$-ésima linha de $A$ com a $j$-ésima coluna de $B$:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1r} \\ b_{21} & \cdots & b_{2j} & \cdots & b_{2r} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nj} & \cdots & b_{nr} \end{bmatrix}$$

Observe que, na expressão $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$, os "subscritos externos" em cada termo $ab$ da soma são sempre $i$ e $j$, enquanto os "subscritos internos" são sempre iguais e variam de 1 a $n$. Vemos esse padrão claramente quando usamos a notação com somatório:

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

**EXEMPLO 6** Calcule $AB$, dados

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

SOLUÇÃO: Como $A$ tem ordem 2×3 e $B$ tem ordem 3×4, o produto $AB$ está definido e será uma matriz 2×4. A primeira linha do produto $C = AB$ é calculada fazendo-se o produto escalar da primeira linha de $A$ com cada uma das colunas de $B$. Assim,

$$\begin{aligned} c_{11} &= 1(-4) + 3(5) + (-1)(-1) = 12 \\ c_{12} &= 1(0) + 3(-2) + (-1)(2) = -8 \\ c_{13} &= 1(3) + 3(-1) + (-1)(0) = 0 \\ c_{14} &= 1(-1) + 3(1) + (-1)(6) = -4 \end{aligned}$$

A segunda linha de $C$ é calculada fazendo-se o produto escalar da segunda linha de $A$ com cada uma das colunas de $B$:

$$\begin{aligned} c_{21} &= (-2)(-4) + (-1)(5) + (1)(-1) = 2 \\ c_{22} &= (-2)(0) + (-1)(-2) + (1)(2) = 4 \\ c_{23} &= (-2)(3) + (-1)(-1) + (1)(0) = -5 \\ c_{24} &= (-2)(-1) + (-1)(1) + (1)(6) = 7 \end{aligned}$$

Assim, a matriz produto é

$$AB = \begin{bmatrix} 12 & -8 & 0 & -4 \\ 2 & 4 & -5 & 7 \end{bmatrix}$$

(Com um pouco de prática, você será capaz de fazer esses cálculos mentalmente sem ter que escrever todos os detalhes, como fizemos aqui. Para exemplos mais complicados, uma calculadora com operações matriciais ou um sistema computacional algébrico é preferível.)

Antes de continuarmos, apresentaremos dois exemplos que justificam nossa escolha da definição de multiplicação de matriz.

**EXEMPLO 7** Ana e Beto estão planejando comprar frutas para a próxima semana. Cada um deles quer comprar algumas maçãs, tangerinas e laranjas, porém em quantidades diferentes. A Tabela 1 mostra o que eles pretendem comprar. Nas proximidades existem duas bancas de frutas — a do Sam e a do Téo — cujos preços estão apresentados na Tabela 2. Quanto gastarão Ana e Beto para fazer suas compras em cada uma das duas bancas?

Tabela 1

| | Maçãs | Tangerinas | Laranjas |
|---|---|---|---|
| Ana | 6 | 3 | 10 |
| Beto | 4 | 8 | 5 |

Tabela 2

| | Sam | Téo |
|---|---|---|
| Maçã | \$ 0,10 | \$ 0,15 |
| Tangerina | \$ 0,40 | \$ 0,30 |
| Laranja | \$ 0,10 | \$ 0,20 |

SOLUÇÃO: Se Ana comprar na banca de Sam, gastará

$$6(0{,}10) + 3(0{,}40) + 10(0{,}10) = \$2{,}80$$

Se ela comprar na banca de Téo, gastará

$$6(0{,}15) + 3(0{,}30) + 10(0{,}20) = \$3{,}80$$

Beto gastará, na banca de Sam,

$$4(0{,}10) + 8(0{,}40) + 5(0{,}10) = \$4{,}10$$

e, na banca de Téo,

$$4(0{,}15) + 8(0{,}30) + 5(0{,}20) = \$4{,}00$$

(Provavelmente, Ana fará suas compras na banca de Sam e Beto na banca de Téo.)

A "forma de produto escalar" desses cálculos mostra que a multiplicação de matrizes funciona aqui. Se organizarmos as informações dadas em uma matriz demanda $D$ e uma matriz de preços $P$, teremos

$$D = \begin{bmatrix} 6 & 3 & 10 \\ 4 & 8 & 5 \end{bmatrix} \quad \text{e} \quad P = \begin{bmatrix} 0{,}10 & 0{,}15 \\ 0{,}40 & 0{,}30 \\ 0{,}10 & 0{,}20 \end{bmatrix}$$

Esses cálculos são equivalentes a efetuar o produto

$$DP = \begin{bmatrix} 6 & 3 & 10 \\ 4 & 8 & 5 \end{bmatrix} \begin{bmatrix} 0{,}10 & 0{,}15 \\ 0{,}40 & 0{,}30 \\ 0{,}10 & 0{,}20 \end{bmatrix} = \begin{bmatrix} 2{,}80 & 3{,}80 \\ 4{,}10 & 4{,}00 \end{bmatrix}$$

Tabela 3

| | Sam | Téo |
|---|---|---|
| Ana | \$ 2,80 | \$ 3,80 |
| Beto | \$ 4,10 | \$ 4,00 |

Assim, o produto da matriz $DP$ nos mostra quanto a compra de cada um em cada banca irá custar (Tabela 3).

**EXEMPLO 8** Considere o sistema linear

$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 5 \\ -x_1 + 3x_2 + x_3 &= 1 \\ 2x_1 - x_2 + 4x_3 &= 14 \end{aligned} \qquad (1)$$

Observe que o lado esquerdo dessa igualdade surge do produto de matrizes

$$\begin{bmatrix} 1 & -2 & 3 \\ -1 & 3 & 1 \\ 2 & -1 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

portanto, o sistema (1) pode ser escrito da forma

$$\begin{bmatrix} 1 & -2 & 3 \\ -1 & 3 & 1 \\ 2 & -1 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ 14 \end{bmatrix}$$

ou $A\mathbf{x} = \mathbf{b}$, em que $A$ é a matriz de coeficientes, $\mathbf{x}$ é o vetor (coluna) das variáveis e $\mathbf{b}$ é o vetor (coluna) dos termos constantes.

Você não deve ter dificuldades em perceber que *todos* os sistemas lineares podem ser escritos na forma $A\mathbf{x} = \mathbf{b}$. Na verdade, a notação $[A \mid \mathbf{b}]$ para a matriz completa de um sistema linear é simplesmente uma abreviação para a equação matricial $A\mathbf{x} = \mathbf{b}$. Essa forma se revelará como um meio tremendamente útil de expressar um sistema de equações lineares, e nós a exploraremos freqüentemente daqui por diante.

Combinando essa idéia com o Teorema 1 da Seção 2.4, vemos que $A\mathbf{x} = \mathbf{b}$ tem solução se, e somente se, $\mathbf{b}$ for uma combinação linear das colunas de $A$.

Também provaremos ser muito útil um outro fato sobre as matrizes: a multiplicação de uma matriz por um vetor canônico unitário pode ser usada para "selecionar" ou "reproduzir" uma coluna ou uma linha de uma matriz. Seja $A = \begin{bmatrix} 4 & 2 & 1 \\ 0 & 5 & -1 \end{bmatrix}$ e considere os produtos $A\mathbf{e}_3$ e $\mathbf{e}_2A$, com os vetores unitários $\mathbf{e}_3$ e $\mathbf{e}_2$ escolhidos de forma que os produtos façam sentido. Assim,

$$A\mathbf{e}_3 = \begin{bmatrix} 4 & 2 & 1 \\ 0 & 5 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{e} \quad \mathbf{e}_2A = [0 \quad 1]\begin{bmatrix} 4 & 2 & 1 \\ 0 & 5 & -1 \end{bmatrix} = [0 \quad 5 \quad -1]$$

Observe que $A\mathbf{e}_3$ nos dá a terceira coluna de $A$, e $\mathbf{e}_2A$ nos dá a segunda linha de $A$. Enunciamos o resultado geral como um teorema.

◈ **TEOREMA 1**

Sejam $A$ uma matriz $m \times n$, $\mathbf{e}_i$ um vetor canônico unitário $1 \times m$ e $\mathbf{e}_j$ um vetor canônico unitário $n \times 1$. Então:

$\mathbf{e}_i A$ é a $i$-ésima linha da matriz $A$ e
$A\mathbf{e}_j$ é a $j$-ésima coluna da matriz $A$.

DEMONSTRAÇÃO: Provamos (b) e deixamos para provar (a) no Exercício 41. Se $\mathbf{a}_1, \ldots, \mathbf{a}_n$ são as colunas de $A$, o produto $A\mathbf{e}_j$ pode ser escrito na forma

$$A\mathbf{e}_j = 0\mathbf{a}_1 + 0\mathbf{a}_2 + \ldots + 1\mathbf{a}_j + \ldots + 0\mathbf{a}_n = \mathbf{a}_j$$

Poderíamos também provar (b) pelo cálculo:

$$A\mathbf{e}_j = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & & \vdots & & \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

pois o 1 em $\mathbf{e}_j$ é o $j$-ésimo elemento. ◈

## *Decomposição de Matrizes*

Muitas vezes é conveniente olhar uma matriz como composta por um número de ***submatrizes*** menores. Traçando-se linhas horizontais e verticais em uma matriz, podemos ***decompô-la*** em ***blocos***. Existe uma maneira natural de decompor matrizes em blocos, particularmente aquelas provenientes de certas aplicações. Por exemplo, considere a matriz

$$A = \begin{bmatrix} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{bmatrix}$$

Parece natural decompor $A$ na forma

$$\left[\begin{array}{ccc:cc} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 & 0 \\ \hdashline 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{array}\right] = \begin{bmatrix} I & B \\ O & C \end{bmatrix}$$

em que $I$ é a matriz identidade $3 \times 3$, $B$ é $3 \times 2$, $O$ é a matriz nula $2 \times 3$ e $C$ é $2 \times 2$. Dessa forma, podemos ver $A$ como uma matriz $2 \times 2$ cujos elementos são matrizes.

Quando se multiplicam matrizes, há freqüentemente vantagem ao vê-las como matrizes decompostas em blocos. Isso não só revela muitas vezes estruturas escondidas como também quase sempre acelera o cálculo, especialmente quando as matrizes são grandes e têm muitos blocos de zeros. Fica claro que a multiplicação de matrizes decompostas em blocos é igual à multiplicação comum de matrizes.

Começamos considerando alguns casos especiais de matrizes decompostas em blocos. Cada um deles mostra um modo diferente de ver o produto de duas matrizes.

Suponha que $A$ seja uma matriz $m \times n$ e $B$, $n \times r$. O produto $AB$, portanto, existe. Se decompusermos $B$ em termos de seus vetores coluna, como $B = [\mathbf{b}_1 \mid \mathbf{b}_2 \mid \cdots \mid \mathbf{b}_r]$, então

$$AB = A[\mathbf{b}_1 \mid \mathbf{b}_2 \mid \cdots \mid \mathbf{b}_r] = [A\mathbf{b}_1 \mid A\mathbf{b}_2 \mid \cdots \mid A\mathbf{b}_r]$$

Esse resultado é uma conseqüência imediata da definição de multiplicação de matriz. A forma da direita é chamada de ***representação matriz-coluna*** do produto.

**EXEMPLO 9** Se

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

então

$$A\mathbf{b}_1 = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 13 \\ 2 \end{bmatrix} \quad \text{e} \quad A\mathbf{b}_2 = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

Portanto, $AB = [A\mathbf{b}_1 \mid A\mathbf{b}_2] = \left[\begin{array}{c|c} 13 & 5 \\ 2 & -2 \end{array}\right]$. (Verifique fazendo a multiplicação usual.)

Observação:

- Note que a representação matriz-coluna de $AB$ nos permite escrever cada coluna de $AB$ como uma combinação das colunas de $A$, tendo os elementos de $B$ como coeficientes. Por exemplo,

$$\begin{bmatrix} 13 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 \\ -1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

(Veja os Exercícios 23 e 26.)

Suponha que $A$ seja uma matriz $m \times n$ e $B$, $n \times r$. O produto $AB$, portanto, existe. Se decompusermos $A$ em termos de seus vetores linha, como

$$A = \begin{bmatrix} \mathbf{A}_1 \\ \hline \mathbf{A}_2 \\ \hline \vdots \\ \hline \mathbf{A}_m \end{bmatrix}$$

então

$$AB = \begin{bmatrix} \mathbf{A}_1 \\ \hline \mathbf{A}_2 \\ \hline \vdots \\ \hline \mathbf{A}_m \end{bmatrix} B = \begin{bmatrix} \mathbf{A}_1 B \\ \hline \mathbf{A}_2 B \\ \hline \vdots \\ \hline \mathbf{A}_m B \end{bmatrix}$$

Novamente, esse resultado é uma conseqüência da definição de multiplicação de matriz. A forma da direita é chamada de ***representação matriz-linha*** do produto.

**EXEMPLO 10** Usando a representação matriz-linha, calcule $AB$ para as matrizes do Exemplo 9.

SOLUÇÃO: Calculamos

$$\mathbf{A}_1B = [1 \quad 3 \quad 2]\begin{bmatrix} 4 & -1 \\ 1 & 2 \\ 3 & 0 \end{bmatrix} = [13 \quad 5] \quad \text{e} \quad \mathbf{A}_2B = [0 \quad -1 \quad 1]\begin{bmatrix} 4 & -1 \\ 1 & 2 \\ 3 & 0 \end{bmatrix} = [2 \quad -2]$$

$$\text{Portanto, } AB = \left[\begin{array}{c} \mathbf{A}_1B \\ \hdashline \mathbf{A}_2B \end{array}\right] = \left[\begin{array}{cc} 13 & 5 \\ \hdashline 2 & -2 \end{array}\right], \text{ como anteriormente.}$$

A definição da matriz produto $AB$ usa a partição natural de $A$ em linhas e de $B$ em colunas; essa forma pode ser chamada de ***representação linha-coluna*** do produto. Podemos também decompor $A$ em colunas e $B$ em linhas; essa forma é chamada de ***representação coluna-linha*** do produto.

Nesse caso, temos

$$A = [\mathbf{a}_1 \mid \mathbf{a}_2 \mid \cdots \mid \mathbf{a}_n] \quad \text{e} \quad B = \left[\begin{array}{c} \mathbf{B}_1 \\ \hdashline \mathbf{B}_2 \\ \hdashline \vdots \\ \hdashline \mathbf{B}_n \end{array}\right]$$

Assim,

$$AB = [\mathbf{a}_1 \mid \mathbf{a}_2 \mid \cdots \mid \mathbf{a}_n]\left[\begin{array}{c} \mathbf{B}_1 \\ \hdashline \mathbf{B}_2 \\ \hdashline \vdots \\ \hdashline \mathbf{B}_n \end{array}\right] = \mathbf{a}_1\mathbf{B}_1 + \mathbf{a}_2\mathbf{B}_2 + \cdots + \mathbf{a}_n\mathbf{B}_n \tag{2}$$

Observe que essa soma se assemelha a uma expansão de produto escalar; a diferença é que seus termos são matrizes, e não escalares. Tenhamos certeza de que isso faz sentido: cada termo $\mathbf{a}_i\mathbf{B}_i$ é o produto de uma matriz $m\times1$ por uma matriz $1\times r$. Assim, cada $\mathbf{a}_i\mathbf{B}_i$ é uma matriz $m\times r$ da mesma ordem que $AB$. Os produtos $\mathbf{a}_i\mathbf{B}_i$ são chamados de ***produtos externos***, e (2) é chamada de ***expansão em produto externo*** de $AB$.

**EXEMPLO 11** Calcule a expansão em produto externo de $AB$ para as matrizes do Exemplo 9.

SOLUÇÃO: Temos

$$A = [\mathbf{a}_1 \mid \mathbf{a}_2 \mid \mathbf{a}_3] = \left[\begin{array}{c:c:c} 1 & 3 & 2 \\ 0 & -1 & 1 \end{array}\right] \quad \text{e} \quad B = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \mathbf{B}_3 \end{bmatrix} = \begin{bmatrix} 4 & -1 \\ 1 & 2 \\ 3 & 0 \end{bmatrix}$$

Os produtos externos são

$$\mathbf{a}_1\mathbf{B}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}[4 \quad -1] = \begin{bmatrix} 4 & -1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{a}_2\mathbf{B}_2 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}[1 \quad 2] = \begin{bmatrix} 3 & 6 \\ -1 & -2 \end{bmatrix},$$

e

$$\mathbf{a}_3\mathbf{B}_3 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}[3 \quad 0] = \begin{bmatrix} 6 & 0 \\ 3 & 0 \end{bmatrix}$$

(Observe que calcular cada produto externo é exatamente como preencher uma tábua de multiplicação.) Portanto, a expansão em produto externo de $AB$ é

$$\mathbf{a}_1\mathbf{B}_1 + \mathbf{a}_2\mathbf{B}_2 + \mathbf{a}_3\mathbf{B}_3 = \begin{bmatrix} 4 & -1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 3 & 6 \\ -1 & -2 \end{bmatrix} + \begin{bmatrix} 6 & 0 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 13 & 5 \\ 2 & -2 \end{bmatrix} = AB$$

Usaremos a expansão em produto externo nos Capítulos 5 e 7, quando discutirmos o Teorema Espectral e a decomposição do valor singular, respectivamente.

Cada uma das partições mencionadas é um caso particular da partição em geral. Uma matriz $A$ é dita decomposta se linhas horizontais e verticais forem introduzidas subdividindo $A$ em submatrizes chamadas de blocos. A partição permite que $A$ seja escrita como uma matriz cujos elementos são seus blocos.

Por exemplo,

$$A = \left[\begin{array}{ccc|cc} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 & 0 \\ \hline 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{array}\right] \quad \text{e} \quad B = \left[\begin{array}{cc|cc|c} 4 & 3 & 1 & 2 & 1 \\ -1 & 2 & 2 & 1 & 1 \\ 1 & -5 & 3 & 3 & 1 \\ \hline 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 3 \end{array}\right]$$

são matrizes decompostas em blocos. Elas têm a estrutura em blocos

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{bmatrix}$$

Se duas matrizes têm a mesma ordem e são decompostas em blocos da mesma forma, é claro que elas podem ser adicionadas e multiplicadas por escalar, bloco por bloco. Menos evidente é o fato de que, com uma partição adequada, matrizes podem também ser multiplicadas por blocos. O próximo exemplo ilustra esse processo.

**EXEMPLO 12** Considere as matrizes $A$ e $B$ que acabamos de ver. Se ignorarmos por um momento o fato de que seus elementos são matrizes, $A$ assemelha-se a uma matriz $2\times2$ e $B$, a uma matriz $2\times3$. O produto delas deve ser, então, uma matriz $2\times3$, dada por

$$\begin{aligned} AB &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{bmatrix} \\ &= \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} & A_{11}B_{13} + A_{12}B_{23} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} & A_{21}B_{13} + A_{22}B_{23} \end{bmatrix} \end{aligned}$$

No entanto, *todos* os produtos nesse cálculo são na verdade produtos de *matrizes*; precisamos nos assegurar de que todos eles estão definidos. Uma investigação rápida revela que esse é realmente o caso, pois o número de *colunas* nos blocos de $A$ (3 e 2) é igual ao número de *linhas* nos blocos de $B$. Dizemos que as matrizes $A$ e $B$ foram ***decompostas adequadamente para multiplicação por blocos***.

Explicitando os cálculos indicados, temos o produto $AB$ na forma:

$$A_{11}B_{11} + A_{12}B_{21} = I_3B_{11} + A_{12}I_2 = B_{11} + A_{12} = \begin{bmatrix} 4 & 3 \\ -1 & 2 \\ 1 & -5 \end{bmatrix} + \begin{bmatrix} 2 & -1 \\ 1 & 3 \\ 4 & 0 \end{bmatrix} = \begin{bmatrix} 6 & 2 \\ 0 & 5 \\ 5 & -5 \end{bmatrix}$$

(Quando alguns dos blocos são matrizes nulas, ou matrizes identidade, como no nosso caso, esses cálculos podem ser feitos muito rapidamente.) Os cálculos para os outros cinco blocos de $AB$ são semelhantes. Verifique que o resultado é

$$\left[\begin{array}{cc:cc:c} 6 & 2 & 1 & 2 & 2 \\ 0 & 5 & 2 & 1 & 12 \\ 5 & -5 & 3 & 3 & 9 \\ \hdashline 1 & 7 & 0 & 0 & 23 \\ 7 & 2 & 0 & 0 & 20 \end{array}\right]$$

(Observe que o bloco do lado superior esquerdo é o resultado do nosso último cálculo.) Verifique que você obtém a mesma resposta multiplicando $A$ por $B$ da forma usual.

## *Potências de Matriz*

Quando $A$ e $B$ forem duas matrizes $n \times n$, o produto delas também será uma matriz $n \times n$. Um caso especial ocorre quando $A = B$. Faz sentido definir $A^2 = AA$ e, em geral, definir $A^k$ como

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

sendo $k$ um inteiro positivo. Assim, $A^1 = A$, e é conveniente definir $A^0 = I_n$.

Antes de fazer muitas suposições, precisamos nos perguntar com que extensão as potências de matrizes se comportam como as potências dos números reais. As propriedades a seguir originam-se imediatamente das definições que acabamos de dar e são matrizes análogas das propriedades correspondentes para potências de números reais.

Se $A$ é uma matriz quadrada e $r$ e $s$ são inteiros não negativos, então
1. $A^rA^s = A^{r+s}$
2. $(A^r)^s = A^{rs}$

Na Seção 3.4, estenderemos a definição e as propriedades para incluir potências com números inteiros negativos.

**EXEMPLO 13** (a) Se $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, então

$$A^2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}, \quad A^3 = A^2A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix}$$

e, em geral,

$$A^n = \begin{bmatrix} 2^{n-1} & 2^{n-1} \\ 2^{n-1} & 2^{n-1} \end{bmatrix} \text{ para todo } n \geq 1$$

Essa afirmação pode ser provada por indução matemática, já que ela é dada para uma coleção *infinita* de afirmações, uma para cada número natural $n$. (O Apêndice B mostra uma breve revisão de indução matemática.) O primeiro passo é provar que a fórmula vale para $n = 1$. Nesse caso,

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

como queríamos.

A hipótese de indução é assumir que

$$A^k = \begin{bmatrix} 2^{k-1} & 2^{k-1} \\ 2^{k-1} & 2^{k-1} \end{bmatrix}$$

para algum número inteiro $k \geq 1$. O passo de indução é para provar que a fórmula vale para $n = k + 1$. Usando a definição de potência de matriz e a hipótese de indução, calculamos

$$\begin{aligned} A^{k+1} = A^k A &= \begin{bmatrix} 2^{k-1} & 2^{k-1} \\ 2^{k-1} & 2^{k-1} \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2^{k-1} + 2^{k-1} & 2^{k-1} + 2^{k-1} \\ 2^{k-1} + 2^{k-1} & 2^{k-1} + 2^{k-1} \end{bmatrix} \\ &= \begin{bmatrix} 2^k & 2^k \\ 2^k & 2^k \end{bmatrix} \\ &= \begin{bmatrix} 2^{(k+1)-1} & 2^{(k+1)-1} \\ 2^{(k+1)-1} & 2^{(k+1)-1} \end{bmatrix} \end{aligned}$$

Assim, a fórmula vale para todo $n \geq 1$, pelo princípio da indução matemática.

(b) Se $B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, então $B^2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. Continuando, encontramos

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

e

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

Assim, $B^5 = B$, e a seqüência de potências de $B$ se repete em um ciclo de quatro:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \ldots$$

## *A Transposta de uma Matriz*

Até agora, todas as operações com matrizes que definimos são análogas a operações com números reais, embora elas nem sempre se comportem da mesma maneira. A próxima operação não tem essa analogia.

A ***transposta*** de uma matriz $A$ $m \times n$ é uma matriz $n \times m$ $A^T$ obtida pela permutação das linhas de $A$ pelas suas colunas — isto é, a $i$-ésima coluna de $A^T$ é a $i$-ésima linha de $A$ para todo $i$.

**EXEMPLO 14** Sejam

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 5 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{e} \quad C = [5 \quad -1 \quad 2]$$

Então, suas transpostas são

$$A^T = \begin{bmatrix} 1 & 5 \\ 3 & 0 \\ 2 & 1 \end{bmatrix}, \quad B^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \quad \text{e} \quad C^T = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$$

A transposta é algumas vezes usada como uma definição alternativa para o produto escalar de dois vetores, em termos de multiplicação de matriz. Se

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad \text{e} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

então

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= u_1v_1 + u_2v_2 + \cdots + u_nv_n \\ &= [u_1 \quad u_2 \quad \cdots \quad u_n] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \\ &= \mathbf{u}^T\mathbf{v} \end{aligned}$$

Uma definição alternativa útil para a transposta é dada elemento a elemento:

$$(A^T)_{ij} = A_{ji} \quad \text{para todos } i \text{ e } j$$

Em palavras, o elemento na linha $i$ e coluna $j$ de $A^T$ é igual ao elemento na linha $j$ e coluna $i$ de $A$.

A transposta é também usada para definir um tipo muito importante de matriz quadrada: a matriz simétrica. Uma matriz é ***simétrica*** quando é igual à sua transposta.

**EXEMPLO 15** Sejam

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

Então, $A$ é simétrica, pois $A^T = A$; entretanto, $B$ não é simétrica, já que $B^T = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} \neq B$.

Uma matriz simétrica tem a propriedade de ser a sua própria "imagem refletida" através da diagonal principal. A Figura 1 ilustra essa propriedade para uma matriz 3×3. As formas geométricas correspondentes representam elementos iguais; os elementos da diagonal (representados em linha tracejada) são arbitrários.

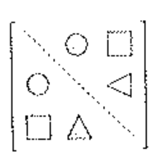


**Figura 1** Uma matriz simétrica

Uma definição de matriz simétrica elemento a elemento também é útil. Ela é simplesmente a descrição algébrica da propriedade da "reflexão".

Uma matriz quadrada $A$ é simétrica se, e somente se, $A_{ij} = A_{ji}$ para todos $i$ e $j$.

## ◈ EXERCÍCIOS 3.2 ◈

*Sejam*

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

$$D = \begin{bmatrix} 0 & -3 \\ -2 & 1 \end{bmatrix}, E = [4 \quad 2], F = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

*Nos Exercícios de 1 a 16, calcule a matriz indicada (se possível).*

**1.** $A + 2D$

**2.** $3D - 2A$

**3.** $B - C$

**4.** $B - C^T$

**5.** $AB$

**6.** $BD$

**7.** $D + BC$

**8.** $B^TB$

**9.** $E(AF)$

**10.** $F(DF)$

**11.** $FE$

**12.** $EF$

**13.** $B^TC^T - (CB)^T$

**14.** $DA - AD$

**15.** $A^3$

**16.** $(I_2 - D)^2$

**17.** Dê um exemplo de uma matriz $A$ 2×2 não nula tal que $A^2 = O$.

**18.** Seja $A = \begin{bmatrix} 2 & 1 \\ 6 & 3 \end{bmatrix}$. Encontre as matrizes 2×2 $B$ e $C$ tais que $AB = AC$, mas $B \neq C$.

**19.** Uma fábrica produz três produtos (banheiras, pias e tanques) e os envia para armazenamento em dois depósitos. O número de unidades enviadas de cada produto para cada depósito é dado pela matriz

$$A = \begin{bmatrix} 200 & 75 \\ 150 & 100 \\ 100 & 125 \end{bmatrix}$$

(em que $a_{ij}$ é o número de unidades enviadas do produto $i$ para o depósito $j$, e os produtos são colocados em ordem alfabética). O custo de remessa de uma unidade de cada produto, por caminhão, é: \$ 1,50 por banheira, \$ 1,00 por pia e \$ 2,00 por tanque. Os custos unitários correspondentes ao envio por trem são: \$ 1,75, \$ 1,50 e \$ 1,00. Organize esses custos em uma matriz $B$ e use essa matriz para mostrar como a fábrica pode comparar os custos de remessa — por caminhão e por trem — de seus produtos para cada um dos dois depósitos.

**20.** Em relação ao Exercício 19, suponha que o custo unitário de distribuição dos produtos para as lojas seja o mesmo para todos os produtos, mas que varie dependendo do depósito por causa das distâncias envolvidas. Custa \$ 0,75 para distribuir uma unidade do depósito 1 e \$ 1,00 para distribuir uma unidade do depósito 2. Organize esses custos em uma matriz $C$ e use multiplicação de matrizes para calcular o custo total de distribuição de cada produto.

*Nos Exercícios 21 e 22, escreva o sistema de equações lineares dado na forma de equação matricial, $A\mathbf{x} = \mathbf{b}$.*

**21.**
$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 0 \\ 2x_1 + x_2 - 5x_3 &= 4 \end{aligned}$$

**22.**
$$\begin{aligned} -x_1 \qquad + 2x_3 &= 1 \\ x_1 - x_2 \qquad &= -2 \\ x_2 + x_3 &= -1 \end{aligned}$$

*Nos Exercícios de 23 a 28, sejam*

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

*e*

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

**23.** Use a representação matriz-coluna do produto para escrever cada coluna de $AB$ como combinação linear das colunas de $A$.

**24.** Use a representação matriz-linha do produto para escrever cada linha de $AB$ como combinação linear das linhas de $B$.

**25.** Calcule o produto externo da expansão de $AB$.

**26.** Use a representação matriz-coluna do produto para escrever cada coluna de $BA$ como combinação linear das colunas de $B$.

**27.** Use a representação matriz-linha do produto para escrever cada linha de $BA$ como combinação linear das linhas de $A$.

**28.** Calcule o produto externo da expansão de $BA$.

*Nos Exercícios 29 e 30, assuma que o produto $AB$ faz sentido.*

**29.** Prove que, se as colunas de $B$ são linearmente independentes, as colunas de $AB$ também o são.

**30.** Prove que, se as linhas de $A$ são linearmente independentes, as linhas de $AB$ também o são.

*Nos Exercícios de 31 a 34, calcule $AB$ usando multiplicação por blocos com a partição indicada.*

**31.** $A = \left[\begin{array}{cc|cc} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 2 & 3 \end{array}\right]$, $B = \left[\begin{array}{cc|c} 2 & 3 & 0 \\ -1 & 1 & 0 \\ \hline 0 & 0 & 1 \\ 0 & 0 & 1 \end{array}\right]$

**32.** $A = \left[\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 4 & 5 & 0 & 1 \end{array}\right]$, $B = \left[\begin{array}{c|cc} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 1 & 5 & 4 \\ -2 & 3 & 2 \end{array}\right]$

**33.** $A = \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \\ \hline 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & -1 \end{array}\right]$, $B = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{array}\right]$

**34.** $A = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \\ \hline 0 & 0 & 0 & 4 \end{array}\right]$, $B = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 1 & 1 \\ \hline 1 & 1 & 1 & -1 \end{array}\right]$

**35.** Seja $A = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$.

(a) Calcule $A^2, A^3, \ldots, A^7$.

(b) O que é $A^{2001}$? Por quê?

**36.** Seja $B = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$. Encontre $B^{2001}$. Justifique.

**37.** Seja $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Encontre uma fórmula para $A^n$ $(n \geq 1)$ e prove a sua fórmula usando indução matemática.

**38.** Seja $A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$.

(a) Mostre que $A^2 = A^2 = \begin{bmatrix} \cos 2\theta & -\operatorname{sen} 2\theta \\ \operatorname{sen} 2\theta & \cos 2\theta \end{bmatrix}$.

(b) Prove, usando indução matemática, que $A^n =$

$\begin{bmatrix} \cos n\theta & -\operatorname{sen} n\theta \\ \operatorname{sen} n\theta & \cos n\theta \end{bmatrix}$ para $n \geq 1$.

**39.** Em cada um dos itens a seguir, ache a matriz $4\times 4$ $A = [a_{ij}]$ que satisfaz a condição dada:

(a) $a_{ij} = (-1)^{i+j}$

(b) $a_{ij} = j - i$

(c) $a_{ij} = (i-1)^j$

(d) $a_{ij} = \operatorname{sen}\left(\dfrac{(i+j-1)\pi}{4}\right)$

**40.** Em cada um dos itens a seguir, ache a matriz $6\times 6$ $A = [a_{ij}]$ que satisfaz a condição dada:

(a) $a_{ij} = \begin{cases} i+j & \text{se} \quad i \leq j \\ 0 & \text{se} \quad i > j \end{cases}$

(b) $a_{ij} = \begin{cases} 1 & \text{se } |i-j| \leq 1 \\ 0 & \text{se } |i-j| > 1 \end{cases}$

(c) $a_{ij} = \begin{cases} 1 & \text{if } 6 \leq i+j \leq 8 \\ 0 & \text{nos outros casos} \end{cases}$

**41.** Prove o Teorema 1(a).

## 3.3 A Álgebra de Matrizes

Em certos aspectos, a aritmética das matrizes generaliza a aritmética dos vetores. Não esperamos surpresas com respeito à adição e à multiplicação por escalar de matrizes, e de fato elas não existem. Isso nos permitirá estender para matrizes vários conceitos com os quais já temos familiaridade, pelo nosso trabalho com os vetores. Em particular, combinações lineares, conjuntos geradores e independência linear serão definidos sem nenhuma dificuldade para matrizes.

Matrizes, entretanto, admitem outras operações — a multiplicação, por exemplo — que vetores não têm. Não devemos esperar que a multiplicação de matrizes se comporte como a multiplicação dos números reais, a menos que possamos provar isso; de fato, ela não se comporta como tal. Nesta seção, resumimos e provamos algumas das propriedades principais das operações com matrizes e começamos a desenvolver uma álgebra de matrizes.

## *Propriedades da Adição e da Multiplicação por Escalar*

Todas as propriedades algébricas da adição e da multiplicação por escalar dos vetores (Teorema 1 da Seção 1.2) valem para matrizes. Resumimos essas propriedades no próximo teorema.

**TEOREMA 1 Propriedades Algébricas da Adição de Matrizes e da Multiplicação por Escalar**

Sejam $A$, $B$ e $C$ matrizes de mesma ordem, e $c$ e $d$, escalares. Então:

a. $A + B = B + A$ — Comutativa
b. $(A + B) + C = A + (B + C)$ — Associativa
c. $A + O = A$
d. $A + (-A) = O$
e. $c(A + B) = cA + cB$ — Distributiva
f. $(c + d)A = cA + dA$ — Distributiva
g. $c(dA) = (cd)A$
h. $1A = A$

As demonstrações dessas propriedades são análogas às demonstrações correspondentes das propriedades dos vetores. Da mesma forma, os comentários que seguem o Teorema 1 da Seção 1.2 são igualmente válidos aqui, e não deveremos ter dificuldades para usar essas propriedades para fazer manipulações algébricas com matrizes. (Consulte o Exemplo 5 da Seção 1.2 e veja os Exercícios 17 e 18 no final desta seção.)

A propriedade associativa nos permite combinar a multiplicação por escalar e a adição sem fazer uso de parênteses. Se $A$, $B$ e $C$ são matrizes de mesma ordem, então

$$(2A + 3B) - C = 2A + (3B - C)$$

e assim podemos simplesmente escrever $2A + 3B - C$. Mais geralmente, se $A_1, A_2, \ldots, A_k$ são matrizes de mesma ordem e $c_1, c_2, \ldots, c_k$ são escalares, fazemos a ***combinação linear***

$$c_1A_1 + c_2A_2 + \ldots + c_kA_k$$

Chamaremos $c_1, c_2, \ldots, c_k$ de ***coeficientes*** da combinação linear. Vamos agora elaborar e responder questões sobre combinações lineares de matrizes.

**EXEMPLO 1** Sejam $A_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, $A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ e $A_3 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

(a) $B = \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix}$ é combinação linear de $A_1$, $A_2$ e $A_3$?

(b) $C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ é combinação linear de $A_1$, $A_2$ e $A_3$?

SOLUÇÃO:

(a) Queremos encontrar escalares $c_1$, $c_2$ e $c_3$ tais que $c_1A_1 + c_2A_2 + c_3A_3 = B$. Assim,

$$c_1\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix}$$

O lado esquerdo dessa equação pode ser escrito na forma:

$$\begin{bmatrix} c_2 + c_3 & c_1 + c_3 \\ -c_1 + c_3 & c_2 + c_3 \end{bmatrix}$$

Comparando os elementos e usando a definição de igualdade de matrizes, temos as quatro equações lineares

$$\begin{aligned} c_2 + c_3 &= 1 \\ c_1 \qquad + c_3 &= 4 \\ -c_1 \qquad + c_3 &= 2 \\ c_2 + c_3 &= 1 \end{aligned}$$

A eliminação de Gauss-Jordan nos dá

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 4 \\ -1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

(verifique isso!), assim, $c_1 = 1$, $c_2 = -2$ e $c_3 = 3$. Portanto, $A_1 - 2A_2 + 3A_3 = B$, o que pode ser facilmente verificado.

(b) Desta vez queremos resolver

$$c_1\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

Procedendo como na parte (a), obtemos o sistema linear

$$\begin{aligned} c_2 + c_3 &= 1 \\ c_1 \qquad + c_3 &= 2 \\ -c_1 \qquad + c_3 &= 3 \\ c_2 + c_3 &= 4 \end{aligned}$$

O escalonamento por linhas nos dá

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ -1 & 0 & 1 & 3 \\ 0 & 1 & 1 & 4 \end{array}\right] \xrightarrow{L_4 - L_1} \left[\begin{array}{ccc|c} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ -1 & 0 & 1 & 3 \\ 0 & 0 & 0 & 3 \end{array}\right]$$

Não precisamos continuar: a última linha implica que o sistema não tem solução. Portanto, nesse caso, $C$ não é combinação linear de $A_1$, $A_2$ e $A_3$.

Observação: ◆ Note que as colunas da matriz completa contêm os elementos das matrizes que nos foram dadas. Se lermos os elementos de cada uma das matrizes da esquerda para a direita e de cima para baixo, obtemos a ordem em que eles aparecem na matriz completa. Por exemplo, lemos $A_1$ como "0, 1, −1, 0", o que corresponde à primeira coluna da matriz completa. É como se tivéssemos "esticado" a matriz dada formando um vetor coluna. Dessa forma, teríamos obtido exatamente o mesmo sistema de equações lineares da parte (a) com esta pergunta:

$$\text{O vetor } \begin{bmatrix} 1 \\ 4 \\ 2 \\ 1 \end{bmatrix} \text{ é combinação linear de } \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \text{ e } \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}?$$

Encontraremos tais paralelos repetidamente daqui por diante. No Capítulo 6, eles serão explorados com mais detalhes.

Podemos definir o ***conjunto gerado*** por um conjunto de matrizes como o conjunto de todas as combinações lineares dessas matrizes.

**EXEMPLO 2** Descreva o conjunto gerado pelas matrizes $A_1$, $A_2$ e $A_3$ do Exemplo 1.

SOLUÇÃO: Um maneira de fazer isso é simplesmente escrever uma combinação linear genérica de $A_1$, $A_2$ e $A_3$. Assim,

$$\begin{aligned} c_1A_1 + c_2A_2 + c_3A_3 &= c_1\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} c_2 + c_3 & c_1 + c_3 \\ -c_1 + c_3 & c_2 + c_3 \end{bmatrix} \end{aligned}$$

(que é análogo à representação paramétrica de um plano). Porém, suponha que queiramos saber quando a matriz $\begin{bmatrix} w & x \\ y & z \end{bmatrix}$ estará no ger $(A_1, A_2, A_3)$. Da representação que acabamos de ver, sabemos que esse fato se dará quando

$$\begin{bmatrix} c_2 + c_3 & c_1 + c_3 \\ -c_1 + c_3 & c_2 + c_3 \end{bmatrix} = \begin{bmatrix} w & x \\ y & z \end{bmatrix}$$

para alguma escolha de escalares $c_1, c_2, c_3$. Isso dará surgimento a um sistema de equações lineares cujo lado esquerdo será exatamente o mesmo do Exemplo 1, mas cujo lado direito será genérico. A matriz completa desse sistema será

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & w \\ 1 & 0 & 1 & x \\ -1 & 0 & 1 & y \\ 0 & 1 & 1 & z \end{array}\right]$$

e o escalonamento por linhas produzirá

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & w \\ 1 & 0 & 1 & x \\ -1 & 0 & 1 & y \\ 0 & 1 & 1 & z \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{1}{2}x - \frac{1}{2}y \\ 0 & 1 & 0 & -\frac{1}{2}x - \frac{1}{2}y + w \\ 0 & 0 & 1 & \frac{1}{2}x + \frac{1}{2}y \\ 0 & 0 & 0 & w - z \end{array}\right]$$

(Verifique isso cuidadosamente.) A única restrição se refere à última linha, onde claramente devemos ter $w - z = 0$ para que haja solução. Portanto, o conjunto gerado por $A_1$, $A_2$ e $A_3$ consiste em todas as matrizes $\begin{bmatrix} w & x \\ y & z \end{bmatrix}$ para as quais $w = z$. Ou seja, o ger$(A_1, A_2, A_3) = \left\{\begin{bmatrix} w & x \\ y & w \end{bmatrix}\right\}$.

Nota: Se tivéssemos observado isso *antes*, olhando o Exemplo 1, teríamos visto imediatamente que $B = \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix}$ é uma combinação linear de $A_1$, $A_2$ e $A_3$, pois $B$ tem a forma necessária (tome $w = 1, x = 4$ e $y = 2$), mas $C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ não pode ser uma combinação linear de $A_1$, $A_2$ e $A_3$, pois não tem a forma apropriada ($1 \neq 4$).

A independência linear também faz sentido para as matrizes. Dizemos que matrizes $A_1, A_2, \ldots, A_k$ de mesma ordem são ***linearmente independentes*** se a única solução da equação

$$c_1A_1 + c_2A_2 + \ldots + c_kA_k = O \tag{1}$$

é a solução trivial: $c_1 = c_2 = \ldots = c_k = 0$. Se existirem coeficientes não triviais que satisfaçam (1), então $A_1, A_2, \ldots, A_k$ serão chamadas de ***linearmente dependentes***.

**EXEMPLO 3** Determine se as matrizes $A_1$, $A_2$ e $A_3$ do Exemplo 1 são linearmente independentes.

SOLUÇÃO: Queremos resolver a equação $c_1A_1 + c_2A_2 + c_3A_3 = O$. Escrevendo as matrizes, temos

$$c_1\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Desta vez, obtemos um sistema linear *homogêneo* cujo lado esquerdo é o mesmo que o dos exemplos 1 e 2. (Você já está conseguindo vislumbrar um padrão?) A matriz completa escalonada por linhas dá

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Portanto, $c_1 = c_2 = c_3 = 0$, e concluímos que as matrizes $A_1$, $A_2$ e $A_3$ são linearmente independentes.

## *Propriedades da Multiplicação de Matrizes*

Sempre que encontramos uma operação nova, tal como a multiplicação de matrizes, devemos ser cuidadosos para não presumir muito sobre ela. Seria bom se a multiplicação de matrizes se comportasse como a multiplicação de números reais. Embora em muitos aspectos isso aconteça, existem algumas diferenças significativas.

**EXEMPLO 4** Considere as matrizes

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

Multiplicando-as, temos

$$AB = \begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 4 \\ -3 & -2 \end{bmatrix} \quad \text{e} \quad BA = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$$

Portanto, $AB \neq BA$. Assim, em contraste com a multiplicação de números reais, a multiplicação de matrizes *não é comutativa* — a ordem dos fatores no produto é importante!

É fácil verificar que $A^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ (faça-o!). Assim, para matrizes, a equação $A^2 = O$ não implica que $A = O$ (diferentemente da situação para números reais, em que a equação $x^2 = 0$ tem somente $x = 0$ como solução).

Apesar de coisas obscuras poderem aparecer depois do último exemplo, a situação não é tão ruim — você precisa somente acostumar-se a trabalhar com as matrizes e lembrar-se constantemente de que elas não são números. O próximo teorema resume as propriedades principais de multiplicação de matrizes.

### ◈ TEOREMA 2 Propriedades da Multiplicação de Matrizes

Sejam $A$, $B$ e $C$ matrizes (cujas ordens possibilitem que as operações indicadas sejam realizadas) e $k$ um escalar. Então:

a. $A(BC) = (AB)C$ — Associativa
b. $A(B + C) = AB + AC$ — Distributiva à esquerda
c. $(A + B)C = AC + BC$ — Distributiva à direita
d. $k(AB) = (kA)B = A(kB)$
e. $I_m A = A = AI_n$ se $A$ for $m \times n$ — Identidade da multiplicação

DEMONSTRAÇÃO: Provamos (b) e encaminhamos a prova de (e). Deixamos a prova da propriedade (a) para a Seção 3.6. As demais propriedades são consideradas nos exercícios.
(b) Para provar que $A(B + C) = AB + AC$, denotaremos as linhas de $A$ por $\mathbf{A}_i$ e as colunas de $B$ e $C$ por $\mathbf{b}_j$ e $\mathbf{c}_j$, respectivamente. Então, a $j$-ésima coluna de $B + C$ é $\mathbf{b}_j + \mathbf{c}_j$ (já que a adição é feita elemento a elemento), e assim

$$\begin{aligned}[A(B + C)]_{ij} &= \mathbf{A}_i \cdot (\mathbf{b}_j + \mathbf{c}_j) \\ &= \mathbf{A}_i \cdot \mathbf{b}_j + \mathbf{A}_i \cdot \mathbf{c}_j \\ &= (AB)_{ij} + (AC)_{ij} \\ &= (AB + AC)_{ij}\end{aligned}$$

Como isso é verdadeiro para todo $i$ e $j$, devemos ter $A(B + C) = AB + AC$.
(e) Para provar que $AI_n = A$, observamos que a matriz identidade $I_n$ pode ser decomposta em colunas por

$$I_n = [\mathbf{e}_1 \mid \mathbf{e}_2 \mid \cdots \mid \mathbf{e}_n]$$

em que $\mathbf{e}_i$ é um vetor canônico unitário. Portanto,

$$\begin{aligned}AI_n &= [A\mathbf{e}_1 \mid A\mathbf{e}_2 \mid \cdots \mid A\mathbf{e}_n] \\ &= [\mathbf{a}_1 \mid \mathbf{a}_2 \mid \cdots \mid \mathbf{a}_n] \\ &= A\end{aligned}$$

pelo Teorema 1(b) da Seção 3.2. ◈

Podemos usar essas propriedades para explorar mais profundamente o quanto a multiplicação de matrizes se assemelha à multiplicação de números reais.

**EXEMPLO 5** Se $A$ e $B$ são matrizes quadradas de mesma ordem, é verdade que $(A + B)^2 = A^2 + 2AB + B^2$?

SOLUÇÃO: Usando as propriedades de multiplicação de matrizes, calculamos

$$\begin{aligned}(A + B)^2 &= (A + B)(A + B) \\ &= (A + B)A + (A + B)B && \text{Distributiva à esquerda} \\ &= A^2 + AB + BA + B^2 && \text{Distributiva à direita}\end{aligned}$$

Portanto, $(A + B)^2 = A^2 + 2AB + B^2$ se, e somente se, $A^2 + BA + AB + B^2 = A^2 + 2AB + B^2$. Subtraindo $A^2$ e $B^2$ dos dois lados, temos $BA = AB = 2AB$. Subtraindo $AB$ dos dois lados temos $BA = AB$. Assim, $(A + B)^2 = A^2 + 2AB + B^2$ se, e somente se, $A$ e $B$ se comutam. (Você consegue achar um exemplo de um par de matrizes que satisfaça essa condição? Você pode dar um exemplo de um par de matrizes que não satisfaça essa propriedade?)

## *Propriedades da Transposta*

**TEOREMA 3 Propriedades da Transposta**

Sejam $A$ e $B$ matrizes (cujas ordens são tais que as operações indicadas podem ser realizadas) e $k$ um escalar. Então:

a. $(A^T)^T = A$

b. $(A + B)^T = A^T + B^T$

c. $(kA)^T = k(A^T)$

d. $(AB)^T = B^T A^T$

e. $(A^r)^T = (A^T)^r$ para todos os inteiros não negativos $r$

DEMONSTRAÇÃO: É relativamente fácil provar as propriedades (a) — (c) (veja o Exercício 30). Provar a propriedade (e) é um bom exercício de indução matemática (veja o Exercício 31). Provaremos (d), já que essa é uma propriedade que você não deve ter esperado. [Você achou que $(AB)^T = A^T B^T$ poderia ser verdadeira?]

Primeiro, se $A$ for $m \times n$ e $B$ for $n \times r$, $B^T$ é $r \times n$, e $A^T$, $n \times m$. Portanto, o produto $B^T A^T$ está definido e é $r \times m$. Como $AB$ é $m \times r$, $(AB)^T$ é $r \times m$ e, assim, $(AB)^T$ e $B^T A^T$ têm a mesma ordem. Devemos provar que os elementos correspondentes são iguais.

Denotamos a $i$-ésima linha de uma matriz $X$ por $\text{lin}_i(X)$, e sua $j$-ésima coluna, por $\text{col}_j(X)$. Usando essas convenções, vemos que

$$\begin{aligned}
[(AB)^T]_{ij} &= (AB)_{ji} \\
&= \text{lin}_j(A) \cdot \text{col}_i(B) \\
&= \text{col}_j(A^T) \cdot \text{lin}_i(B^T) \\
&= \text{lin}_i(B^T) \cdot \text{col}_j(A^T) = [B^T A^T]_{ij}
\end{aligned}$$

(Note que usamos a definição de multiplicação de matriz, a definição de transposta e o fato de o produto escalar ser comutativo.) Como $i$ e $j$ são arbitrários, concluímos que $(AB)^T = B^T A^T$. ◈

Observação:

- As propriedades (b) e (d) do Teorema 3 podem ser generalizadas para somas e produtos de um número finito de matrizes:

$$(A_1 + A_2 + \cdots + A_k)^T = A_1^T + A_2^T + \cdots + A_k^T \quad \text{e} \quad (A_1 A_2 \cdots A_k)^T = A_k^T \cdots A_2^T A_1^T$$

assumindo que as ordens das matrizes são tais que as operações indicadas podem ser realizadas. Fica para você provar esses fatos usando indução matemática nos Exercícios 32 e 33.

**EXEMPLO 6** Sejam

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

Então, $A^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$, por isso $A + A^T = \begin{bmatrix} 2 & 5 \\ 5 & 8 \end{bmatrix}$, uma matriz simétrica.

Temos

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

assim,

$$BB^T = \begin{bmatrix} 4 & -1 & 0 \\ 2 & 3 & 1 \end{bmatrix}\begin{bmatrix} 4 & 2 \\ -1 & 3 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 5 \\ 5 & 14 \end{bmatrix}$$

e

$$B^TB = \begin{bmatrix} 4 & 2 \\ -1 & 3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 4 & -1 & 0 \\ 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 20 & 2 & 2 \\ 2 & 10 & 3 \\ 2 & 3 & 1 \end{bmatrix}$$

Portanto, $BB^T$ e $B^TB$ são simétricas, embora $B$ nem mesmo seja quadrada! (Verifique que $AA^T$ e $A^TA$ são também simétricas.) ◈

O próximo teorema diz que os resultados do Exemplo 6 são em geral verdadeiros.

### ◈ TEOREMA 4

a. Se $A$ é uma matriz quadrada, então $A + A^T$ é uma matriz simétrica.
b. Para toda matriz $A$, $AA^T$ e $A^TA$ são matrizes simétricas.

DEMONSTRAÇÃO: Provamos (a) e deixamos para provar (b) no Exercício 34. Simplesmente observamos que

$$(A + A^T)^T = A^T + (A^T)^T = A^T + A = A + A^T$$

(usando propriedades da transposta e a comutativa da adição de matrizes). Assim, $A + A^T$ é igual à sua transposta, por isso, por definição, é simétrica. ◈

## ◈ EXERCÍCIOS 3.3 ◈

*Nos Exercícios de 1 a 4, ache X, dadas* $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ *e* $B = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}$.

**1.** $X - 2A + 3B = O$

**2.** $2X = A - B$

**3.** $2(A + 2B) = 3X$

**4.** $2(A - B + X) = 3(X - A)$

*Nos Exercícios de 5 a 8, escreva B como combinação linear das outras matrizes, se possível.*

**5.** $B = \begin{bmatrix} 2 & 5 \\ 0 & 3 \end{bmatrix}, A_1 = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$

**6.** $B = \begin{bmatrix} 2 & 3 \\ -4 & 2 \end{bmatrix}, A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$
$A_3 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

**7.** $B = \begin{bmatrix} 3 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}, A_1 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix},$
$A_2 = \begin{bmatrix} -1 & 2 & 0 \\ 0 & 1 & 0 \end{bmatrix}, A_3 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

**8.** $B = \begin{bmatrix} 2 & -2 & 3 \\ 0 & 0 & -2 \\ 0 & 0 & 2 \end{bmatrix}, A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$
$A_2 = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, A_3 = \begin{bmatrix} -1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix},$
$A_4 = \begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$

*Nos Exercícios de 9 a 12, encontre a expressão geral do conjunto gerado pelas matrizes dadas, como no Exemplo 2.*

9. ger($A_1, A_2$) do Exercício 5

10. ger($A_1, A_2, A_3$) do Exercício 6

11. ger($A_1, A_2, A_3$) do Exercício 7

12. ger($A_1, A_2, A_3, A_4$) do Exercício 8

*Nos Exercícios de 13 a 16, determine se as matrizes dadas são linearmente independentes.*

13. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$

14. $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$

15. $\begin{bmatrix} 0 & 1 \\ 5 & 2 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 2 & 3 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} -2 & -1 \\ 0 & 1 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} -1 & -3 \\ 1 & 9 \\ 4 & 5 \end{bmatrix}$

16. $\begin{bmatrix} 1 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & 2 & 6 \end{bmatrix}, \begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 4 & 9 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 5 \end{bmatrix},$
$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -4 \end{bmatrix}$

17. Prove o Teorema 1(a) – (d).

18. Prove o Teorema 1(e) – (h).

19. Prove o Teorema 2(c).

20. Prove o Teorema 2(d).

21. Prove a parte do Teorema 2(e) que não foi provada no texto.

22. Prove que, para matrizes quadradas $A$ e $B$, $AB = BA$ se, e somente se, $(A - B)(A + B) = A^2 - B^2$.

*Nos Exercícios de 23 a 25, se* $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, *encontre condições em a, b, c e d para que* $AB = BA$.

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

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

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

26. Encontre condições em $a$, $b$, $c$ e $d$ para que $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ comute com $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ e $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$.

27. Encontre condições em $a$, $b$, $c$ e $d$ para que $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ comute com todas as matrizes $2\times2$.

28. Prove que, se $AB$ e $BA$ estiverem definidas, $AB$ e $BA$ são matrizes quadradas.

*Uma matriz quadrada é chamada de **triangular superior** quando todos os elementos abaixo da diagonal principal são zeros. Assim, a forma de uma matriz triangular superior é*

$$\begin{bmatrix} * & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & & * & * \\ 0 & 0 & \cdots & 0 & * \end{bmatrix}$$

*em que os elementos marcadas por * são arbitrários. Uma definição mais formal dessa matriz* $A = [a_{ij}]$ é que $a_{ij} = 0$ *se* $i > j$.

29. Prove que o produto de duas matrizes triangulares superiores $n\times n$ é uma matriz triangular superior.

30. Prove o Teorema 3(a) – (c).

31. Prove o Teorema 3(e).

32. Usando indução, prove que $(A_1 + A_2 + \cdots + A_n)^T = A_1^T + A_2^T + \cdots + A_n^T$ para todo $n \geq 1$.

33. Usando indução, prove que $(A_1 A_2 \cdots A_n)^T = A_n^T \cdots A_2^T A_n^T$ para todo $n \geq 1$.

34. Prove o Teorema 4(b).

35. (a) Prove que, se $A$ e $B$ são matrizes $n\times n$ simétricas, $A + B$ também é.

(b) Prove que, se $A$ é uma matriz $n\times n$ simétrica, $kA$ também é para todo escalar $k$.

36. (a) Dê um exemplo para mostrar que, se $A$ e $B$ são matrizes $n\times n$ simétricas, $AB$ não é necessariamente simétrica.

(b) Prove que, se $A$ e $B$ são matrizes $n\times n$ simétricas, $AB$ é simétrica se, e somente se, $AB = BA$.

*Uma matriz quadrada é chamada de **anti-simétrica** se* $A^T = -A$.

37. Quais das seguintes matrizes são anti-simétricas?

(a) $\begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 3 & -1 \\ -3 & 0 & 2 \\ 1 & -2 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 5 \\ 2 & 5 & 0 \end{bmatrix}$

38. Dê uma definição de matriz anti-simétrica elemento a elemento.

39. Prove que a diagonal principal de uma matriz anti-simétrica é formada inteiramente por zeros.

40. Prove que, se $A$ e $B$ são matrizes $n\times n$ anti-simétricas, $A + B$ também é.

41. Se $A$ e $B$ são matrizes $2\times2$ anti-simétricas, sob que condições $AB$ é anti-simétrica?

42. Prove que, se $A$ é uma matriz $n\times n$, $A - A^T$ é anti-simétrica.

**43.** (a) Prove que toda matriz quadrada $A$ pode ser escrita como a soma de uma matriz simétrica com uma matriz anti-simétrica. (*Sugestão:* use o Teorema 4 e o Exercício 42).

(b) Ilustre o item usando a matriz $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$.

*O **traço** de uma matriz $n \times n$ $A = [a_{ij}]$ é a soma dos elementos da sua diagonal principal e é denotado por* $\text{tr}(A)$:

$$\text{tr}(A) = a_{11} + a_{22} + \ldots + a_{nn}$$

**44.** Se $A$ e $B$ são matrizes $n \times n$, prove as seguintes propriedades do traço:

(a) $\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B)$

(b) $\text{tr}(kA) = k\text{tr}(A)$, em que $k$ é um escalar

**45.** Prove que, se $A$ e $B$ são matrizes $n \times n$, $\text{tr}(AB) = \text{tr}(BA)$.

**46.** Se $A$ é uma matriz qualquer, o que é $\text{tr}(AA^T)$?

**47.** Mostre que não existem matrizes $2 \times 2$ $A$ e $B$ tais que $AB - BA = I_2$.

## 3.4 A Inversa de uma Matriz

Nesta seção, retornamos para a descrição matricial $A\mathbf{x} = \mathbf{b}$ de um sistema de equações lineares e apresentamos meios de usar a álgebra das matrizes para resolver o sistema. Para fazer uma analogia, considere a equação $ax = b$, em que $a$, $b$ e $x$ representam números reais e queremos resolver a equação em $x$. Rapidamente, vemos que $x = b/a$ é a solução, mas precisamos nos lembrar de que isso só é verdade se $a \neq 0$. Calculando mais lentamente, assumindo que $a \neq 0$, acharemos a solução nesta seqüência de passos:

$$ax = b \Rightarrow \frac{1}{a}(ax) = \frac{1}{a}(b) \Rightarrow \left(\frac{1}{a}(a)\right)x = \frac{b}{a} \Rightarrow 1 \cdot x = \frac{b}{a} \Rightarrow x = \frac{b}{a}$$

(Esse exemplo mostra quanto nossa cabeça faz e quantas propriedades da aritmética e da álgebra assumimos!)

Para imitar esse procedimento para a equação matricial $A\mathbf{x} = \mathbf{b}$, do que precisamos? Precisamos encontrar uma matriz $A'$ (a análoga a $1/a$) tal que $A'A = I$, a matriz identidade (análoga a 1). Se tal matriz existir (analogamente à imposição que a $\neq 0$), então podemos fazer a seqüência de cálculos:

$$A\mathbf{x} = \mathbf{b} \Rightarrow A'(A\mathbf{x}) = A\mathbf{b} \Rightarrow (A'A)\,\mathbf{x} = A\mathbf{b} \Rightarrow I\mathbf{x} = A'\mathbf{b} \Rightarrow \mathbf{x} = A'\mathbf{b}$$

(Como se justifica cada um desses passos?)

Nosso objetivo, nesta seção, é determinar precisamente quando podemos achar a matriz $A'$. Vamos insistir um pouco mais: não queremos somente que $A'A = I$, mas queremos também que $AA' = I$. Essa imposição força $A$ e $A'$ a serem matrizes quadradas. (Por quê?)

**Definição** Se $A$ é uma matriz $n \times n$, uma ***inversa*** de $A$ é uma matriz $n \times n$ $A'$ que satisfaz a seguinte propriedade:

$$AA' = I \quad \text{e} \quad A'A = I$$

em que $I = I_n$ é a matriz identidade $n \times n$. Se essa matriz $A'$ existir, $A$ será chamada de ***invertível***.

**EXEMPLO 1** Se $A = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$, então $A' = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}$ é a inversa de $A$, pois

$$AA' = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{e} \quad A'A = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

**EXEMPLO 2** Mostre que as seguintes matrizes não são invertíveis:

(a) $O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ (b) $B = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

SOLUÇÃO:
(a) É fácil ver que a matriz nula $O$ não tem inversa. Se tivesse, existiria uma matriz $O'$ tal que $OO' = I = O'O$. Entretanto, o produto da matriz nula por qualquer outra matriz é a matriz nula, e assim $OO'$ nunca será igual à matriz identidade $I$. (Observe que essa demonstração não faz referência à ordem das matrizes, por isso ela é verdadeira para matrizes $n \times n$ em geral.)

(b) Suponha que $B$ tenha uma inversa $B' = \begin{bmatrix} w & x \\ y & z \end{bmatrix}$. A equação $BB' = I$ nos dá

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

da qual obtemos as equações

$$\begin{aligned} w \quad\quad + 2y \quad\quad &= 1 \\ x \quad\quad + 2z &= 0 \\ 2w \quad\quad + 4y \quad\quad &= 0 \\ 2x \quad\quad + 4z &= 1 \end{aligned}$$

Subtraindo duas vezes a primeira equação da terceira, obtemos $0 = -2$, que obviamente é um absurdo, por isso não existe solução. (O escalonamento por linhas fornece o mesmo resultado, mas ele não é realmente necessário nesse caso.) Deduzimos que a matriz $B'$ não existe, ou seja, que $B$ não é invertível. (De fato, ela não tem nem mesmo uma inversa em um dos lados, imagine então nos dois!) ◈

Observações:

- Embora tenhamos visto que a multiplicação de matrizes não é em geral comutativa, $A'$ (se existir) deve satisfazer $A'A = AA'$.
- Os exemplos que acabamos de ver levantaram duas questões que iremos responder nesta seção:
  (1) Como podemos saber quando uma matriz tem uma inversa?
  (2) Se uma matriz tem uma inversa, como podemos encontrá-la?
- Não questionamos ainda a possibilidade de uma matriz $A$ ter mais de uma inversa. O próximo teorema nos assegura que isso não pode acontecer.

## ◈ TEOREMA 1

Se $A$ é uma matriz invertível, então a sua inversa é única.

DEMONSTRAÇÃO: Em matemática, um modo padrão de mostrar que existe apenas uma de alguma coisa é mostrar que não pode existir mais de uma. Assim, suponha que $A$ tenha duas inversas — digamos, $A'$ e $A''$. Então:

$$AA' = I = A'A \quad \text{e} \quad AA'' = I = A''A$$

Assim,

$$A' = A'I = A'(AA'') = (A'A)A'' = IA'' = A''$$

Portanto, $A' = A''$, e a inversa é única. ◈

Graças a esse teorema, podemos agora nos referir *à* inversa de uma matriz invertível. De agora em diante, quando $A$ for invertível, denotaremos a sua inversa (única) por $A^{-1}$ (chamada "inversa de $A$").

Aviso: Não pense em escrever $A^{-1} = \frac{1}{A}$! Não existe a operação "divisão por matriz". Mesmo que ela existisse, como poderíamos dividir o *escalar* 1 pela *matriz* $A$? Se você alguma vez se sentir tentado a "dividir" por uma matriz, o que você realmente quer fazer é multiplicar pela sua inversa.

Podemos agora completar a analogia que mencionamos no começo desta seção.

### ◈ TEOREMA 2

Se $A$ é uma matriz invertível $n \times n$, o sistema de equações lineares dado por $A\mathbf{x} = \mathbf{b}$ tem uma única solução $\mathbf{x} = A^{-1}\mathbf{b}$ para todo $\mathbf{b}$ em $\mathbb{R}^n$.

DEMONSTRAÇÃO: O Teorema 2 essencialmente formaliza a observação que fizemos no começo desta seção. Voltaremos a ela mais cuidadosamente desta vez. Precisamos provar duas coisas: que $A\mathbf{x} = \mathbf{b}$ *tem* uma solução e que essa solução é *única*. (Em matemática, essa demonstração é chamada de "demonstração de existência e unicidade".)

Para mostrar que uma solução existe, precisamos apenas verificar que $\mathbf{x} = A^{-1}\mathbf{b}$ funciona. Calculamos

$$A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\,\mathbf{b} = \mathbf{b}$$

Assim, $A^{-1}\mathbf{b}$ satisfaz a equação $A\mathbf{x} = \mathbf{b}$, que, portanto, tem *pelo menos* essa solução.

Para mostrar que essa solução é única, suponha que $\mathbf{y}$ seja uma outra solução. Então, $A\mathbf{y} = \mathbf{b}$, e, multiplicando à esquerda ambos os lados da equação por $A^{-1}$, obtemos

$$A^{-1}(A\mathbf{y}) = A^{-1}\mathbf{b} \Rightarrow (A^{-1}A)\mathbf{y} = A^{-1}\mathbf{b} \Rightarrow I\mathbf{y} = A^{-1}\mathbf{b} \Rightarrow \mathbf{y} = A^{-1}\mathbf{b}$$

Portanto, $\mathbf{y}$ é a solução que tínhamos, por isso a solução é única. ◈

Agora, retornando às questões que levantamos nas observações anteriores ao Teorema 1, como podemos saber quando uma matriz tem uma inversa e como podemos achar a inversa quando a matriz é invertível? Daremos brevemente um procedimento geral, mas a situação para matrizes $2\times2$ é suficientemente simples para ser destacada.

### ◈ TEOREMA 3

Se $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, então $A$ será invertível se $ad - bc \neq 0$, caso em que

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Se $ad - bc = 0$, $A$ não será invertível.

A expressão $ad - bc$, chamada de ***determinante*** de $A$, é denotada por det $A$. A fórmula para a inversa de $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ (quando ela existe) é, portanto, $\frac{1}{\det A}$ vezes a matriz obtida permutando-se os elementos da diagonal principal e trocando-se os sinais dos outros dois elementos. Além de dar essa fórmula, o Teorema 3 nos diz que uma matriz 2×2 é invertível se, e somente se, $\det A \neq 0$. Veremos, no Capítulo 4, que o determinante pode ser definido para todas as matrizes quadradas e que esse resultado permanece verdadeiro, embora não exista uma fórmula simples para a inversa de matrizes quadradas de ordem maior.

DEMONSTRAÇÃO: Suponha que $\det A = ad - bc \neq 0$. Então:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} ad - bc & -ab + ba \\ cd - dc & -cb + da \end{bmatrix} = \begin{bmatrix} ad - bc & 0 \\ 0 & ad - bc \end{bmatrix} = \det A\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Analogamente,

$$\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \det A\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Já que $\det A \neq 0$, podemos multiplicar ambos os lados da equação por $1/\det A$ e obter

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\left(\frac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

ou, equivalentemente,

$$\left(\frac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\right)\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

[Note que usamos a propriedade (d) do Teorema 2 da Seção 3.3.] Assim, a matriz

$$\frac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

satisfaz a definição de inversa, por isso $A$ é invertível. Como a inversa de $A$ é única, pelo Teorema 1, devemos ter

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Reciprocamente, assuma que $ad - bc = 0$. Consideraremos separadamente os casos em que $a \neq 0$ e $a = 0$. Se $a \neq 0$, então $d = bc/a$, e assim a matriz pode ser escrita da forma

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ ac/a & bc/a \end{bmatrix} = \begin{bmatrix} a & b \\ ka & kb \end{bmatrix}$$

sendo $k = c/a$. Em outras palavras, a segunda linha de $A$ é um múltiplo da primeira. Voltando ao Exemplo 2(b), vemos que, se $A$ tem uma inversa $\begin{bmatrix} w & x \\ y & z \end{bmatrix}$, então

$$\begin{bmatrix} a & b \\ ka & kb \end{bmatrix}\begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

e o sistema de equações lineares correspondente

$$\begin{aligned} aw \quad &+ \quad by \quad &&= 1 \\ ax \quad &+ \quad bz &&= 0 \\ kaw \quad &+ kby &&= 0 \\ kax \quad &+ kbz &&= 1 \end{aligned}$$

não tem solução. (Por quê?)

Se a = 0, $ad - bc = 0$ implica que $bc = 0$, e, portanto, $b$ ou $c$ é 0. Assim, $A$ é da forma

$$\begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix} \quad \text{ou} \quad \begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix}$$

No primeiro caso, $\begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix}\begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ * & * \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Analogamente, $\begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix}$ não pode ter uma inversa. (Verifique isso.)

Conseqüentemente, se $ad - bc = 0$, $A$ não é invertível. ◈

**EXEMPLO 3** Ache as inversas de $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ e $B = \begin{bmatrix} 12 & -15 \\ 4 & -5 \end{bmatrix}$, se elas existirem.

SOLUÇÃO: Temos det $A = 1(4) - 2(3) = -2 \neq 0$, portanto, $A$ é invertível, com

$$A^{-1} = \frac{1}{-2}\begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$$

(Verifique isso.)

Por outro lado, det $B = 12(-5) - (-15)(4) = 0$, portanto, $B$ não é invertível.

**EXEMPLO 4** Use a inversa da matriz dos coeficientes para resolver o sistema linear

$$\begin{aligned} x + 2y &= 3 \\ 3x + 4y &= -2 \end{aligned}$$

SOLUÇÃO: A matriz dos coeficientes é $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, cuja inversa calculamos no Exemplo 3. Pelo Teorema 2, $A\mathbf{x} = \mathbf{b}$ tem a solução única $\mathbf{x} = A^{-1}\mathbf{b}$. Aqui, temos $\mathbf{b} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$; portanto, a solução do sistema dado é

$$\mathbf{x} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -8 \\ \frac{11}{2} \end{bmatrix}$$

Observação: ◆ Resolver um sistema linear $A\mathbf{x} = \mathbf{b}$ via $\mathbf{x} = A^{-1}\mathbf{b}$ pode parecer um bom método. Infelizmente, exceto para matrizes dos coeficientes 2×2 e para matrizes que tenham certas formas especiais, é quase sempre mais rápido usar o método de escalonamento de Gauss ou Gauss-Jordan para encontrar a solução. (Veja o Exercício 13.) Além disso, a técnica do Exemplo 4 funciona somente quando a matriz dos coeficientes é quadrada e invertível, enquanto o método do escalonamento pode sempre ser aplicado.

## *Propriedades das Matrizes Invertíveis*

O teorema seguinte enuncia algumas das propriedades mais importantes das matrizes invertíveis.

**◈ TEOREMA 4**

a. Se $A$ é uma matriz invertível, então $A^{-1}$ é invertível e

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

b. Se $A$ é uma matriz invertível e $c$ é um escalar não nulo, então $cA$ é uma matriz invertível e

$$(cA)^{-1} = \frac{1}{c}A^{-1}$$

c. Se $A$ e $B$ são matrizes invertíveis, então $AB$ é invertível e

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

d. Se $A$ é uma matriz invertível, então $A^T$ é invertível e

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

e. Se $A$ é uma matriz invertível, então $A^n$ é invertível para todo inteiro não negativo $n$ e

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

DEMONSTRAÇÃO: Vamos provar as propriedades (a), (c) e (e), deixando as propriedades (b) e (d) para serem provadas nos Exercícios 14 e 15.
(a) Para mostrar que $A^{-1}$ é invertível, devemos procurar por uma matriz $X$ tal que

$$A^{-1}X = I = XA^{-1}$$

A matriz $A$ certamente satisfaz essas equações, por isso $A^{-1}$ é invertível e $A$ é uma inversa de $A^{-1}$. Como inversas são únicas, temos que $(A^{-1})^{-1} = A$.
(c) Aqui, devemos mostrar que existe uma matriz $X$ tal que

$$(AB)X = I = X(AB)$$

A afirmação diz que, substituindo-se $B^{-1}A^{-1}$ por $X$, teremos as igualdades satisfeitas. Verificamos que

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

onde usamos a associatividade para mudar os parênteses. Analogamente, $(B^{-1}A^{-1})(AB) = I$ (verifique!), portanto, $AB$ é invertível e sua inversa é $B^{-1}A^{-1}$.
(e) A idéia básica aqui é relativamente fácil. Por exemplo, quando $n = 2$, temos

$$A^2(A^{-1})^2 = AAA^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I$$

Da mesma forma, $(A^{-1})^2A^2 = I$. Portanto, $(A^{-1})^2$ é a inversa de $A^2$. Não é difícil ver que um argumento análogo funciona para todos os valores maiores do inteiro $n$. Entretanto, a indução matemática é a forma de fazer a demonstração corretamente.

O primeiro passo é para $n = 0$, no qual devemos provar que $A^0$ é invertível e que

$$(A^0)^{-1} = (A^{-1})^0$$

Isso é equivalente a mostrar que $I$ é invertível e que $I^{-1} = I$, o que é claramente verdadeiro.

(Por quê? Veja o Exercício 16.)

Agora assuma que o resultado seja verdadeiro para $n = k$, sendo $k$ um inteiro não negativo específico. A hipótese de indução diz que $A^k$ é invertível e que

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

O passo de indução requer que provemos que $A^{k+1}$ é invertível e que $(A^{k+1})^{-1} = (A^{-1})^{k+1}$. Sabemos, por (c), que $A^{k+1} = A^kA$ é invertível, já que $A$ e (por hipótese de indução) $A^k$ são invertíveis. Mais ainda,

$$\begin{aligned}(A^{-1})^{k+1} &= (A^{-1})^kA^{-1} \\ &= (A^k)^{-1}A^{-1} && \text{Por hipótese de indução} \\ &= (AA^k)^{-1} && \text{Pela propriedade (c)} \\ &= (A^{k+1})^{-1}\end{aligned}$$

Portanto, $A^n$ é invertível para todo inteiro não negativo $n$, e $(A^n)^{-1} = (A^{-1})^n$, pelo princípio da indução matemática. ◈

Observações:

- Apesar de todas as propriedades do Teorema 4 serem úteis, a (c) é a de maior destaque: ela é talvez a propriedade algébrica mais importante das matrizes invertíveis. É também a mais fácil de ser aplicada erroneamente. No Exercício 17, você deve dar um contra-exemplo para mostrar que, ao contrário do que imaginávamos, em geral $(AB)^{-1} \neq A^{-1}B^{-1}$. A propriedade correta, $(AB)^{-1} = B^{-1}A^{-1}$, é algumas vezes chamada de regra da meia e sapato: apesar de calçarmos as meias antes dos sapatos, nós os retiramos na ordem inversa.
- A propriedade (c) é generalizada para produtos de um número finito de matrizes invertíveis: se $A_1, A_2, \ldots, A_n$ são matrizes invertíveis de mesma ordem, então $A_1A_2 \ldots A_n$ é invertível e

$$(A_1A_2 \cdots A_n)^{-1} = A_n^{-1} \cdots A_2^{-1}A_1^{-1}$$

(Veja o Exercício 18.) Assim, podemos enunciar:

A inversa de um produto de matrizes invertíveis é o produto das inversas na ordem reversa.

- Como, para números reais, $\dfrac{1}{a+b} \neq \dfrac{1}{a} + \dfrac{1}{b}$, não devemos esperar que, para matrizes quadradas, $(A + B)^{-1} = A^{-1} + B^{-1}$ (e isso realmente não é verdade em geral; veja o Exercício 19). Na realidade, exceto para matrizes especiais, não existe nenhuma fórmula para $(A + B)^{-1}$.
- A propriedade (e) nos permite definir potências inteiras negativas de uma matriz invertível:

Se $A$ é uma matriz invertível e $n$ é um inteiro positivo, então $A^{-n}$ é definida por

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

Com essa definição, podem ser demonstradas as regras para exponenciação, $A^rA^s = A^{r+s}$ e $(A^r)^s = A^{rs}$, para todos os inteiros $r$ e $s$, quando $A$ é invertível.

Um dos objetivos das propriedades algébricas das matrizes é a resolução de equações que envolvem matrizes. O próximo exemplo ilustra o processo. Note que devemos dar uma atenção especial à ordem das matrizes no produto.

**EXEMPLO 5** Resolva a seguinte equação matricial em $X$ (assuma que as matrizes são tais que as operações indicadas estão definidas):

$$A^{-1}(BX)^{-1} = (A^{-1}B^3)^2$$

SOLUÇÃO: Existem vários procedimentos aqui. Uma solução é

$$\begin{aligned} A^{-1}(BX)^{-1} = (A^{-1}B^3)^2 &\Rightarrow ((BX)\,A)^{-1} = (A^{-1}B^3)^2 \\ &\Rightarrow [((BX)\,A)^{-1}]^{-1} = [(A^{-1}B^3)^2]^{-1} \\ &\Rightarrow (BX)\,A = [(A^{-1}B^3)(A^{-1}B^3)]^{-1} \\ &\Rightarrow (BX)\,A = B^{-3}(A^{-1})^{-1}B^{-3}(A^{-1})^{-1} \\ &\Rightarrow BXA = B^{-3}AB^{-3}A \\ &\Rightarrow B^{-1}BXAA^{-1} = B^{-1}B^{-3}AB^{-3}AA^{-1} \\ &\Rightarrow IXI = B^{-4}AB^{-3}I \\ &\Rightarrow X = B^{-4}AB^{-3} \end{aligned}$$

(Você consegue justificar cada passo?) Observe o uso cuidadoso do Teorema 4(c) e a expansão de $(A^{-1}B^3)^2$. Fizemos também um uso liberal da associatividade da multiplicação de matrizes para simplificar (ou eliminar) a colocação de parênteses.

## *Matrizes Elementares*

Vamos usar multiplicação de matrizes para ter uma perspectiva diferente do escalonamento de matrizes. Nesse processo, você descobrirá muitas idéias novas e importantes da natureza das matrizes invertíveis.
Se

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{e} \quad A = \begin{bmatrix} 5 & 7 \\ -1 & 0 \\ 8 & 3 \end{bmatrix}$$

calculamos

$$EA = \begin{bmatrix} 5 & 7 \\ 8 & 3 \\ -1 & 0 \end{bmatrix}$$

Em outras palavras, multiplicar $A$ por $E$ (à esquerda) é equivalente a permutarmos as linhas 2 e 3 de $A$. O que é significante em $E$? $E$ é simplesmente a matriz obtida aplicando-se a mesma operação elementar nas linhas, $L_2 \leftrightarrow L_3$, à matriz identidade $I_3$. É fato que isso sempre funciona.

**Definição** Uma ***matriz elementar*** é uma matriz obtida por meio de operações elementares nas linhas de uma matriz identidade.

Como existem três tipos de operações elementares nas linhas de uma matriz, há três tipos correspondentes de matrizes elementares. Apresentamos mais algumas matrizes elementares.

**EXEMPLO 6** Sejam

$$E_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{e} \quad E_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -2 & 0 & 1 \end{bmatrix}$$

Cada uma dessas matrizes foi obtida da matriz identidade $I_4$, pela aplicação de uma única operação elementar em suas linhas. A matriz $E_1$ corresponde a $3L_2$, $E_2$ corresponde a $L_1 \leftrightarrow L_3$, e $E_3$, a $L_4 - 2L_2$. Observe que, quando multiplicamos à esquerda uma matriz $4 \times n$ por uma dessas matrizes elementares, a operação elementar é realizada nas linhas dessa matriz. Por exemplo, se

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}$$

então

$$E_1A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 3a_{21} & 3a_{22} & 3a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}, \quad E_2A = \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{41} & a_{42} & a_{43} \end{bmatrix},$$

e

$$E_3A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} - 2a_{21} & a_{42} - 2a_{22} & a_{43} - 2a_{23} \end{bmatrix}$$

O Exemplo 6 e os Exercícios de 24 a 30 devem convencê-lo de que *toda* operação elementar nas linhas de *qualquer* matriz pode ser realizada por uma multiplicação à esquerda por uma matriz elementar adequada. Enunciamos esse fato em um teorema cuja demonstração omitimos.

**TEOREMA 5**

Seja $E$ a matriz elementar obtida fazendo-se uma operação elementar nas linhas de $I_n$. Se a mesma operação elementar for feita nas linhas de uma matriz $n \times r$ $A$, o resultado será igual à matriz $EA$.

Observação: Do ponto de vista computacional, não é boa idéia usar matrizes elementares para fazer operações elementares — faça-as diretamente. Entretanto, matrizes elementares nos dão algumas idéias valiosas sobre matrizes invertíveis e na resolução de sistemas de equações lineares.

Já observamos que toda operação elementar pode s er "desfeita" ou "revertida". Essa mesma observação, aplicada a matrizes elementares, nos mostra que elas são invertíveis.

**EXEMPLO 7** Sejam

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{e} \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$$

Então, $E_1$ corresponde a $L_2 \leftrightarrow L_3$, que é desfeita fazendo-se $L_2 \leftrightarrow L_3$ novamente. Assim $E_1^{-1} = E_1$. (Verifique isso provando que $E_1^2 = E_1E_1 = I$.) A matriz $E_2$ está relacionada com $4L_2$, que pode ser revertida fazendo-se $\frac{1}{4}L_2$. Assim,

$$E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

que pode ser facilmente verificada. Finalmente, $E_3$ corresponde à operação elementar $L_3 - 2L_1$, que pode ser desfeita pela operação elementar $L_3 + 2L_1$. Assim, nesse caso,

$$E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$$

(Novamente, é fácil verificar isso confirmando que o produto dessa matriz por $E_3$, em ambos os lados, será igual a $I$.)

Note que não só toda matriz elementar é invertível, mas a sua inversa é também uma matriz elementar e do mesmo tipo. Enunciamos esse resultado no próximo teorema.

**TEOREMA 6**

Toda matriz elementar é invertível, e a sua inversa é uma matriz elementar do mesmo tipo.

## *O Teorema Fundamental das Matrizes Invertíveis*

Estamos prontos para provar um dos resultados principais deste livro — uma lista de caracterizações equivalentes do significado de matriz invertível. De certa forma, grande parte da álgebra linear está relacionada com esse teorema, no desenvolvimento dessas caracterizações ou nas suas aplicações. Como você deve esperar, depois dessa introdução, usaremos muito esse teorema. Familiarize-se com ele!

Referimo-nos ao Teorema 7 como a primeira versão do Teorema Fundamental, já que ele será ampliado nos próximos capítulos. Lembre-se de que, quando dizemos que afirmações sobre uma matriz são equivalentes, isso significa que, para uma matriz $A$ dada, as afirmações são todas verdadeiras ou todas falsas.

**TEOREMA 7 O Teorema Fundamental das Matrizes Invertíveis: Versão 1**

Seja $A$ uma matriz $n \times n$. As seguintes afirmações são equivalentes:

a. $A$ é invertível.
b. $A\mathbf{x} = \mathbf{b}$ tem uma única solução para todo $\mathbf{b}$ em $\mathbb{R}^n$.
c. $A\mathbf{x} = \mathbf{0}$ tem apenas uma solução trivial.
d. A forma escalonada reduzida de $A$ é $I_n$.
e. $A$ é um produto de matrizes elementares.

DEMONSTRAÇÃO: Provaremos o teorema fazendo a cadeia circular de implicações

$$(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (e) \Rightarrow (a)$$

(a) $\Rightarrow$ (b) Já provamos que, se $A$ é invertível, então $A\mathbf{x} = \mathbf{b}$ tem uma única solução $\mathbf{x} = A^{-1}\mathbf{b}$ para todo $\mathbf{b}$ em $\mathbb{R}^n$ (Teorema 2).
(b) $\Rightarrow$ (c) Assuma que $A\mathbf{x} = \mathbf{b}$ tem uma única solução para todo $\mathbf{b}$ em $\mathbb{R}^n$. Isso implica, em particular, que $A\mathbf{x} = \mathbf{0}$ tem uma única solução. Mas um sistema homogêneo $A\mathbf{x} = \mathbf{0}$ tem sempre $\mathbf{x} = \mathbf{0}$ como *uma* solução. Portanto, nesse caso, $\mathbf{x} = \mathbf{0}$ deve ser *a* solução.
(c) $\Rightarrow$ (d) Suponha que $A\mathbf{x} = \mathbf{0}$ tenha apenas uma solução trivial. O sistema de equações lineares correspondente é

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= 0 \end{aligned}$$

e assumimos que a solução seja

$$\begin{aligned} x_1 \qquad\qquad &= 0 \\ x_2 \qquad &= 0 \\ \ddots \quad & \\ x_n &= 0 \end{aligned}$$

Em outras palavras, o escalonamento de Gauss-Jordan aplicado à matriz completa do sistema dá

$$[A|\mathbf{0}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{array}\right] = [I_n|\mathbf{0}]$$

Assim, a forma escalonada reduzida de $A$ é $I_n$.

(d) $\Rightarrow$ (e) Se assumirmos que a forma escalonada reduzida de $A$ é $I_n$, $A$ pode ser transformada em $I_n$ por meio de uma seqüência finita de operações elementares em suas linhas. Pelo Teorema 5, cada uma dessas operações elementares pode ser realizada pela multiplicação à esquerda por uma matriz elementar adequada. Se a seqüência dessas matrizes elementares for $E_1, E_2, \ldots, E_k$ (nessa ordem), teremos

$$E_k \ldots E_2E_1A = I_n$$

Conforme o Teorema 6, essas matrizes elementares são todas invertíveis. Portanto, o produto delas também o é, e temos

$$A = (E_k \ldots E_2E_1)^{-1} I_n = (E_k \ldots E_2E_1)^{-1} = E_1^{-1} E_2^{-1} \ldots E_k^{-1}$$

Novamente, cada $E_i^{-1}$ é uma matriz elementar pelo Teorema 6, por isso escrevemos $A$ como um produto de matrizes elementares, como queríamos.

(e) $\Rightarrow$ (a) Se $A$ é um produto de matrizes elementares, $A$ é invertível, já que matrizes elementares são invertíveis e produtos de matrizes invertíveis são invertíveis. ◈

Este é um teorema extremamente poderoso. Para ilustrar sua força, consideraremos duas de suas conseqüências. A primeira é que, embora a definição de matrizes invertíveis afirme que uma matriz $A$ é invertível se existe uma matriz $B$ tal que *ambas* as equações, $AB = I$ e $BA = I$, sejam satisfeitas, precisamos verificar somente uma delas. Dessa forma, eliminamos metade do nosso trabalho!

### ◈ TEOREMA 8

Seja $A$ uma matriz quadrada. Se $B$ é uma matriz quadrada tal que $AB = I$ ou $BA = I$, então $A$ é invertível e $B = A^{-1}$.

DEMONSTRAÇÃO: Suponha que $BA = I$. Considere a equação $A\mathbf{x} = \mathbf{0}$. Multiplicando-se à esquerda por $B$, temos $BA\mathbf{x} = B\mathbf{0}$. Isso significa que $\mathbf{x} = I\mathbf{x} = \mathbf{0}$. Assim, o sistema representado por $A\mathbf{x} = \mathbf{0}$ tem a única solução $\mathbf{x} = \mathbf{0}$. Da equivalência entre (c) e (a) no Teorema Fundamental, sabemos que A é invertível ($A^{-1}$ existe e satisfaz $AA^{-1} = I = A^{-1}A$).

Se multiplicarmos agora ambos os lados de $BA = I$ por $A^{-1}$ pela direita, obtemos

$$BAA^{-1} = I\,A^{-1} \Rightarrow BI = A^{-1} \Rightarrow B = A^{-1}$$

(A demonstração no caso de $AB = I$ fica como Exercício 41.) ◈

A próxima conseqüência do Teorema Fundamental é base para um método eficiente de cálculo da inversa de uma matriz.

### ◈ TEOREMA 9

Seja $A$ uma matriz quadrada. Se uma seqüência de operações elementares nas suas linhas reduz $A$ a $I$, a mesma seqüência de operações elementares transforma $I$ em $A^{-1}$.

DEMONSTRAÇÃO: Se $A$ for linha equivalente a $I$, podemos conseguir essa redução multiplicando $A$ à esquerda por uma seqüência $E_1, E_2, \ldots, E_k$ de matrizes elementares. Portanto, temos $E_k \ldots E_2E_1A = I$. Denotando $B = E_k \ldots E_2E_1$, temos $BA = I$. Pelo Teorema 8, $A$ é invertível e $A^{-1} = B$. Agora, aplicar a mesma seqüência de operações elementares em $I$ é equivalente a multiplicar $I$ à esquerda por $E_k \ldots E_2E_1 = B$. O resultado é

$$E_k \ldots E_2E_1I = BI = B = AB^{-1}$$

Dessa forma, $I$ é transformada em $A^{-1}$ pela mesma seqüência de operações elementares de linhas. ◈

## *O Método de Gauss-Jordan para o Cálculo da Inversa*

Podemos efetuar operações linha em $A$ e $I$ simultaneamente, construindo uma "matriz supercompleta" $[A \mid I]$. O Teorema 9 mostra que, se $A$ é linha equivalente a $I$ [o que, pelo Teorema Fundamental (d) $\Leftrightarrow$ (a), significa que $A$ é invertível], operações elementares em suas linhas resultarão em

$$[A|I] \longrightarrow [I|A^{-1}]$$

Se $A$ não pode ser reduzida a $I$, o Teorema Fundamental nos garante que $A$ não é invertível.

O método que acabamos de descrever é simplesmente o escalonamento de Gauss-Jordan efetuado em uma matriz completa $n\times 2n$, não em uma matriz completa $n\times(n+1)$. Outra maneira de encarar esse método é olhar para o problema de encontrar $A^{-1}$ através da resolução da equação matricial $AX = I_n$ para uma matriz $X$ $n \times n$. (Isso é suficiente pelo Teorema Fundamental, já que uma inversa de $A$ à direita deve ser a sua inversa.) Se denotarmos as colunas de $X$ por $\mathbf{x}_1, \ldots, \mathbf{x}_n$, essa equação matricial será equivalente à resolução nas colunas de $X$, uma de cada vez. Agora, sendo as colunas de $I_n$ os vetores canônicos unitários $\mathbf{e}_1, \ldots, \mathbf{e}_n$, temos $n$ sistemas de equações lineares, todos com $A$ sendo a matriz dos coeficientes:

$$A\mathbf{x}_1 = \mathbf{e}_1, \ldots, A\mathbf{x}_n = \mathbf{e}_n$$

Como é necessária a mesma seqüência de operações linha para transformar $A$ na sua forma escalonada reduzida, as matrizes completas desses sistemas, $[A \mid \mathbf{e}_1], \ldots, [A \mid \mathbf{e}_n]$, podem ser combinadas por

$$[A \mid \mathbf{e}_1\, \mathbf{e}_2 \cdots \mathbf{e}_n] = [A \mid I_n]$$

Agora efetuamos operações linha para tentar reduzir $A$ a $I_n$, as quais, se bem-sucedidas, transformarão $I_n$ em $A^{-1}$.

Ilustramos com três exemplos esse uso do escalonamento de Gauss-Jordan.

**EXEMPLO 8** Encontre a inversa de

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

se ela existir.

SOLUÇÃO: O escalonamento de Gauss-Jordan fornece

$$[A \mid I] = \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 2 & 2 & 4 & 0 & 1 & 0 \\ 1 & 3 & -3 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[L_3 - L_1]{L_2 - 2L_1} \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & -2 & 6 & -2 & 1 & 0 \\ 0 & 1 & -2 & -1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{(-\frac{1}{2})L_2} \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 1 & -\frac{1}{2} & 0 \\ 0 & 1 & -2 & -1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{L_3 - L_2} \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & -2 & \frac{1}{2} & 1 \end{array}\right]$$

$$\xrightarrow[L_2 + 3L_3]{L_1 + L_3} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -1 & \frac{1}{2} & 1 \\ 0 & 1 & 0 & -5 & 1 & 3 \\ 0 & 0 & 1 & -2 & \frac{1}{2} & 1 \end{array}\right]$$

$$\xrightarrow{L_1 - 2L_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 9 & -\frac{3}{2} & -5 \\ 0 & 1 & 0 & -5 & 1 & 3 \\ 0 & 0 & 1 & -2 & \frac{1}{2} & 1 \end{array}\right]$$

Portanto,

$$A^{-1} = \begin{bmatrix} 9 & -\frac{3}{2} & -5 \\ -5 & 1 & 3 \\ -2 & \frac{1}{2} & 1 \end{bmatrix}$$

(Você deve sempre verificar que $AA^{-1} = I$ pela multiplicação direta. Pelo Teorema 8, não precisamos verificar que $A^{-1}A = I$.)

Observação: Note que usamos uma variante do escalonamento de Gauss-Jordan que introduz primeiro todos os zeros *abaixo* dos pivôs 1, da esquerda para a direita e de cima para baixo, e depois conseguimos os zeros *acima* dos pivôs 1, da direita para a esquerda e de baixo para cima. Essa forma economiza cálculos, como já observamos no Capítulo 2, mas você pode achar mais fácil, quando fizer as contas à mão, criar *todos* os zeros em cada coluna conforme for trabalhando. A resposta obviamente será a mesma.

**EXEMPLO 9** Encontre a inversa de

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

se ela existir.

SOLUÇÃO: Procederemos como no Exemplo 8, colocando a matriz identidade ao lado da matriz $A$ e manipulando $[A \mid I]$ tentando transformá-la em $[I \mid A^{-1}]$.

$$[A|I] = \left[\begin{array}{rrr|rrr} 2 & 1 & -4 & 1 & 0 & 0 \\ -4 & -1 & 6 & 0 & 1 & 0 \\ -2 & 2 & -2 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[L_3+L_1]{L_2+2L_1} \left[\begin{array}{rrr|rrr} 2 & 1 & -4 & 1 & 0 & 0 \\ 0 & 1 & -2 & 2 & 1 & 0 \\ 0 & 3 & -6 & 1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{L_3-3L_2} \left[\begin{array}{rrr|rrr} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 2 & 1 & 0 \\ 0 & 0 & 0 & -5 & -3 & 1 \end{array}\right]$$

Nesse ponto, vemos que não é possível reduzir $A$ a $I$, já que encontramos uma linha de zeros do lado esquerdo da matriz completa. Conseqüentemente, $A$ não é invertível.

Como o próximo exemplo ilustra, tudo funciona da mesma maneira em $\mathbb{Z}_p$, em que $p$ é um inteiro primo.

**EXEMPLO 10** Encontre a inversa de

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

sobre $\mathbb{Z}_3$, se ela existir.

SOLUÇÃO 1: Usamos o método de Gauss-Jordan, lembrando que os cálculos são feitos em $\mathbb{Z}_3$.

$$
\begin{aligned}
[A \mid I] &= \left[\begin{array}{cc|cc} 2 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{array}\right] \\
&\xrightarrow{2L_1} \left[\begin{array}{cc|cc} 1 & 1 & 2 & 0 \\ 2 & 0 & 0 & 1 \end{array}\right] \\
&\xrightarrow{L_2+L_1} \left[\begin{array}{cc|cc} 1 & 1 & 2 & 0 \\ 0 & 1 & 2 & 1 \end{array}\right] \\
&\xrightarrow{L_1+2L_2} \left[\begin{array}{cc|cc} 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 1 \end{array}\right]
\end{aligned}
$$

Portanto, $A^{-1} = \begin{bmatrix} 0 & 2 \\ 2 & 1 \end{bmatrix}$, e é fácil verificar que, em $\mathbb{Z}_3$, $AA^{-1} = I$.

SOLUÇÃO 2: Como $A$ é uma matriz $2 \times 2$, podemos calcular a sua inversa usando a fórmula dada no Teorema 3. O determinante de $A$ é

$$\det A = 2(0) - 2(2) = -1 = 2$$

em $\mathbb{Z}_3$ (pois $2 + 1 = 0$). Portanto, $A^{-1}$ existe e é dada pela fórmula do Teorema 3. Devemos ser cuidadosos aqui, já que a fórmula introduz a "fração" $1/\det A$, e não há frações em $\mathbb{Z}_3$. Precisamos usar inversos multiplicativos.

No lugar de $1/\det A = \frac{1}{2}$, usamos $2^{-1}$ — isto é, achamos um número $x$ em $\mathbb{Z}_3$ que satisfaz a equação $2x = 1$. É fácil ver que $x = 2$ é a solução que queremos: em $\mathbb{Z}_3$, $2^{-1} = 2$, pois $2(2) = 1$. A fórmula para $A^{-1}$ é

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

que concorda com a nossa primeira solução.

## ◈ EXERCÍCIOS 3.4 ◈

*Nos Exercícios de 1 a 10, ache a inversa da matriz dada (se ela existir) usando o Teorema 3.*

**1.** $\begin{bmatrix} 4 & 7 \\ 1 & 2 \end{bmatrix}$

**2.** $\begin{bmatrix} 4 & -2 \\ 2 & 0 \end{bmatrix}$

**3.** $\begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$

**4.** $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

**5.** $\begin{bmatrix} \frac{3}{4} & \frac{3}{5} \\ \frac{5}{6} & \frac{2}{3} \end{bmatrix}$

**6.** $\begin{bmatrix} \sqrt{2}/2 & -\sqrt{2} \\ 2\sqrt{2} & \sqrt{2} \end{bmatrix}$

**7.** $\begin{bmatrix} -1{,}5 & -4{,}2 \\ 0{,}5 & 2{,}4 \end{bmatrix}$

**8.** $\begin{bmatrix} 2{,}54 & 8{,}128 \\ 0{,}25 & 0{,}8 \end{bmatrix}$

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

**10.** $\begin{bmatrix} 1/a & 1/b \\ 1/c & 1/d \end{bmatrix}$, em que $a$, $b$, $c$ e $d$ não são 0.

*Nos Exercícios 11 e 12, resolva os sistemas dados usando o método do Exemplo 4.*

**11.** $\begin{aligned} 2x + \phantom{3}y &= -1 \\ 5x + 3y &= \phantom{-}2 \end{aligned}$

**12.** $\begin{aligned} x_1 - x_2 &= 1 \\ 2x_1 + x_2 &= 2 \end{aligned}$

**13.** Sejam $A = \begin{bmatrix} 1 & 2 \\ 2 & 6 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ e $\mathbf{b}_3 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$.

(a) Encontre $A^{-1}$ e use-a para resolver os três sistemas $A\mathbf{x} = \mathbf{b}_1$, $A\mathbf{x} = \mathbf{b}_2$ e $A\mathbf{x} = \mathbf{b}_3$.

(b) Resolva todos os três sistemas simultaneamente reduzindo por linhas a matriz completa $[A \mid \mathbf{b}_1\, \mathbf{b}_2\, \mathbf{b}_3]$, por meio do escalonamento de Gauss-Jordan.

(c) Conte cuidadosamente o número total de multiplicações individuais que você efetuou em (a) e em (b). Você deve descobrir que, mesmo para esse exemplo 2×2, um método usa menos operações. Para sistemas maiores, a diferença é ainda mais acentuada, e isso explica por que sistemas computacionais não usam um desses métodos para resolver sistemas lineares.

**14.** Prove o Teorema 4(b).

**15.** Prove o Teorema 4(d).

**16.** Prove que a matriz identidade $I_n$ $n\times n$ é invertível e que $I_n^{-1} = I_n$.

**17.** (a) Dê um contra-exemplo para mostrar que, em geral, $(AB)^{-1} \neq A^{-1}B^{-1}$.

(b) Sob que condições em $A$ e $B$ é verdade que $(AB)^{-1} = A^{-1}B^{-1}$? Prove sua resposta.

**18.** Por indução, prove que, se $A_1, A_2, \ldots, A_n$ são matrizes invertíveis do mesmo tamanho, o produto $A_1A_2\cdots A_n$ é invertível e $(A_1A_2\cdots A_n)^{-1} = A_n^{-1}\cdots A_2^{-1}A_1^{-1}$.

**19.** Dê um contra-exemplo para mostrar que, em geral, $(A+B)^{-1} \neq A^{-1} + B^{-1}$.

*Nos Exercícios de 20 a 23, resolva a equação matriz para X. Simplifique suas respostas tanto quanto possível. (Nas palavras de Albert Einstein, "Tudo deve ser feito tão simples quanto possível, mas não mais simples".) Assuma que todas as matrizes são invertíveis.*

**20.** $XA^2 = A^{-1}$ **21.** $AXB = (BA)^2$

**22.** $(A^{-1}X)^{-1} = A(B^{-2}A)^{-1}$

**23.** $ABXA^{-1}B^{-1} = I + A$

*Nos Exercícios de 24 a 30, sejam*

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & -1 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 1 \\ 2 & 1 & -1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 2 & -1 \\ -3 & -1 & 3 \\ 2 & 1 & -1 \end{bmatrix}$$

*Em cada caso, encontre a matriz elementar E que satisfaça a equação dada.*

**24.** $EA = B$ **25.** $EB = A$ **26.** $EA = C$

**27.** $EC = A$ **28.** $EC = D$ **29.** $ED = C$

**30.** Existe uma matriz elementar $E$ tal que $EA = D$? Por quê?

*Nos Exercícios de 31 a 38, encontre a inversa da matriz elementar dada.*

**31.** $\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$ **32.** $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

**33.** $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ **34.** $\begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & 1 \end{bmatrix}$

**35.** $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$ **36.** $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

**37.** $\begin{bmatrix} 1 & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{bmatrix}, c \neq 0$ **38.** $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}, c \neq 0$

*Nos Exercícios 39 e 40, encontre a seqüência de matrizes elementares $E_1, E_2, \ldots, E_k$ tais que $E_k \cdots E_2E_1A = I$. Use essa seqüência para escrever $A$ e $A^{-1}$ como produtos de matrizes elementares.*

**39.** $A = \begin{bmatrix} 1 & 0 \\ -1 & -2 \end{bmatrix}$ **40.** $A = \begin{bmatrix} 2 & 4 \\ 1 & 1 \end{bmatrix}$

**41.** Prove o Teorema 8 para o caso de $AB = I$.

**42.** (a) Prove que, se $A$ é invertível e $AB = O$, então $B = O$.
(b) Dê um contra-exemplo para mostrar que o resultado da parte (a) pode falhar se $A$ não for invertível.

**43.** (a) Prove que, se $A$ é invertível e $BA = CA$, então $B = C$.

(b) Dê um contra-exemplo para mostrar que o resultado da parte (a) pode falhar se $A$ não for invertível.

**44.** A matriz quadrada $A$ é chamada ***idempotente*** se $A^2 = A$. (A palavra *idempotente* vem do Latim *idem*, que significa "mesmo", e *potere*, que significa "ter potência". Assim, algo que é idempotente tem a "mesma potência" quando elevado ao quadrado.)

(a) Ache três matrizes 2×2 idempotentes.

(b) Prove que a única matriz $n\times n$ idempotente invertível é a matriz identidade.

**45.** Mostre que, se $A$ é uma matriz quadrada que satisfaz a equação $A^2 - 2A + I = O$, então $A^{-1} = 2I - A$.

**46.** Prove que, se uma matriz simétrica é invertível, sua inversa também é simétrica.

**47.** Prove que, se $A$ e $B$ são matrizes quadradas e $AB$ é invertível, $A$ e $B$ são invertíveis.

*Nos Exercícios de 48 a 63, use o método de Gauss-Jordan para achar a inversa da matriz dada (se existir).*

**48.** $\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}$ **49.** $\begin{bmatrix} -2 & 4 \\ 3 & -1 \end{bmatrix}$

**50.** $\begin{bmatrix} 6 & -4 \\ -3 & 2 \end{bmatrix}$ **51.** $\begin{bmatrix} 1 & a \\ -a & 1 \end{bmatrix}$

**52.** $\begin{bmatrix} 2 & 3 & 0 \\ 1 & -2 & -1 \\ 2 & 0 & -1 \end{bmatrix}$ **53.** $\begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & 2 \\ 2 & 3 & -1 \end{bmatrix}$

**54.** $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

**55.** $\begin{bmatrix} a & 0 & 0 \\ 1 & a & 0 \\ 0 & 1 & a \end{bmatrix}$

**56.** $\begin{bmatrix} 0 & a & 0 \\ b & 0 & c \\ 0 & d & 0 \end{bmatrix}$

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

**58.** $\begin{bmatrix} \sqrt{2} & 0 & 2\sqrt{2} & 0 \\ -4\sqrt{2} & \sqrt{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 1 \end{bmatrix}$

**59.** $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ a & b & c & d \end{bmatrix}$

**60.** $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ sobre $\mathbb{Z}_2$

**61.** $\begin{bmatrix} 4 & 2 \\ 3 & 4 \end{bmatrix}$ sobre $\mathbb{Z}_5$

**62.** $\begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$ sobre $\mathbb{Z}_3$

**63.** $\begin{bmatrix} 1 & 5 & 0 \\ 1 & 2 & 4 \\ 3 & 6 & 1 \end{bmatrix}$ sobre $\mathbb{Z}_7$

*Decompor matrizes quadradas grandes pode às vezes tornar suas inversas mais fáceis de se calcular, particularmente se os blocos tiverem uma forma adequada. Nos Exercícios de 64 a 68, verifique, por multiplicação por blocos, que a inversa de uma matriz, decomposta como apresentado, resulta na matriz dada. (Assuma que todas as inversas existem, quando necessário.)*

**64.** $\begin{bmatrix} A & B \\ O & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}BD^{-1} \\ O & D^{-1} \end{bmatrix}$

**65.** $\begin{bmatrix} O & B \\ C & I \end{bmatrix}^{-1} = \begin{bmatrix} -(BC)^{-1} & (BC)^{-1}B \\ C(BC)^{-1} & I - C(BC)^{-1}B \end{bmatrix}$

**66.** $\begin{bmatrix} I & B \\ C & I \end{bmatrix}^{-1} = \begin{bmatrix} (I - BC)^{-1} & -(I - BC)^{-1}B \\ -C(I - BC)^{-1} & I + C(I - BC)^{-1}B \end{bmatrix}$

**67.** $\begin{bmatrix} O & B \\ C & D \end{bmatrix}^{-1}$

$$= \begin{bmatrix} -(BD^{-1}C)^{-1} & (BD^{-1}C)^{-1}BD^{-1} \\ D^{-1}C(BD^{-1}C)^{-1} & D^{-1} - D^{-1}C(BD^{-1}C)^{-1}BD^{-1} \end{bmatrix}$$

**68.** $\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} P & Q \\ R & S \end{bmatrix}$, em que $P = (A - BD^{-1}C)^{-1}$, $Q = -PBD^{-1}$, $R = -D^{-1}CP$ e $S = D^{-1} + D^{-1}CPBD^{-1}$

*Nos Exercícios de 69 a 72, divida as matrizes dadas de modo que você possa aplicar uma das fórmulas dos Exercícios 64 a 68, e então calcule a inversa usando a fórmula.*

**69.** $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 3 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{bmatrix}$

**70.** A matriz do Exercício 58.

**71.** $\begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$

**72.** $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 3 & 1 \\ -1 & 5 & 2 \end{bmatrix}$

# 3.5 Subespaços, Base, Dimensão e Posto

Esta seção introduz talvez as idéias mais importantes de todo o livro. Já vimos que há uma interação entre geometria e álgebra: freqüentemente, podemos usar intuição geométrica e lógica para obter resultados algébricos, e o poder da álgebra normalmente nos permite estender nossos achados bem além das configurações geométricas em que eles primeiramente surgiram.

Em nosso estudo de vetores, já encontramos informalmente todos os conceitos desta seção. Aqui, começaremos a nos tornar mais formais, dando definições para as idéias-chave. Como você verá, a noção de um *subespaço* é simplesmente uma generalização algébrica dos exemplos geométricos de retas e planos que passam pela origem. O conceito fundamental de uma *base* para um subespaço é então derivado da idéia de vetores diretores para tais retas e planos. O conceito de base nos permitirá dar uma definição precisa de *dimensão* que estará de acordo com uma idéia intuitiva e geométrica do termo, apesar de ser suficientemente flexível para permitir a generalização de outras configurações.

Você também começará a ver que essas idéias dão mais luz ao que você já sabe sobre as matrizes e a solução de sistemas de equações lineares. No Capítulo 6, encontraremos todas essas idéias fundamentais novamente, mais detalhadas. Considere esta seção uma "introdução ao pensamento abstrato".

Um plano que passa pela origem em $\mathbb{R}^3$ "assemelha-se" a uma cópia de $\mathbb{R}^2$. Intuitivamente, concordaríamos que ambos são "bidimensionais". Aprofundando-nos, podemos dizer também que todo cálculo que pode ser feito com vetores em $\mathbb{R}^2$ também pode ser feito em um plano que passa pela origem. Em particular, podemos somar e multiplicar por escalares (e, mais geralmente, fazer combinações lineares) de vetores em tal plano e os resultados serão vetores *no mesmo plano*. Dizemos que, como em $\mathbb{R}^2$, um plano que passa pela origem é *fechado* em relação às operações de adição e de multiplicação por escalar. (Veja Figura 1.)

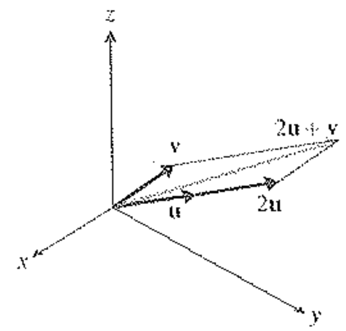



**Figura 1**

Os vetores desse plano são objetos bi ou tridimensionais? Podemos argumentar que eles são tridimensionais, porque estão em $\mathbb{R}^3$ e, conseqüentemente, têm três coordenadas. Por outro lado, esses vetores podem ser descritos como combinações lineares de apenas dois vetores — os vetores diretores do plano —, e desse ponto de vista eles são objetos bidimensionais dentro de um plano bidimensional. A noção de subespaço é a chave para resolver essa questão.

**Definição** Um ***subespaço*** de $\mathbb{R}^n$ é uma coleção $S$ de vetores de $\mathbb{R}^n$ tal que

1. O vetor $\mathbf{0}$ pertence a $S$.
2. Se $\mathbf{u}$ e $\mathbf{v}$ pertencem a $S$, então $\mathbf{u} + \mathbf{v}$ também pertence a $S$. ($S$ é ***fechado pela adição***.)
3. Se $\mathbf{v}$ está em $S$ e $c$ é um escalar, então $c\mathbf{v}$ também pertence a $S$. ($S$ é ***fechado pela multiplicação por escalar***.)

Podemos combinar as propriedades (2) e (3) e exigir equivalentemente que $S$ seja ***fechado por combinações lineares***:

Se $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ pertencem a $S$ e $c_1, c_2, \ldots, c_k$ são escalares,
então $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_k\mathbf{u}_k$ pertencem a $S$.

**EXEMPLO 1** Toda reta e todo plano que passam pela origem em $\mathbb{R}^3$ são um subespaço de $\mathbb{R}^3$. Geometricamente, deve ficar claro que as propriedades de (1) a (3) estão satisfeitas. Vamos dar uma demonstração algébrica para o caso de um plano que passa pela origem. Deixamos para você a demonstração correspondente para o caso de uma reta que passa pela origem, no Exercício 5.

Seja $\mathcal{P}$ um plano que passa pela origem com vetores diretores $\mathbf{v}_1$ e $\mathbf{v}_2$. Então, $\mathcal{P} = \text{ger}(\mathbf{v}_1, \mathbf{v}_2)$. O vetor nulo $\mathbf{0}$ pertence a $\mathcal{P}$, pois $\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2$. Agora, considere

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \quad \text{e} \quad \mathbf{v} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2$$

como dois vetores em $\mathcal{P}$. Então:

$$\mathbf{u} + \mathbf{v} = (c_1\mathbf{v}_1 + c_2\mathbf{v}_2) + (d_1\mathbf{v}_1 + d_2\mathbf{v}_2) = (c_1 + d_1)\,\mathbf{v}_1 + (c_2 + d_2)\,\mathbf{v}_2$$

Portanto, $\mathbf{u} + \mathbf{v}$ é uma combinação linear de $\mathbf{v}_1$ e $\mathbf{v}_2$, e por isso pertence a $\mathcal{P}$.

Agora, considere $c$ um escalar. Então:

$$c\mathbf{u} = c(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = (cc_1)\mathbf{v}_1 + (cc_2)\mathbf{v}_2$$

que mostra que $c\mathbf{u}$ é também uma combinação linear de $\mathbf{v}_1$ e $\mathbf{v}_2$ e está, portanto, em $\mathcal{P}$. Mostramos que $\mathcal{P}$ satisfaz as propriedades de (1) a (3), e por essa razão é um subespaço de $\mathbb{R}^3$.

Se você olhar cuidadosamente os detalhes do Exemplo 1, notará que o fato de $\mathbf{v}_1$ e $\mathbf{v}_2$ serem vetores em $\mathbb{R}^3$ não tem influência nenhuma na verificação das propriedades. Dessa maneira, o método algébrico que utilizamos deve generalizar além de $\mathbb{R}^3$ e ser aplicável em situações que não mais podemos visualizar na geometria. Ele o é. Além disso, o método do Exemplo 1 pode servir como "modelo" para situações mais gerais. Quando generalizamos o Exemplo 1 para um conjunto arbitrário de vetores em $\mathbb{R}^n$, o resultado é suficientemente importante para ser chamado de teorema.

### TEOREMA 1

Sejam $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ vetores em $\mathbb{R}^n$. Então, ger$(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ é um subespaço de $\mathbb{R}^n$.

DEMONSTRAÇÃO: Seja $S = \text{ger}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$. Para verificar a propriedade (1) da definição, simplesmente observamos que o vetor $\mathbf{0}$ está em $S$, já que $\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \ldots + 0\mathbf{v}_k$.

Agora, considere

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \ \text{ e } \ \mathbf{v} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_k\mathbf{v}_k$$

como dois vetores em $S$. Então:

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= (c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k) + (d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_k\mathbf{v}_k)\\ &= (c_1 + d_1)\mathbf{v}_1 + (c_2 + d_2)\mathbf{v}_2 + \cdots + (c_k + d_k)\mathbf{v}_k\end{aligned}$$

Dessa maneira, $\mathbf{u} + \mathbf{v}$ é uma combinação linear de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$; logo, está em $S$. Isso verifica a propriedade (2).

Para mostrar a propriedade (3), considere $c$ um escalar. Então:

$$\begin{aligned}c\mathbf{u} &= c(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k)\\ &= (cc_1)\mathbf{v}_1 + (cc_2)\mathbf{v}_2 + \ldots + (cc_k)\mathbf{v}_k\end{aligned}$$

que mostra que $c\mathbf{u}$ é também uma combinação linear de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ e está, portanto, em $S$. Mostramos que $S$ satisfaz as propriedades de (1) a (3), e por essa razão é um subespaço de $\mathbb{R}^n$. ◈

Vamos nos referir a ger$(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ como o ***subespaço gerado por*** $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. Podemos freqüentemente economizar bastante trabalho reconhecendo quando o Teorema 1 pode ser aplicado.

**EXEMPLO 2** Mostre que o conjunto de todos os vetores $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ que satisfazem as condições $x = 3y$ e $z = -2y$ forma um subespaço de $\mathbb{R}^3$.

SOLUÇÃO: Substituindo as duas condições em $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$, obtemos

$$\begin{bmatrix} 3y \\ y \\ -2y \end{bmatrix} = y\begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$$

Como $y$ é arbitrário, o conjunto dado de vetores é ger$\left(\begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}\right)$ e é, desse modo, um subespaço de $\mathbb{R}^3$, pelo Teorema 1.

Geometricamente, o conjunto de vetores do Exemplo 2 representa a reta que passa pela origem em $\mathbb{R}^3$ com o vetor diretor $\begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$.

**EXEMPLO 3** Determine se o conjunto de todos os vetores $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ que satisfazem as condições $x = 3y + 1$ e $z = -2y$ é um subespaço de $\mathbb{R}^3$.

SOLUÇÃO: Desta vez, temos todos os vetores na forma

$$\begin{bmatrix} 3y + 1 \\ y \\ -2y \end{bmatrix}$$

O vetor nulo não possui essa forma. (Por que não? Tente resolver $\begin{bmatrix} 3y + 1 \\ y \\ -2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.) A propriedade (1), portanto, não é válida; esse conjunto não pode ser um subespaço de $\mathbb{R}^3$.

**EXEMPLO 4** Determine se o conjunto de todos os vetores $\begin{bmatrix} x \\ y \end{bmatrix}$, em que $y = x^2$, é um subespaço de $\mathbb{R}^2$.

SOLUÇÃO: Esses são os vetores da forma $\begin{bmatrix} x \\ x^2 \end{bmatrix}$ — chame esse conjunto de $S$. Desta vez, $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ pertence a $S$ (assuma $x = 0$), por isso a propriedade (1) é válida. Sejam $\mathbf{u} = \begin{bmatrix} x_1 \\ x_1^2 \end{bmatrix}$ e $\mathbf{v} = \begin{bmatrix} x_2 \\ x_2^2 \end{bmatrix}$, pertencentes a $S$. Então:

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} x_1 + x_2 \\ x_1^2 + x_2^2 \end{bmatrix}$$

que, em geral, não está em $S$, já que não tem a forma correta, ou seja, $x_1^2 + x_2^2 \neq (x_1 + x_2)^2$. Para sermos específicos, procuramos um contra-exemplo. Se

$$\mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{e} \quad \mathbf{v} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

então $\mathbf{u}$ e $\mathbf{v}$ estão em $S$, mas sua soma $\mathbf{u} + \mathbf{v} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$ não está em $S$, já que $5 \neq 3^2$. A propriedade (2), portanto, falha e $S$ não é um subespaço de $\mathbb{R}^2$.

Observação:

- Para que um conjunto $S$ seja um subespaço de algum $\mathbb{R}^n$, é necessário que se *prove* que as propriedades de (1) a (3) são confirmadas *genericamente*. Entretanto, para $S$ não ser um subespaço de $\mathbb{R}^n$, é suficiente mostrar que apenas uma das três propriedades não é confirmada. Normalmente, o caminho mais fácil é achar um único e específico *contra-exemplo* para ilustrar a falha da propriedade. Feito isso, não há necessidade de considerar as outras propriedades.

## *Subespaços Associados a Matrizes*

Muitos exemplos de subespaços existem no contexto de matrizes. Já encontramos o mais importante deles no Capítulo 2; agora, nós os revisamos com a noção de subespaço.

**Definição** Seja $A$ uma matriz $m \times n$.

1. O ***espaço linha*** de $A$ é o subespaço lin($A$) de $\mathbb{R}^n$ gerado pelas linhas de $A$.
2. O ***espaço coluna*** de $A$ é o subespaço col($A$) de $\mathbb{R}^m$ gerado pelas colunas de $A$.

**EXEMPLO 5** Considere a matriz

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

(a) Determine se $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ está no espaço coluna de $A$.

(b) Determine se $\mathbf{w} = [4 \quad 5]$ está no espaço linha de $A$.

(c) Descreva lin($A$) e col($A$).

SOLUÇÃO: (a) Pelo Teorema 1 da Seção 2.4 e sua discussão precedente, $\mathbf{b}$ é a combinação linear das colunas de $A$ se, e somente se, o sistema linear $A\mathbf{x} = \mathbf{b}$ for possível. Escalonamos por linhas a matriz completa da seguinte forma:

$$\left[\begin{array}{cc|c} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 3 & -3 & 3 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right]$$

O sistema, portanto, é possível (e, de fato, tem uma única solução), e $\mathbf{b}$ está em col($A$). (Este é simplesmente o Exemplo 1 da Seção 2.4, utilizando-se à terminologia desta seção.)

(b) Como vimos também na Seção 2.4, operações elementares com linhas simplesmente produzem combinações lineares das linhas de uma matriz — elas produzem vetores apenas no espaço linha da matriz. Se o vetor $\mathbf{w}$ está em lin($A$), então $\mathbf{w}$ é uma combinação linear das linhas de $A$. Logo, se completarmos $A$ com $\mathbf{w}$ como

$\left[\begin{array}{c} A \\ \hline \mathbf{w} \end{array}\right]$, será possível aplicarmos operações elementares com linhas a essa matriz completa para reduzi-la à forma $\left[\begin{array}{c} A' \\ \hline \mathbf{0} \end{array}\right]$ *sem fazer nenhuma permutação envolvendo a última linha.* (Por quê?)

Neste exemplo, temos

$$\left[\begin{array}{c} A \\ \hline \mathbf{w} \end{array}\right] = \left[\begin{array}{cc} 1 & -1 \\ 0 & 1 \\ 3 & -3 \\ \hline 4 & 5 \end{array}\right] \xrightarrow[L_4-4L_1]{L_3-3L_1} \left[\begin{array}{cc} 1 & -1 \\ 0 & 1 \\ 0 & 0 \\ \hline 0 & 9 \end{array}\right] \xrightarrow{L_4-9L_2} \left[\begin{array}{cc} 1 & -1 \\ 0 & 1 \\ 0 & 0 \\ \hline 0 & 0 \end{array}\right]$$

Portanto, **w** é uma combinação linear das linhas de $A$ (de fato, esses cálculos mostram que $\mathbf{w} = 4[1\ -1] + 9[0\ \ 1]$ — como?), e, portanto, **w** está em lin($A$).

(c) É fácil verificar que, para qualquer vetor $\mathbf{w} = [x\ \ y]$, a matriz completa $\left[\dfrac{A}{\mathbf{w}}\right]$ é reduzida a

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ \hline 0 & 0 \end{bmatrix}$$

sem permutação das linhas envolvendo a última linha. Portanto, todo vetor de $\mathbb{R}^2$ está em lin($A$), por isso lin($A$) $= \mathbb{R}^2$.

Achar col($A$) é um processo idêntico à solução do Exemplo 4 da Seção 2.4, no qual determinamos que ele coincide com o plano (que passa pela origem) em $\mathbb{R}^3$ com equação $3x - z = 0$. (Descobriremos logo outras maneiras de responder a esse tipo de questão.)

Comentário: Poderíamos também ter respondido à parte (b) e à primeira parte da parte (c) observando que qualquer pergunta sobre as *linhas* de $A$ é a questão correspondente sobre as *colunas* de $A^T$. Então, para exemplificar, **w** está em lin($A$) se, e somente se, $\mathbf{w}^T$ estiver na col($A^T$). Isso é verdade se, e somente se, o sistema $A^T\mathbf{x} = \mathbf{w}^T$ for consistente. Podemos agora proceder como na parte (a). (Ver Exercícios de 17 a 20.)

As observações que fizemos sobre a relação entre as operações elementares com linhas e o espaço linha estão sumarizadas no seguinte teorema:

### ◈ TEOREMA 2

Seja $B$ qualquer matriz linha equivalente a uma matriz $A$. Então: lin($B$) = lin($A$).

DEMONSTRAÇÃO: A matriz $A$ pode ser transformada em $B$ por uma seqüência de operações em suas linhas. Conseqüentemente, as linhas de $B$ são combinações lineares das linhas de $A$; por essa razão, combinações lineares das linhas de $B$ são combinações lineares das linhas de $A$. (Veja o Exercício 21 na Seção 2.4.) Segue que lin($B$) $\subseteq$ lin($A$).

Por outro lado, a reversão dessas operações nas linhas transforma $B$ em $A$. Portanto, o argumento do parágrafo anterior mostra que lin($A$) $\subseteq$ lin($B$). Combinando esses resultados, temos que lin($A$) = lin($B$). ◈

Existe outro subespaço importante que já encontramos: o conjunto de soluções de um sistema homogêneo de equações lineares. É fácil provar que esse subespaço satisfaz as três propriedades de subespaço.

### ◈ TEOREMA 3

Sejam $A$ uma matriz $m \times n$ e $N$ o conjunto de soluções do sistema linear homogêneo $A\mathbf{x} = \mathbf{0}$. Então, $N$ é subespaço de $\mathbb{R}^n$.

DEMONSTRAÇÃO: [Note que **x** precisa ser um vetor (coluna) em $\mathbb{R}^n$ para que $A\mathbf{x}$ seja definida, e que $\mathbf{0} = \mathbf{0}_m$ é o vetor nulo de $\mathbb{R}^m$.] Já que $A\mathbf{0}_n = \mathbf{0}_m$, $\mathbf{0}_n$ está em $N$. Agora, estejam **u** e **v** em $N$. Com isso, $A\mathbf{u} = \mathbf{0}$ e $A\mathbf{v} = \mathbf{0}$. Segue que

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

Por essa razão, $\mathbf{u} + \mathbf{v}$ está em $N$. Finalmente, para qualquer escalar $c$,

$$A(c\mathbf{u}) = c(A\mathbf{u}) = c\mathbf{0} = \mathbf{0}$$

e $c\mathbf{u}$ também está em $N$. Conseqüentemente, $N$ é subespaço de $\mathbb{R}^n$. ◈

**Definição** Seja $A$ uma matriz $m \times n$. O ***espaço anulado*** por $A$ é o subespaço de $\mathbb{R}^n$ que consiste nas soluções do sistema linear homogêneo $A\mathbf{x} = \mathbf{0}$. É denotado por anul($A$).

O fato de um espaço anulado por uma matriz ser um subespaço nos permite provar o que a intuição e os exemplos nos levaram a entender sobre as soluções de sistemas lineares: eles têm ou nenhuma solução, ou uma única solução ou infinitas soluções.

### ◈ TEOREMA 4

Seja $A$ uma matriz cujos elementos são números reais. Para qualquer sistema de equações lineares $A\mathbf{x} = \mathbf{b}$, apenas uma das seguintes afirmações é verdadeira:
a. Não há solução.
b. Há uma única solução.
c. Há infinitas soluções.

À primeira vista, não fica muito claro como devemos proceder para provar esse teorema. Uma pequena reflexão deve persuadi-lo de que a questão é provar que, se (a) e (b) não são verdadeiros, (c) é a única possibilidade. Se existe mais de uma solução, não é possível que existam apenas duas ou um número finito — tem de haver infinitas soluções.

DEMONSTRAÇÃO: Se o sistema $A\mathbf{x} = \mathbf{b}$ tem exatamente uma solução, ou nenhuma solução, estamos feitos. Assuma, então, que existam pelo menos duas soluções distintas de $A\mathbf{x} = \mathbf{b}$ — digamos, $\mathbf{x}_1$ e $\mathbf{x}_2$. Assim,

$$A\mathbf{x}_1 = \mathbf{b} \ \text{ e } \ A\mathbf{x}_2 = \mathbf{b}$$

com $\mathbf{x}_1 \neq \mathbf{x}_2$. Segue que

$$A(\mathbf{x}_1 - \mathbf{x}_2) = A\mathbf{x}_1 - A\mathbf{x}_2 = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

Chame $\mathbf{x}_0 = \mathbf{x}_1 - \mathbf{x}_2$. Logo, $\mathbf{x}_0 \neq \mathbf{0}$ e $A\mathbf{x}_0 = \mathbf{0}$. Por essa razão, o espaço anulado por $A$ é não trivial, e, já que anul($A$) é fechado para multiplicação por escalar, $c\mathbf{x}_0$ está em anul($A$) para qualquer escalar $c$. Conseqüentemente, o espaço anulado por $A$ contém um número infinito de vetores (já que contém *pelo menos* todo vetor da forma $c\mathbf{x}_0$, e existem infinitos deste).

Agora, considere os vetores (infinitos) da forma $\mathbf{x}_1 + c\mathbf{x}_0$, com $c$ variando no conjunto dos números reais. Temos

$$A(\mathbf{x}_1 + c\mathbf{x}_0) = A\mathbf{x}_1 + cA\mathbf{x}_0 = \mathbf{b} + c\mathbf{0} = \mathbf{b}$$

Portanto, existem infinitas soluções para a equação $A\mathbf{x} = \mathbf{b}$. ◈

## Base

Podemos extrair um pouco mais da idéia intuitiva de que subespaços são generalizações de planos que passam pela origem em $\mathbb{R}^3$. Um plano é gerado por quaisquer dois vetores paralelos ao plano, mas não paralelos entre si. Na linguagem algébrica, esses dois vetores geram o plano e são linearmente independentes. Não dá certo com menos de dois vetores, e mais de dois vetores não são necessários. Essa é a essência de uma *base* para um subespaço.

**Definição** Uma ***base*** de um subespaço de $\mathbb{R}^n$ é um conjunto de vetores em $S$ que

1. Gera $S$.
2. É linearmente independente.

**EXEMPLO 6** Na Seção 2.4, vimos que os vetores canônicos unitários $\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_n$ em $\mathbb{R}^n$ são linearmente independentes e geram $\mathbb{R}^n$. Portanto, eles formam uma base para $\mathbb{R}^n$, chamada de ***base canônica***.

**EXEMPLO 7** No Exemplo 2 da Seção 2.4, mostramos que $\mathbb{R}^2 = \text{ger}\left(\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}\right)$. Já que $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ e $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ são também linearmente independentes (pois não são múltiplos), eles formam uma base de $\mathbb{R}^2$.

Um subespaço pode (e vai) ter mais de uma base. Por exemplo, acabamos de ver que $\mathbb{R}^2$ tem a base canônica $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$ e a base $\left\{\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}\right\}$. De qualquer modo, provaremos em breve que o *número* de vetores em uma base para um dado subespaço será sempre o mesmo.

**EXEMPLO 8** Ache uma base para $S = \text{ger}(\mathbf{u}, \mathbf{v}, \mathbf{w})$, onde

$$\mathbf{u} = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \quad \text{e} \quad \mathbf{w} = \begin{bmatrix} 0 \\ -5 \\ 1 \end{bmatrix}$$

SOLUÇÃO: Os vetores $\mathbf{u}$, $\mathbf{v}$ e $\mathbf{w}$ já geram $S$, por isso serão uma base de $S$ se também forem linearmente independentes. É fácil determinar que eles não são; de fato, $\mathbf{w} = 2\mathbf{u} - 3\mathbf{v}$. Portanto, podemos ignorar $\mathbf{w}$, já que qualquer combinação linear envolvendo $\mathbf{u}$, $\mathbf{v}$ e $\mathbf{w}$ pode ser reescrita envolvendo apenas $\mathbf{u}$ e $\mathbf{v}$. (Veja também o Exercício 47 da Seção 2.4.) Isso implica que $S = \text{ger}(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \text{ger}(\mathbf{u}, \mathbf{v})$, e, já que certamente $\mathbf{u}$ e $\mathbf{v}$ são linearmente independentes (por quê?), eles formarão uma base para $S$. (Geometricamente, isso significa que $\mathbf{u}$, $\mathbf{v}$ e $\mathbf{w}$ pertencem ao mesmo plano, e $\mathbf{u}$ e $\mathbf{v}$ servem como um conjunto de vetores diretores para esse plano.)

**EXEMPLO 9** Ache uma base do espaço linha de

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

SOLUÇÃO: A forma escalonada reduzida por linhas de $A$ é

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

Pelo Teorema 2, $\text{lin}(A) = \text{lin}(R)$, então é suficiente achar uma base do espaço linha de $R$. Mas $\text{lin}(R)$ é claramente gerado pelas linhas não nulas, e é fácil verificar que a forma escalonada força as três primeiras linhas de $R$ a serem linearmente independentes. (Isso é um fato geral, que você precisará verificar para provar o Exercício 29.) Portanto, uma base do espaço linha de $A$ é

$$\{[1\ \ 0\ \ 1\ \ 0\ \ -1],\ [0\ \ 1\ \ 2\ \ 0\ \ 3],\ [0\ \ 0\ \ 0\ \ 1\ \ 4]\}$$

Podemos utilizar o método do Exemplo 9 para achar uma base do subespaço gerado por um conjunto de vetores dado.

**EXEMPLO 10** Refaça o Exemplo 8 usando o método do Exemplo 9.

SOLUÇÃO: Consideramos os transpostos de $\mathbf{u}$, $\mathbf{v}$ e $\mathbf{w}$ para obter vetores linha e então formar uma matriz com esses vetores como suas linhas:

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

Continuando como no Exemplo 9, reduzimos $B$ à sua forma escalonada reduzida por linhas

$$\begin{bmatrix} 1 & 0 & \frac{8}{5} \\ 0 & 1 & -\frac{1}{5} \\ 0 & 0 & 0 \end{bmatrix}$$

e usamos os vetores de linha não nulos como uma base para o espaço linha. Já que começamos com vetores coluna, precisamos transpor de novo. Desse modo, uma base para $\text{ger}(\mathbf{u}, \mathbf{v}, \mathbf{w})$ é

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ \frac{8}{5} \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -\frac{1}{5} \end{bmatrix} \right\}$$

Observações:

- De fato, não é necessário chegar até a forma escalonada *reduzida* por linhas — a forma escalonada por linhas já é o suficiente. Se $U$ é uma forma escalonada reduzida por linhas de $A$, então os vetores linha não nulos de $U$ formarão uma base para $\text{lin}(A)$ (veja o Exercício 29). Esse enfoque tem a vantagem de (freqüentemente) nos permitir evitar o uso de frações. No Exemplo 10, $B$ pode ser reduzida a

$$U = \begin{bmatrix} 3 & -1 & 5 \\ 0 & -5 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

que nos dá a base

$$\left\{ \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ -5 \\ 1 \end{bmatrix} \right\}$$

para $\text{ger}(\mathbf{u}, \mathbf{v}, \mathbf{w})$.

- Observe que os métodos usados no Exemplo 8, no Exemplo 10 e nos comentários que acabamos de ver irão geralmente produzir bases diferentes.

Veremos agora o problema de achar uma base para o espaço coluna de uma matriz $A$. Um método consiste em simplesmente transpor a matriz: os vetores coluna de $A$ se transformam nos vetores linha de $A^T$, e podemos aplicar o método do Exemplo 9 para achar uma base para $\text{lin}(A^T)$. Transpondo esses vetores, teremos uma base para $\text{col}(A)$. (Você fará isso nos Exercícios de 17 a 20.) Essa abordagem, entretanto, requer a realização de um novo conjunto de operações com linhas em $A^T$.

Em vez disso, preferimos utilizar uma abordagem que nos permita usar a forma escalonada reduzida de $A$ que já calculamos. Lembre-se de que um produto $A\mathbf{x}$ de uma matriz e um vetor corresponde a uma combinação linear das colunas de $A$ com as entradas de $\mathbf{x}$ como coeficientes. Dessa maneira, uma solução não trivial para $A\mathbf{x} = \mathbf{0}$ representa uma *relação de dependência* entre as colunas da $A$. Já que as operações elementares com linhas não afetam o conjunto solução do sistema, se $A$ é linha equivalente a $R$, *as colunas de $A$ têm as mesmas relações de dependência que as colunas de $R$*. Essa observação importante é a base (nenhum trocadilho intencional!) para a técnica que usamos agora para achar uma base para $\text{col}(A)$.

**EXEMPLO 11** Encontre uma base do espaço coluna da matriz do Exemplo 9,

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

SOLUÇÃO: Sejam $\mathbf{a}_i$ um vetor coluna de $A$ e $\mathbf{r}_i$ um vetor coluna da forma escalonada reduzida

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

Podemos verificar facilmente que $\mathbf{r}_3 = \mathbf{r}_1 + 2\mathbf{r}_2$ e $\mathbf{r}_5 = -\mathbf{r}_1 + 3\mathbf{r}_2 + 4\mathbf{r}_4$. (Verifique que, como previsto, os vetores coluna de $A$ correspondentes satisfazem as mesmas relações de dependência.) Dessa maneira, $\mathbf{r}_3$ e $\mathbf{r}_5$ não contribuem em nada para $\text{col}(R)$. Os vetores coluna restantes, $\mathbf{r}_1$, $\mathbf{r}_2$ e $\mathbf{r}_4$, são linearmente independentes, por serem apenas vetores unitários canônicos. As afirmações correspondentes são, portanto, verdadeiras para os vetores coluna de $A$.

Dessa maneira, dentre os vetores coluna de $A$, eliminamos os dependentes ($\mathbf{a}_3$ e $\mathbf{a}_5$) e os restantes serão linearmente independentes, por isso formarão uma base para $\text{col}(A)$. Qual é o modo mais rápido de achar uma base? Use as colunas de $A$ que correspondem às colunas de $R$ que *contêm os 1 líderes*. Uma base para $\text{col}(A)$ é

$$\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_4\} = \left\{ \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -2 \\ 1 \end{bmatrix} \right\}$$

Aviso: As operações elementares com linhas alteram o espaço coluna! Em nosso exemplo, $\text{col}(A) \neq \text{col}(R)$, já que todo vetor em $\text{col}(R)$ tem seu quarto componente igual a 0, mas isso certamente não é verdade para $\text{col}(A)$. Então, precisamos voltar à matriz original $A$ para obter os vetores coluna para uma base de $\text{col}(A)$. Para ser mais específico, no Exemplo 11, $\mathbf{r}_1$, $\mathbf{r}_2$ e $\mathbf{r}_4$ *não* formam uma base para o espaço coluna de $A$.

**12** Encontre uma base do espaço anulado pela matriz $A$ do Exemplo 11.

Na verdade, não há nada de novo aqui a não ser a terminologia. Simplesmente temos de encon-
rever as soluções do sistema homogêneo $A\mathbf{x} = \mathbf{0}$. Já temos a forma escalonada reduzida $R$ de $A$,
o que resta ser feito pelo método de eliminação de Gauss-Jordan é expressar as variáveis que li-
termos das três variáveis. A matriz completa final é

$$[R \mid \mathbf{0}] = \left[\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

1 líderes estão nas colunas 1, 2 e 4, por isso expressamos $x_1$, $x_2$ e $x_4$ em termos das variáveis livres $x_3$
temos $x_1 = -x_3 + x_5$, $x_2 = -2x_3 - 3x_5$ e $x_4 = -4x_5$. Colocando $x_3 = s$ e $x_5 = t$, temos

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -s + t \\ -2s - 3t \\ s \\ -4t \\ t \end{bmatrix} = s\begin{bmatrix} -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 1 \\ -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} = s\mathbf{u} + t\mathbf{v}$$

forma, $\mathbf{u}$ e $\mathbf{v}$ geram anul($A$), e, por eles serem linearmente independentes, formam uma base de

eguir, apresentamos um resumo do procedimento mais efetivo para achar bases para o espaço linha, o
coluna e a espaço anulado por uma matriz $A$.

1. Encontrar a forma escalonada reduzida $R$ de $A$.
2. Use os vetores linha não nulos de $R$ (que contenham os 1 líderes) para formar uma base de lin($A$).
3. Use os vetores coluna de $A$ correspondentes às colunas de $R$ que contenham os 1 líderes (colunas pivô) para formar uma base de col($A$).
4. Expresse as variáveis líderes de $R\mathbf{x} = \mathbf{0}$ em termos das variáveis livres, coloque as variáveis livres iguais a parâmetros, substitua novamente em $\mathbf{x}$ e escreva o resultado como uma combinação linear de vetores $f$ (em que $f$ é o número de variáveis livres). Esses vetores $f$ formam uma base para anul($A$).

uando não há necessidade de achar o espaço anulado, é mais rápido apenas reduzir $A$ à sua forma
onada por linhas para achar bases dos seus espaços linha e coluna. Os passos 2 e 3 que acabamos de ver
nuam válidos (com a substituição da palavra "pivô" no lugar de "1 líderes").

## *Dimensão e Posto*

Observamos que, embora um subespaço tenha diferentes bases, cada base tem o mesmo número de vetores. Esse fato fundamental será de vital importância a partir daqui neste livro.

**◈ TEOREMA 5 O Teorema da Base**

Seja $S$ um subespaço de $\mathbb{R}^n$. Logo, quaisquer duas bases em $S$ terão o mesmo número de vetores.

DEMONSTRAÇÃO: Sejam $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ e $\mathcal{C} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s\}$ bases em $S$. Precisamos provar que $r = s$. Fazemos isso mostrando que nenhuma das duas possibilidades, $r < s$ ou $r > s$, pode ocorrer.[4]

Admita que $r < s$. Mostraremos que isso força $\mathcal{C}$ a ser um conjunto de vetores linearmente dependentes. Com essa finalidade, seja

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_s\mathbf{v}_s = \mathbf{0} \tag{1}$$

Já que $\mathcal{B}$ é uma base para $S$, podemos escrever cada $\mathbf{v}_i$ como uma combinação linear dos elementos $\mathbf{u}_j$:

$$\begin{aligned}
\mathbf{v}_1 &= a_{11}\mathbf{u}_1 + a_{12}\mathbf{u}_2 + \ldots + a_{1r}\mathbf{u}_r \\
\mathbf{v}_2 &= a_{21}\mathbf{u}_1 + a_{22}\mathbf{u}_2 + \ldots + a_{2r}\mathbf{u}_r \\
&\vdots \\
\mathbf{v}_s &= a_{s1}\mathbf{u}_1 + a_{s2}\mathbf{u}_2 + \ldots + a_{sr}\mathbf{u}_r
\end{aligned} \tag{2}$$

Substituindo as equações (2) em (1), obtemos

$$c_1(a_{11}\mathbf{u}_1 + \ldots + a_{1r}\mathbf{u}_r) + c_2(a_{21}\mathbf{u}_1 + \ldots + a_{2r}\mathbf{u}_r) + \ldots + c_s(a_{s1}\mathbf{u}_1 + \ldots + a_{sr}\mathbf{u}_r) = \mathbf{0}$$

Reagrupando, temos

$$(c_1a_{11} + c_2a_{21} + \ldots + c_sa_{s1})\mathbf{u}_1 + (c_1a_{12} + c_2a_{22} + \ldots + c_sa_{s2})\mathbf{u}_2 + \ldots + (c_1a_{1r} + c_2a_{2r} + \ldots + c_sa_{sr})\mathbf{u}_r = \mathbf{0}$$

Agora, já que $\mathcal{B}$ é uma base, os $\mathbf{u}_j$'s são linearmente independentes. Logo, cada expressão dentro dos parênteses deve ser igual a zero:

$$\begin{aligned}
c_1a_{11} + c_2a_{21} + \cdots + c_sa_{s1} &= 0 \\
c_1a_{12} + c_2a_{22} + \cdots + c_sa_{s2} &= 0 \\
&\vdots \\
c_1a_{1r} + c_2a_{2r} + \cdots + c_sa_{sr} &= 0
\end{aligned}$$

Esse é um sistema homogêneo de $r$ equações lineares nas $s$ variáveis $c_1, c_2, \ldots, c_s$. (O fato de as variáveis aparecerem à esquerda do coeficiente não faz diferença nenhuma.) Já que $r < s$, sabemos, pelo Teorema 3 da Seção 2.3, que existem infinitas soluções. Em particular, existe uma solução não trivial, que resulta em uma relação de dependência não trivial na equação (1). Desse modo, $\mathcal{C}$ é um conjunto de vetores linearmente dependentes. Mas isso contradiz o fato de $\mathcal{C}$ ser uma base e, portanto, linearmente independente. Concluímos que $r < s$ não é possível. Analogamente (invertendo os papéis de $\mathcal{B}$ e $\mathcal{C}$), vemos que $r > s$ nos leva a uma contradição. Por isso, precisamos ter $r = s$, como desejado. ◈

Como todas as bases de um subespaço devem ter o mesmo número de vetores, fixaremos um nome para esse número.

---

[4] Como Sherlock Holmes dizia: "Quando você elimina o impossível, o que sobra, mesmo que improvável, necessariamente é a verdade" (em *O signo dos quatro*, de *Sir* Arthur Conan Doyle).

**Definição** Se $S$ é um subespaço de $\mathbb{R}^n$, o número de vetores em uma base de $S$ é chamado de ***dimensão*** de $S$, denotado como $\dim S$.

Observação: • O vetor nulo $\mathbf{0}$ por ele mesmo é sempre um subespaço de $\mathbb{R}^n$. (Por quê?) Porém, qualquer conjunto que contenha o vetor nulo (e, em particular, $\{\mathbf{0}\}$) é linearmente dependente, por isso $\{\mathbf{0}\}$ não pode ter uma base. Definimos $\dim\{\mathbf{0}\}$ como 0.

**EXEMPLO 13** Já que a base canônica de $\mathbb{R}^n$ tem $n$ vetores, $\dim \mathbb{R}^n = n$. (Perceba que esse resultado concorda com nosso entendimento intuitivo de dimensão para $n \le 3$.)

**EXEMPLO 14** Nos Exemplos de 9 a 12, achamos que lin($A$) tem uma base com três vetores, col($A$) tem uma base com três vetores e anul($A$) tem uma base com dois vetores. Por essa razão, dim (lin($A$)) = 3, dim (col($A$)) = 3 e dim (anul($A$)) = 2.

Um único exemplo não é o suficiente para especularmos, mas o fato de os espaços linha e coluna, no Exemplo 14, terem a mesma dimensão não é acidental. Nem mesmo o fato de a soma de dim (col($A$)) e dim (anul($A$)) ser 5, o número de colunas de $A$. Provaremos agora que essas relações são verdadeiras em geral.

### TEOREMA 6

Os espaços linha e coluna de uma matriz $A$ têm a mesma dimensão.

DEMONSTRAÇÃO: Seja $R$ a forma escalonada reduzida por linhas de $A$. Pelo Teorema 2, lin($A$) = lin($R$), então

$$\begin{aligned}\dim(\text{lin}(A)) &= \dim(\text{lin}(R))\\ &= \text{número de linhas não nulas}\\ &= \text{número de 1 líderes de } R\end{aligned}$$

Chame esse número de $r$.

Agora, col($A$) $\neq$ col($R$), mas as colunas de $A$ e $R$ possuem as mesmas relações de dependência. Portanto, dim (col($A$)) = dim (col($R$)). Já que existem $r$ 1 líderes, $R$ tem $r$ colunas que são vetores canônicos unitários, $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_r$. (Estes serão vetores em $\mathbb{R}^m$ se $A$ e $R$ forem matrizes $m \times n$.) Esses $r$ vetores são linearmente independentes, e as colunas restantes de $R$ são combinações lineares deles. Dessa maneira, dim (col($R$)) = $r$. Segue que dim (lin($A$)) = $r$ = dim (col($A$)), como queríamos demonstrar. ◆

O posto de uma matriz foi definido pela primeira vez por Georg Frobenius (1849–1917), embora ele o tenha definido utilizando determinantes, e não como fizemos aqui. (Veja o Capítulo 4.) Frobenius foi um matemático alemão que obteve seu doutorado e ensinou na Universidade de Berlim. Mais conhecido por suas contribuições à teoria de grupos, utilizava matrizes em seu trabalho em representação de grupos.

**Definição** O ***posto*** de uma matriz $A$, denotado por posto($A$), é a dimensão de seus espaços linha e coluna.

Para o Exemplo 14, podemos então escrever posto($A$) = 3.

Observações:
- A definição anterior concorda com a definição mais informal de posto introduzida no Capítulo 2. A vantagem de nossa nova definição é que ela é muito mais flexível.
- O posto de uma matriz nos dá simultaneamente informações sobre a dependência linear entre os vetores linha da matriz *e* entre seus vetores coluna. Particularmente, nos diz o número de linhas e colunas linearmente independentes (e esse número é o mesmo em cada caso!).

Já que os vetores linha de $A$ são os vetores coluna de $A^T$, o Teorema 6 tem o seguinte corolário imediato:

### ◆ TEOREMA 7

Para qualquer matriz $A$,

$$\text{posto}(A^T) = \text{posto}(A)$$

DEMONSTRAÇÃO: Temos

$$\begin{aligned}\text{posto}(A^T) &= \dim(\text{col}(A^T)) \\ &= \dim(\text{lin}(A)) \\ &= \text{posto}(A)\end{aligned}$$

◆

**Definição** A ***nulidade*** de uma matriz $A$, denotada por nulidade($A$), é a dimensão do espaço anulado por $A$.

Em outras palavras, nulidade($A$) é a dimensão do espaço solução de $A\mathbf{x} = \mathbf{0}$, que é o mesmo número de variáveis livres na solução do sistema. Podemos agora rever o Teorema do Posto (Teorema 2 da Seção 2.3) em função das nossas novas definições.

### ◆ TEOREMA 8 O Teorema do Posto

Se $A$ é uma matriz $m \times n$, então

$$\text{posto}(A) + \text{nulidade}(A) = n$$

DEMONSTRAÇÃO: Considere $R$ a forma escalonada reduzida por linhas de $A$ e suponha que posto($A$) $= r$. Então, $R$ tem $r$ 1 líderes, por isso há $r$ variáveis dependentes e $n - r$ variáveis livres na solução para $A\mathbf{x} = \mathbf{0}$. Como $\dim(\text{anul}(A)) = n - r$, temos

$$\begin{aligned}\text{posto}(A) + \text{nulidade}(A) &= r + (n - r) \\ &= n\end{aligned}$$

◆

Geralmente, quando precisamos saber a nulidade de uma matriz, não é necessário conhecer o conjunto solução de $A\mathbf{x} = \mathbf{0}$. O Teorema do Posto é extremamente útil em tais situações, como o exemplo a seguir ilustra.

**EXEMPLO 15** Encontre a nulidade de cada uma das seguintes matrizes:

$$M = \begin{bmatrix} 2 & 3 \\ 1 & 5 \\ 4 & 7 \\ 3 & 6 \end{bmatrix} \quad \text{e}$$

$$N = \begin{bmatrix} 2 & 1 & -2 & -1 \\ 4 & 4 & -3 & 1 \\ 2 & 7 & 1 & 8 \end{bmatrix}$$

SOLUÇÃO: Já que as duas colunas de $M$ são, sem dúvida, linearmente independentes, posto$(M) = 2$. Assim, pelo Teorema do Posto, nulidade$(M) = 2 - $ posto$(M) = 2 - 2 = 0$.

Não há nenhuma dependência óbvia entre as linhas ou colunas de $N$, por isso aplicamos operações em suas linhas para reduzir $N$ a

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

Reduzimos a matriz o suficiente (não precisamos da forma escalonada *reduzida* por linhas aqui, já que não estamos procurando uma base do espaço anulado). Vemos que existem apenas duas linhas não nulas, logo, posto$(N) = 2$. Por isso, nulidade$(N) = 4 - $ posto$(N) = 4 - 2 = 2$. ◈

Os resultados desta seção nos permitem estender o Teorema Fundamental das Matrizes Invertíveis (Teorema 7 da Seção 3.4).

### ◈ TEOREMA 9 O Teorema Fundamental das Matrizes Invertíveis: Versão 2

Seja $A$ uma matriz $n \times n$. As seguintes afirmações são equivalentes:

a. $A$ é invertível.
b. $A\mathbf{x} = \mathbf{b}$ tem uma solução única para todo $\mathbf{b}$ em $\mathbb{R}^n$.
c. $A\mathbf{x} = \mathbf{0}$ tem apenas a solução trivial.
d. A forma escalonada reduzida por linhas de $A$ é $I_n$.
e. $A$ é um produto de matrizes elementares.
f. posto$(A) = n$
g. nulidade$(A) = 0$
h. Os vetores coluna de $A$ são linearmente independentes.
i. Os vetores coluna de $A$ geram $\mathbb{R}^n$.
j. Os vetores coluna de $A$ formam uma base de $\mathbb{R}^n$.
k. Os vetores linha de $A$ são linearmente independentes.
l. Os vetores linha de $A$ geram $\mathbb{R}^n$.
m. Os vetores linha de $A$ formam uma base de $\mathbb{R}^n$.

DEMONSTRAÇÃO: Já verificamos a equivalência de (a) até (e). Precisamos mostrar que as afirmações de (f) a (m) são equivalentes às primeiras cinco.

(f) $\Leftrightarrow$ (g) Como posto$(A)$ + nulidade$(A) = n$ quando $A$ é uma matriz $n \times n$, segue, do Teorema do Posto, que posto$(A) = n$ se, e somente se, nulidade$(A) = 0$.

(f) $\Rightarrow$ (d) $\Rightarrow$ (c) $\Rightarrow$ (h) Se posto$(A) = n$, a forma escalonada reduzida por linhas de $A$ tem $n$ 1 líderes e é então $I_n$. De (d) $\Rightarrow$ (c) sabemos que $A\mathbf{x} = \mathbf{0}$ tem apenas a solução trivial, o que implica que os vetores coluna de $A$ são linearmente independentes, já que $A\mathbf{x}$ é uma combinação linear dos vetores coluna de $A$.

(h) $\Rightarrow$ (i) Se os vetores coluna de $A$ são linearmente independentes, $A\mathbf{x} = \mathbf{0}$ tem apenas a solução trivial. Dessa forma, por (c) $\Rightarrow$ (b), $A\mathbf{x} = \mathbf{b}$ tem uma única solução para cada $\mathbf{b}$ em $\mathbb{R}^n$. Isso significa que todo vetor $\mathbf{b}$ em $\mathbb{R}^n$ pode ser escrito como uma combinação linear dos vetores coluna de $A$, provando (i).

(i) $\Rightarrow$ (j) Se os vetores coluna de $A$ geram $\mathbb{R}^n$, $\text{col}(A) = \mathbb{R}^n$ por definição, logo, $\text{posto}(A) = \dim(\text{col}(A)) = n$. Isso é (f), e já provamos que (f) $\Rightarrow$ (h). Concluímos, então, que os vetores coluna de $A$ são linearmente independentes e formam uma base de $\mathbb{R}^n$, já que, por hipótese, eles também geram $\mathbb{R}^n$.

(j) $\Rightarrow$ (f) Se os vetores coluna de $A$ formam uma base de $\mathbb{R}^n$, então, em particular, eles são linearmente independentes. Segue que a forma escalonada reduzida por linhas de $A$ contém $n$ 1 líderes, e, dessa maneira, $\text{posto}(A) = n$.

A nulidade de uma matriz foi definida em 1884 por James Joseph Sylvester (1814–1887), que era interessado em invariantes — propriedades de matrizes que não se alteram sob certos tipos de transformação. Nascido na Inglaterra, Sylvester se tornou o segundo presidente da London Mathematical Society. Em 1878, enquanto ensinava na Universidade John Hopkins, em Baltimore, fundou o *American Journal Of Mathematics*, o primeiro jornal de matemática dos Estados Unidos.

Essa discussão mostra que (f) $\Rightarrow$ (d) $\Rightarrow$ (c) $\Rightarrow$ (h) $\Rightarrow$ (i) $\Rightarrow$ (j) $\Rightarrow$ (f) $\Leftrightarrow$ (g). Agora, lembre-se de que, pelo Teorema 7, $\text{posto}(A^T) = \text{posto}(A)$, por isso o que acabamos de provar nos dá resultados correspondentes sobre os vetores coluna de $A^T$. Esses são resultados sobre os vetores *linha* de $A$, trazendo (k), (l) e (m) para a rede de equivalências e completando a demonstração. ◆

Teoremas como o Teorema Fundamental não atingem interesse teórico apenas. Eles também são poderosos instrumentos "economizadores" de trabalho. O Teorema Fundamental já nos permitiu cortar pela metade o trabalho necessário para verificar que duas matrizes quadradas são inversas uma da outra. Também simplifica a tarefa de mostrar que alguns conjuntos de vetores são bases de $\mathbb{R}^n$. De fato, quando tivermos um conjunto de $n$ vetores em $\mathbb{R}^n$, ele será uma base de $\mathbb{R}^n$ se *apenas uma* das propriedades necessárias de independência linear ou de conjunto gerador for verdadeira. O exemplo seguinte nos mostra quão fáceis os cálculos podem se tornar.

**EXEMPLO 16** Mostre que os vetores

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 4 \\ 9 \\ 7 \end{bmatrix}$$

formam uma base de $\mathbb{R}^n$.

SOLUÇÃO: De acordo com o Teorema Fundamental, os vetores formarão uma base de $\mathbb{R}^n$ se, e somente se, a matriz que apresentar esses vetores como suas colunas (ou linhas) tiver posto 3. Fazemos então apenas o mínimo de operações nas linhas para determinar isso:

$$A = \begin{bmatrix} 1 & -1 & 4 \\ 2 & 0 & 9 \\ 3 & 1 & 7 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -1 & 4 \\ 0 & 2 & 1 \\ 0 & 0 & -7 \end{bmatrix}$$

Vemos que $A$ tem posto 3, logo, os vetores dados são uma base de $\mathbb{R}^3$ pela equivalência entre (f) e (j).

O próximo teorema é uma aplicação do Teorema do Posto e do Teorema Fundamental. Precisaremos desse resultado nos Capítulos 5 e 7.

### ◈ TEOREMA 10

Seja $A$ uma matriz $m \times n$. Então:

a. $\text{posto}(A^TA) = \text{posto}(A)$
b. A matriz $n \times n$ $A^TA$ será invertível se, e somente se, $\text{posto}(A) = n$.

DEMONSTRAÇÃO: (a) Já que $A^TA$ é $n \times n$, ela tem o mesmo número de colunas que $A$. O Teorema do Posto nos diz que

$$\text{posto}(A) + \text{nulidade}(A) = n = \text{posto}(A^TA) + \text{nulidade}(A^TA)$$

Assim, para mostrarmos que $\text{posto}(A) = \text{posto}(A^TA)$, é suficiente mostrar que $\text{nulidade}(A) = \text{nulidade}(A^TA)$. Faremos isso verificando que os espaços anulados por $A$ e por $A^TA$ são os mesmos.

Para isso, considere $\mathbf{x}$ em $\text{anul}(A)$, de modo que com isso, $A\mathbf{x} = \mathbf{0}$. Assim, $A^TA\mathbf{x} = A^T\mathbf{0} = \mathbf{0}$, e, portanto, $\mathbf{x}$ está em $\text{anul}(A^TA)$. Reciprocamente, considere $\mathbf{x}$ em $\text{anul}(A^TA)$. Então, $A^TA\mathbf{x} = \mathbf{0}$, logo, $\mathbf{x}^TA^TA\mathbf{x} = \mathbf{x}^T\mathbf{0} = \mathbf{0}$. Mas então

$$(A\mathbf{x}) \cdot (A\mathbf{x}) = (A\mathbf{x})^T(A\mathbf{x}) = \mathbf{x}^TA^TA\mathbf{x} = \mathbf{0}$$

e concluímos que $A\mathbf{x} = \mathbf{0}$, pelo Teorema 1(d) da Seção 1.3. Portanto, $\mathbf{x}$ está em $\text{anul}(A)$, logo, $\text{anul}(A) = \text{anul}(A^TA)$, como queríamos.
(b) Pelo Teorema Fundamental, a matriz $n \times n$ $A^TA$ será invertível se, e somente se, $\text{posto}(A^TA) = n$. Mas por (a) isso será verdade se, e somente se, $\text{posto}(A) = n$. ◈

## *Coordenadas*

Voltamos agora a uma das questões propostas logo no início desta seção: como descrever vetores de $\mathbb{R}^3$ que pertencem a um plano que passa pela origem? Eles têm duas ou três dimensões? As noções de base e dimensão nos ajudarão a deixar as coisas mais claras.

Um plano que passa pela origem é um subespaço de dimensão dois de $\mathbb{R}^3$, com qualquer subconjunto de dois vetores com direções diferentes servindo como base. Os vetores da base dão a direção dos eixos coordenados no plano/subespaço, e assim nos permitem ver o plano como uma "cópia" de $\mathbb{R}^2$. Antes de ilustrarmos essa abordagem, provaremos um teorema que garante que as "coordenadas" que surgem desse modo são únicas.

### ◈ TEOREMA 11

Considere $S$ um subespaço de $\mathbb{R}^n$ e $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ uma base de $S$. Para todo vetor $\mathbf{v}$ em $S$, existe exatamente um modo de escrever $\mathbf{v}$ como combinação linear dos vetores da base em $\mathcal{B}$:

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k$$

DEMONSTRAÇÃO: Por ser uma base, $\mathcal{B}$ gera $S$, logo, $\mathbf{v}$ pode ser escrito de *pelo menos uma* forma como uma combinação linear de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. Considere uma dessas combinações como

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k$$

Nossa tarefa é mostrar que essa é a única forma de escrever $\mathbf{v}$ como combinação linear de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. Para isso, suponha também que

$$\mathbf{v} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \ldots + d_k\mathbf{v}_k$$

Então:
$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \ldots + d_k\mathbf{v}_k$$

Rearranjando (usando propriedades de álgebra vetorial), obtemos

$$(c_1 - d_1)\mathbf{v}_1 + (c_2 - d_2)\mathbf{v}_2 + \ldots + (c_k - d_k)\mathbf{v}_k = \mathbf{0}$$

Já que $\mathcal{B}$ é uma base, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ são linearmente independentes. Portanto,

$$(c_1 - d_1) = (c_2 - d_2) = \ldots = (c_k - d_k) = 0$$

Em outras palavras, $c_1 = d_1$, $c_2 = d_2$, $\ldots$, $c_k = d_k$, e as duas combinações são exatamente a mesma. Desse modo, existe exatamente um modo de escrever $\mathbf{v}$ como uma combinação linear dos vetores base em $\mathcal{B}$. ◈

**Definição** Sejam $S$ um subespaço de $\mathbb{R}^n$ e $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ uma base de $S$. Considere $\mathbf{v}$ um vetor em $S$ e escreva $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k$. Então, $c_1, c_2, \ldots, c_k$ são chamadas de ***coordenadas de* v *em relação à base*** $\mathcal{B}$, e o vetor coluna

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix}$$

é chamado de ***vetor das coordenadas de* v *em relação à base*** $\mathcal{B}$.

**EXEMPLO 17** Considere $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ a base canônica de $\mathbb{R}^3$. Encontre o vetor das coordenadas de

$$\mathbf{v} = \begin{bmatrix} 2 \\ 7 \\ 4 \end{bmatrix}$$

em relação à base $\mathcal{E}$.

SOLUÇÃO: Já que $\mathbf{v} = 2\mathbf{e}_1 + 7\mathbf{e}_2 + 4\mathbf{e}_3$,

$$[\mathbf{v}]_{\mathcal{E}} = \begin{bmatrix} 2 \\ 7 \\ 4 \end{bmatrix}$$

Deve ficar claro que o vetor das coordenadas de um vetor qualquer (coluna) em $\mathbb{R}^n$, em relação à base canônica, é simplesmente o próprio vetor.

**EXEMPLO 18** No Exemplo 8, vimos que $\mathbf{u} = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$ e $\mathbf{w} = \begin{bmatrix} 0 \\ -5 \\ 1 \end{bmatrix}$ são três vetores pertencentes ao mesmo subespaço (plano que passa pela origem) $S$ de $\mathbb{R}^3$, e que $\mathcal{B} = \{\mathbf{u}, \mathbf{v}\}$ é uma base de $S$. Já que $\mathbf{w} = 2\mathbf{u} - 3\mathbf{v}$, temos

$$[\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$$

Veja a Figura 2.

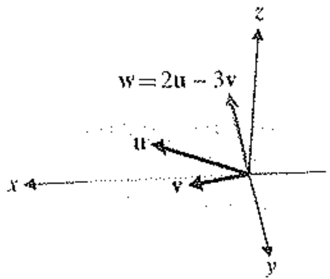



**Figura 2** As coordenadas de um vetor em relação a uma base

## ◈ EXERCÍCIOS 3.5 ◈

*Nos Exercícios de 1 a 4, considere S o conjunto de vetores* $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ *em* $\mathbb{R}^3$ *que satisfazem a propriedade dada. Em cada caso, prove que S forma um subespaço de* $\mathbb{R}^3$*, ou dê um contra-exemplo para mostrar que ele não forma.*

**1.** $x = y = z$

**2.** $z = 2x,\ y = 0$

**3.** $x - y + z = 1$

**4.** $|x - y| = |y - z|$

**5.** Prove que toda reta que passa pela origem em $\mathbb{R}^3$ é um subespaço de $\mathbb{R}^3$.

**6.** Suponha que $S$ consiste em todos os pontos em $\mathbb{R}^2$ que estão no eixo das abscissas ou das coordenadas (ou ambos). ($S$ é chamado de *união* dos dois eixos.) $S$ é um subespaço de $\mathbb{R}^2$? Por quê?

*Nos Exercícios 7 e 8, determine se* **b** *está em col(A) e se* **w** *está em lin(A), como no Exemplo 5.*

**7.** $A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix},\ \mathbf{b} = \begin{bmatrix} 3 \\ 2 \end{bmatrix},\ \mathbf{w} = \begin{bmatrix} -1 & 1 & 1 \end{bmatrix}$

**8.** $A = \begin{bmatrix} 1 & 1 & -3 \\ 0 & 2 & 1 \\ 1 & -1 & -4 \end{bmatrix},\ \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},\ \mathbf{w} = \begin{bmatrix} 2 & 4 & -5 \end{bmatrix}$

**9.** No Exercício 7, determine se **w** está em lin($A$), utilizando o método descrito na Observação após o Exemplo 5.

**10.** No Exercício 8, determine se **w** está em lin($A$), utilizando o método descrito na Observação após o Exemplo 5.

**11.** Se $A$ é a matriz do Exercício 7, então $\mathbf{v} = \begin{bmatrix} -1 \\ 3 \\ -1 \end{bmatrix}$ está em anul($A$)?

**12.** Se $A$ é a matriz do Exercício 8, então $\mathbf{v} = \begin{bmatrix} 7 \\ -1 \\ 2 \end{bmatrix}$ está em anul($A$)?

*Nos Exercícios de 13 a 16, dê bases para lin(A), col(A) e anul(A).*

**13.** $A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}$

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

**15.** $A = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 1 & -1 & -1 \end{bmatrix}$

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

*Nos Exercícios de 17 a 20, ache bases para lin(A) e col(A) dos exercícios citados usando* $A^T$.

**17.** Exercício 13

**18.** Exercício 14

**19.** Exercício 15

**20.** Exercício 16

**21.** Explique cuidadosamente por que suas respostas para os Exercícios 13 e 17 estão ambas corretas, embora pareçam diferentes.

**22.** Explique cuidadosamente por que suas respostas para os Exercícios 14 e 18 estão ambas corretas, embora pareçam diferentes.

*Nos Exercícios de 23 a 26, ache uma base para o conjunto gerado pelos vetores dados.*

**23.** $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$

**24.** $\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$

**25.** $\begin{bmatrix} 2 & -3 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 & 0 \end{bmatrix}, \begin{bmatrix} 4 & -4 & 1 \end{bmatrix}$

**26.** $\begin{bmatrix} 0 & 1 & -2 & 1 \end{bmatrix}, \begin{bmatrix} 3 & 1 & -1 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 1 & 5 & 1 \end{bmatrix}$

*Para os Exercícios 27 e 28, encontre bases, entre os próprios vetores, para os conjuntos gerados pelos vetores nos exercícios citados.*

**27.** Exercício 25

**28.** Exercício 26

**29.** Prove que, se $R$ é uma matriz escalonada, então uma base para lin($A$) consiste nas linhas não nulas de $R$.

**30.** Prove que, se as colunas de $A$ são linearmente independentes, então elas formam necessariamente uma base de col($A$).

*Para os Exercícios de 31 a 34, dê o posto e a nulidade das matrizes nos Exercícios citados.*

**31.** Exercício 13

**32.** Exercício 14

**33.** Exercício 15

**34.** Exercício 16

**35.** Se $A$ é uma matriz 3×5, explique por que as colunas de $A$ devem ser linearmente dependentes.

**36.** Se $A$ é uma matriz 4×2, explique por que as linhas de $A$ devem ser linearmente dependentes.

**37.** Se $A$ é uma matriz 3×5, quais são os possíveis valores de nulidade($A$)?

**38.** Se $A$ é uma matriz 4×2, quais são os possíveis valores de nulidade($A$)?

*Nos Exercícios 39 e 40, encontre todos os valores possíveis de posto(A) em função de a.*

**39.** $A = \begin{bmatrix} 1 & 2 & a \\ -2 & 4a & 2 \\ a & -2 & 1 \end{bmatrix}$

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

*Responda aos Exercícios de 41 a 44 considerando a matriz com os vetores dados como suas colunas.*

**41.** $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ formam uma base de $\mathbb{R}^3$?

**42.** $\begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix}$ formam uma base de $\mathbb{R}^3$?

**43.** $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ formam uma base de $\mathbb{R}^4$?

**44.** $\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ formam uma base de $\mathbb{R}^4$?

*Nos Exercícios 45 e 46, mostre que* $\mathbf{w}$ *está em ger(B) e ache o vetor das coordenadas* $[\mathbf{w}]_B$.

**45.** $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}, \mathbf{w} = \begin{bmatrix} 1 \\ 6 \\ 2 \end{bmatrix}$

**46.** $\mathcal{B} = \left\{ \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \\ 6 \end{bmatrix} \right\}, \mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$

*Nos Exercícios de 47 a 50, calcule o posto e a nulidade das matrizes dadas sobre o* $\mathbb{Z}_p$ *indicado.*

**47.** $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$ sobre $\mathbb{Z}_2$

**48.** $\begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 2 \\ 2 & 0 & 0 \end{bmatrix}$ sobre $\mathbb{Z}_3$

**49.** $\begin{bmatrix} 1 & 3 & 1 & 4 \\ 2 & 3 & 0 & 1 \\ 1 & 0 & 4 & 0 \end{bmatrix}$ sobre $\mathbb{Z}_5$

**50.** $\begin{bmatrix} 2 & 4 & 0 & 0 & 1 \\ 6 & 3 & 5 & 1 & 0 \\ 1 & 0 & 2 & 2 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$ sobre $\mathbb{Z}_7$

**51.** Se $A$ é uma matriz $m \times n$, prove que todo vetor em anul($A$) é ortogonal a todo vetor em lin($A$).

**52.** Se $A$ e $B$ são matrizes $n \times n$ de posto $n$, prove que $AB$ tem posto $n$.

**53.** (a) Prove que posto($AB$) $\leq$ posto($B$). (*Sugestão*: reveja o Exercício 29 da Seção 3.2.)

(b) Dê um exemplo em que posto($AB$) < posto($B$).

**54.** (a) Prove que posto($AB$) $\leq$ posto($A$). [*Sugestão*: reveja o Exercício 30 da Seção 3.2 ou use transpostas e o Exercício 53(a).]

(b) Dê um exemplo em que posto($AB$) < posto($A$).

**55.** (a) Prove que, se $U$ é invertível, então posto($UA$) = posto($A$). [*Sugestão*: $A = U^{-1}(UA)$.]

(b) Prove que, se $V$ é invertível, então posto($AV$) = posto($A$).

**56.** Prove que uma matriz $A$ $m \times n$ tem posto 1 se, e somente se, $A$ pode ser escrita como o produto externo $\mathbf{u}\mathbf{v}^T$ de um vetor $\mathbf{u}$ em $\mathbb{R}^m$ e $\mathbf{v}$ em $\mathbb{R}^n$.

**57.** Se uma matriz $m \times n$ $A$ tem posto $r$, prove que $A$ pode ser escrita como a soma de $r$ matrizes, cada uma tendo posto 1. (*Sugestão*: encontre um meio de utilizar o Exercício 56.)

**58.** Prove que, para matrizes $m \times n$ $A$ e $B$, posto($A + B$) $\leq$ posto($A$) + posto($B$).

## 3.6 Introdução às Transformações Lineares

Nesta seção, começamos a explorar um dos temas da introdução deste capítulo. Lá, vimos que matrizes podem ser usadas para transformar vetores, agindo como um tipo de "função" da forma $\mathbf{w} = T(\mathbf{v})$, em que a variável independente $\mathbf{v}$ e a variável dependente $\mathbf{w}$ são vetores.Tornaremos agora mais precisa essa noção, e veremos diversos exemplos de tais transformações de matrizes, que nos guiarão para o conceito de *transformação linear* — uma idéia poderosa que encontraremos constantemente daqui em diante.

Começamos por relembrar alguns dos principais conceitos associados a funções. Você se sentirá familiarizado com a maior parte das idéias, vindas de outros cursos nos quais você encontrou funções da forma $f:\mathbb{R} \to \mathbb{R}$ [tal como $f(x) = x^2$], que transformam números reais em números reais.

O que é novo aqui é que os vetores estão envolvidos e estamos interessados somente em funções "compatíveis" com operações vetoriais de adição e de multiplicação por escalar.

Considere um exemplo. Sejam

$$A = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix} \quad \text{e} \quad \mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Então

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

Isso mostra que $A$ transforma $\mathbf{v}$ em $\mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$.

Podemos descrever de forma mais geral essa transformação. A equação matricial

$$\begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$$

gera uma fórmula que nos mostra como $A$ transforma um vetor arbitrário $\begin{bmatrix} x \\ y \end{bmatrix}$ de $\mathbb{R}^2$ em um vetor $\begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$ de $\mathbb{R}^3$. Denotamos essa transformação por $T_A$ e escrevemos

$$T_A\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$$

(Embora tecnicamente incorreto, omitir os parênteses em definições como essa é uma convenção comum que economiza escrita. A descrição de $T_A$ passa a ser

$$T_A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$$

com essa convenção.)

Com esse exemplo em mente, consideramos agora alguma terminologia. Uma ***transformação*** (ou ***aplicação*** ou ***função***) $T$ de $\mathbb{R}^n$ em $\mathbb{R}^m$ é uma relação que designa, para cada vetor $\mathbf{v}$ de $\mathbb{R}^n$, um único vetor $T(\mathbf{v})$ em $\mathbb{R}^m$. O ***domínio*** de $T$ é $\mathbb{R}^n$, e o ***contradomínio*** de $T$ é $\mathbb{R}^m$. Indicamos isso escrevendo $T: \mathbb{R}^n \to \mathbb{R}^m$. Para um vetor $\mathbf{v}$ no domínio de $T$, o vetor $T(\mathbf{v})$ no contradomínio é chamado de ***imagem*** de $\mathbf{v}$ sob (a ação de) $T$. O conjunto de todas as imagens possíveis $T(\mathbf{v})$ (com $\mathbf{v}$ variando pelo domínio de $T$) é chamado de ***imagem*** de $T$.

Em nosso exemplo, o domínio de $T_A$ é $\mathbb{R}^2$, e seu contradomínio é $\mathbb{R}^3$. Escrevemos, portanto, $T_A: \mathbb{R}^2 \to \mathbb{R}^3$.

A imagem de $\mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ é $\mathbf{w} = T_A(\mathbf{v}) = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$. Qual é a imagem de $T_A$? Ela consiste em todos os vetores no contradomínio $\mathbb{R}^3$ que são da forma

$$T_A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y\begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$$

que descreve o conjunto de todas as combinações lineares dos vetores coluna $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ e $\begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$ de $A$. Em outras palavras, a imagem de $T$ é o espaço coluna de $A$! (Falaremos mais sobre isso; por enquanto, deixamos como uma simples observação.) Geometricamente, isso mostra que a imagem de $T_A$ é o plano que passa pela origem em $\mathbb{R}^3$ com vetores diretores dados pelos vetores coluna de $A$. Note que a imagem de $T_A$ é estritamente menor que o contradomínio de $T_A$.

## *Transformações Lineares*

O exemplo $T_A$ que acabamos de ver é um caso especial de um tipo mais geral de transformação chamado *transformação linear*. Consideraremos a definição geral no Capítulo 6, mas sua essência é que essas transformações "preservam" as operações vetoriais de adição e de multiplicação por escalar.

**Definição** Uma transformação $T: \mathbb{R}^n \to \mathbb{R}^m$ é chamada de ***transformação linear*** se

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ para todo $\mathbf{u}$ e $\mathbf{v}$ em $\mathbb{R}^n$.
2. $T(c\mathbf{v}) = cT(\mathbf{v})$ para todo $\mathbf{v}$ em $\mathbb{R}^n$ e todos os escalares $c$.

**EXEMPLO 1** Considere novamente a transformação $T: \mathbb{R}^2 \to \mathbb{R}^3$, definida por

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$$

Vamos verificar se $T$ é uma transformação linear. Para verificar (1), considere

$$\mathbf{u} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \quad \text{e} \quad \mathbf{v} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

Então

$$T(\mathbf{u}+\mathbf{v}) = T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}\right)$$

$$= \begin{bmatrix} x_1 + x_2 \\ 2(x_1 + x_2) - (y_1 + y_2) \\ 3(x_1 + x_2) + 4(y_1 + y_2) \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ 2x_1 + 2x_2 - y_1 - y_2 \\ 3x_1 + 3x_2 + 4y_1 + 4y_2 \end{bmatrix}$$

$$= \begin{bmatrix} x_1 + x_2 \\ (2x_1 - y_1) + (2x_2 - y_2) \\ (3x_1 + 4y_1) + (3x_2 + 4y_2) \end{bmatrix} = \begin{bmatrix} x_1 \\ 2x_1 - y_1 \\ 3x_1 + 4y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ 2x_2 - y_2 \\ 3x_2 + 4y_2 \end{bmatrix}$$

$$= T\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + T\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = T(\mathbf{u}) + T(\mathbf{v})$$

Para mostrar (2), considere $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ e $c$ um escalar. Então:

$$T(c\mathbf{v}) = T\left(c\begin{bmatrix} x \\ y \end{bmatrix}\right) = T\left(\begin{bmatrix} cx \\ cy \end{bmatrix}\right)$$

$$= \begin{bmatrix} cx \\ 2(cx) - (cy) \\ 3(cx) + 4(cy) \end{bmatrix} = \begin{bmatrix} cx \\ c(2x - y) \\ c(3x + 4y) \end{bmatrix}$$

$$= c\begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix} = cT\begin{bmatrix} x \\ y \end{bmatrix} = cT(\mathbf{v})$$

Dessa maneira, $T$ é uma transformação linear.

Observação: A definição de uma transformação linear pode ser resumida por combinação de (1) e (2), como a seguir.

$T: \mathbb{R}^n \to \mathbb{R}^m$ é uma transformação linear se

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) \text{ para todo } \mathbf{v}_1, \mathbf{v}_2 \text{ em } \mathbb{R}^n \text{ e escalares } c_1, c_2$$

No Exercício 53, você precisará mostrar que essa afirmação é equivalente à definição original. Na prática, essa formulação equivalente pode economizar alguma escrita — tente!

Embora a transformação linear $T$ do Exemplo 1 tenha originalmente surgido como uma transformação por meio da *matriz* $T_A$, é simples recuperar a matriz $A$ a partir da definição de $T$ dada no exemplo. Observamos que

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y\begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

então, $T = T_A$, em que $A = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix}$. (Perceba que, quando as variáveis $x$ e $y$ estão alinhadas, a matriz $A$ é simplesmente a matriz dos seus coeficientes.)

Reconhecer que uma transformação dada é uma transformação por meio de uma matriz é importante, já que, como o próximo teorema irá mostrar, todas as transformações por meio de uma matriz são transformações lineares.

## ◆ TEOREMA 1

Considere $A$ como uma matriz $m \times n$. Então a transformação por meio da matriz $A$, $T_A$: $\mathbb{R}^n \to \mathbb{R}^m$, definida por

$$T_A(\mathbf{x}) = A\mathbf{x} \quad (\text{para } \mathbf{x} \text{ em } \mathbf{R}^n)$$

é uma transformação linear.

DEMONSTRAÇÃO: Considere $\mathbf{u}$ e $\mathbf{v}$ como vetores em $\mathbb{R}^n$ e $c$ como um escalar. Então:

$$T_A(\mathbf{u} + \mathbf{v}) = A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = T_A(\mathbf{u}) + T_A(\mathbf{v})$$

e

$$T_A(c\mathbf{v}) = A(c\mathbf{v}) = c\,(A\mathbf{v}) = cT_A(\mathbf{v})$$

Por essa razão, $T_A$ é uma transformação linear. ◆

**EXEMPLO 2** $F$: $\mathbb{R}^2 \to \mathbb{R}^2$ é a transformação que leva cada ponto ao seu simétrico em relação ao eixo $x$. Mostre que $F$ é uma transformação linear.

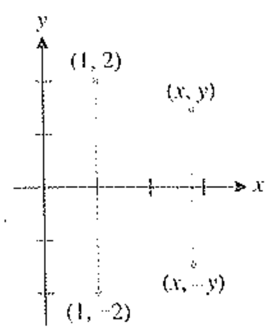



**Figura 1** Simetria em relação ao eixo $x$

SOLUÇÃO: Pela Figura 1, fica claro que $F$ leva o ponto $(x, y)$ ao ponto $(x, -y)$. Dessa forma, podemos escrever

$$F\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix}$$

Poderíamos continuar a verificação de que $F$ é linear, como no Exemplo 1 (este é ainda mais fácil de se verificar!), mas é mais rápido observar que

$$\begin{bmatrix} x \\ -y \end{bmatrix} = x\begin{bmatrix} 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

Portanto, $F\begin{bmatrix} x \\ y \end{bmatrix} = A\begin{bmatrix} x \\ y \end{bmatrix}$, em que $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, logo, $F$ é uma transformação por meio de uma matriz. Segue agora, pelo Teorema 1, que $F$ é uma transformação linear.

**EXEMPLO 3** Seja $R$: $\mathbb{R}^2 \to \mathbb{R}^2$ a transformação que faz a rotação de ângulo 90° no sentido anti-horário em relação à origem. Mostre que $R$ é uma transformação linear.

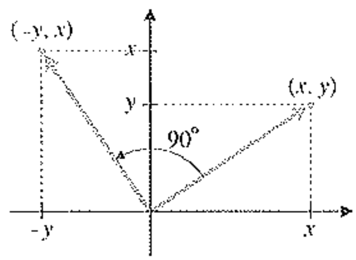


**Figura 2** Uma rotação de 90°

SOLUÇÃO: Como a Figura 2 mostra, $R$ leva o ponto $(x, y)$ ao ponto $(-y, x)$. Dessa forma, temos

$$R\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix} = x\begin{bmatrix} 0 \\ 1 \end{bmatrix} + y\begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

Assim, $R$ é uma transformação por meio de uma matriz e, portanto, é linear.

Observe que, se multiplicarmos uma matriz por vetores da base canônica, obteremos as colunas da matriz. Por exemplo,

$$\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ c \\ e \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b \\ d \\ f \end{bmatrix}$$

Podemos usar essa observação para mostrar que *toda* transformação linear de $\mathbb{R}^n$ em $\mathbb{R}^m$ surge como uma transformação por meio de uma matriz.

### ◆ TEOREMA 2

Considere $T$: $\mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Então, $T$ é uma transformação por meio de uma matriz. Mais especificamente, $T = T_A$, em que $A$ é uma matriz $m \times n$.

$$A = [T(\mathbf{e}_1) \mid T(\mathbf{e}_2) \mid \cdots \mid T(\mathbf{e}_n)]$$

DEMONSTRAÇÃO: Sejam $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ os vetores da base canônica de $\mathbb{R}^n$ e $\mathbf{x}$ um vetor de $\mathbb{R}^n$. Podemos escrever $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n$ (em que os $x_i$'s são as coordenadas de $\mathbf{x}$). Sabemos também que $T(\mathbf{e}_1)$, $T(\mathbf{e}_2), \ldots, T(\mathbf{e}_n)$ são vetores (coluna) em $\mathbb{R}^m$. Considere $A = [T(\mathbf{e}_1) \mid T(\mathbf{e}_2) \mid \cdots \mid T(\mathbf{e}_n)]$ uma matriz $m \times n$ que tem esses vetores como suas colunas. Então:

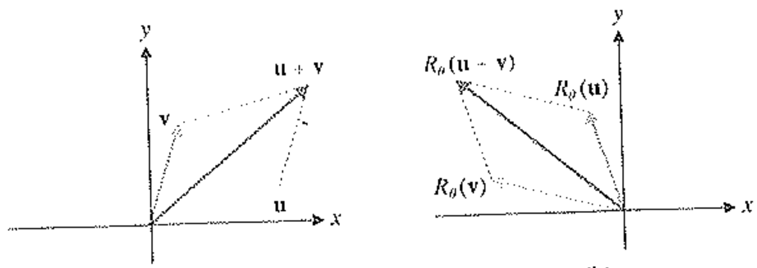
$$
\begin{aligned}
T(\mathbf{x}) &= T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n) \\
&= x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \ldots + x_nT(\mathbf{e}_n) \\
&= [T(\mathbf{e}_1) \mid T(\mathbf{e}_2) \mid \ldots \mid T(\mathbf{e}_n)] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \mathbf{x} = A
\end{aligned}
$$

como queríamos. ◈

A matriz $A$ do Teorema 2 é chamada de ***matriz canônica da transformação linear T***.

**EXEMPLO 4** Mostre que uma rotação de ângulo $\theta$ em torno da origem define uma transformação linear de $\mathbb{R}^2$ em $\mathbb{R}^2$ e ache sua matriz canônica.

SOLUÇÃO: Seja $R_\theta$ uma rotação. Daremos um argumento geométrico para provar que $R_\theta$ é linear. Considere $\mathbf{u}$ e $\mathbf{v}$ como vetores em $\mathbb{R}^2$. Se eles não forem paralelos, a Figura 3(a) mostra a regra do paralelogramo que determina $\mathbf{u} + \mathbf{v}$. Se aplicarmos $R_\theta$, todo o paralelogramo é rodado de ângulo $\theta$, como mostra a Figura 3(b). Mas a diagonal desse paralelogramo precisa ser $R_\theta(\mathbf{u}) + R_\theta(\mathbf{v})$, novamente pela regra do paralelogramo. Por isso, $R_\theta(\mathbf{u} + \mathbf{v}) = R_\theta(\mathbf{u}) + R_\theta(\mathbf{v})$. (O que acontece se $\mathbf{u}$ e $\mathbf{v}$ forem paralelos?)




(a) (b)

**Figura 3**

Analogamente, se aplicarmos $R_\theta$ em $\mathbf{v}$ e $c\mathbf{v}$, obteremos $R_\theta(\mathbf{v})$ e $R_\theta(c\mathbf{v})$, como mostra a Figura 4. Entretanto, como as rotações não afetam o comprimento, precisamos ter $R_\theta(c\mathbf{v}) = cR_\theta(\mathbf{v})$, como exigido. (Desenhe diagramas para os casos $0 < c < 1$, $-1 < c < 0$ e $c < -1$.)

$R_\theta$, portanto, é uma transformação linear. De acordo com o Teorema 2, podemos achar sua matriz pela determinação de seu valor nos vetores da base canônica, $\mathbf{e}_1$ e $\mathbf{e}_2$, de $\mathbb{R}^2$. Agora, como mostra a Figura 5, $R_\theta\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos\theta \\ \operatorname{sen}\theta \end{bmatrix}$.

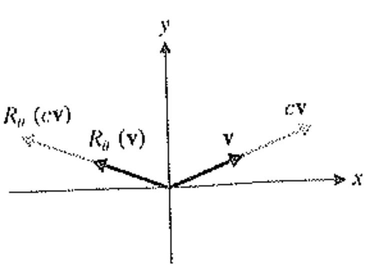



**Figura 4**

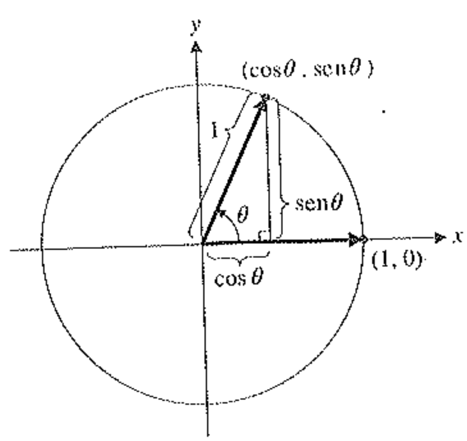



**Figura 5** $R_\theta(\mathbf{e}_1)$

Podemos calcular $R_\theta\begin{bmatrix}0\\1\end{bmatrix}$ similarmente, mas é mais rápido observar que $R_\theta\begin{bmatrix}0\\1\end{bmatrix}$ é necessariamente perpendicular (no sentido anti-horário) a $R_\theta\begin{bmatrix}1\\0\end{bmatrix}$ e, então, pelo Exemplo 3, $R_\theta\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}-\operatorname{sen}\theta\\ \cos\theta\end{bmatrix}$ (Figura 6).

de

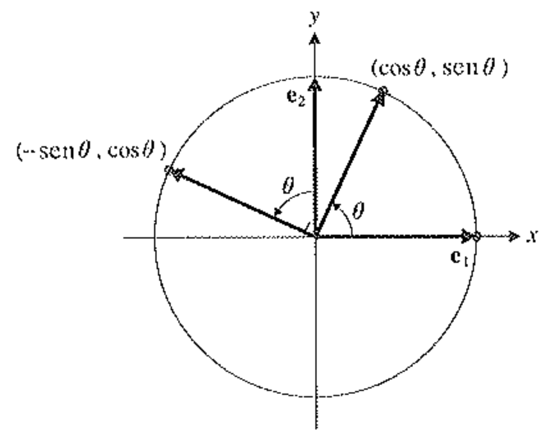



**Figura 6** $R_\theta(\mathbf{e}_2)$

A matriz canônica de $R_\theta$, portanto, é $\begin{bmatrix}\cos\theta & -\operatorname{sen}\theta\\ \operatorname{sen}\theta & \cos\theta\end{bmatrix}$.

O resultado do Exemplo 4 pode agora ser utilizado para o cálculo do efeito de qualquer rotação. Por exemplo, suponha que queiramos fazer a rotação do ponto (2, −1) de ângulo 60° em torno da origem. (A convenção é que um ângulo positivo corresponde a uma rotação no sentido anti-horário, e um ângulo negativo, a uma rotação no sentido horário.) Já que $\cos 60° = \frac{1}{2}$ e $\operatorname{sen} 60° = \sqrt{3}/2$, calculamos

$$R_{60}\begin{bmatrix}2\\-1\end{bmatrix} = \begin{bmatrix}\cos 60° & -\operatorname{sen} 60°\\ \operatorname{sen} 60° & \cos 60°\end{bmatrix}\begin{bmatrix}2\\-1\end{bmatrix} = \begin{bmatrix}1/2 & -\sqrt{3}/2\\ \sqrt{3}/2 & 1/2\end{bmatrix}\begin{bmatrix}2\\-1\end{bmatrix} = \begin{bmatrix}(2+\sqrt{3})/2\\ (2\sqrt{3}-1)/2\end{bmatrix}$$

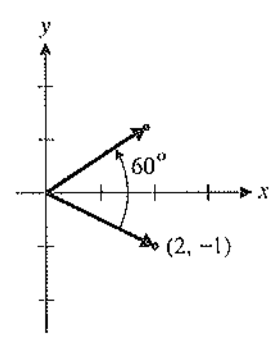



**Figura 7** Uma rotação de ângulo de 60°

Dessa maneira, a imagem do ponto $(2, -1)$ sob essa rotação é o ponto $((2 + \sqrt{3})/2, (2\sqrt{3} - 1)/2) \approx (1{,}87, 1{,}23)$, como mostra a Figura 7.

**EXEMPLO 5** (a) Mostre que a transformação $P: \mathbb{R}^2 \to \mathbb{R}^2$, que projeta um ponto sobre o eixo $x$, é uma trans-

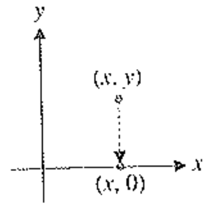



**Figura 8** Uma projeção

formação linear, e ache a matriz canônica da transformação.

(b) Genericamente, se $\ell$ é uma reta que passa pela origem em $\mathbb{R}^2$, mostre que a transformação $P_\ell: \mathbb{R}^2 \to \mathbb{R}^2$, que projeta um ponto sobre $\ell$, é uma transformação linear, e ache sua matriz canônica.

SOLUÇÃO: (a) Como mostra a Figura 8, $P$ leva o ponto $(x, y)$ no ponto $(x, 0)$. Dessa forma,

$$P\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix} = x\begin{bmatrix} 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

Conseqüentemente, $P$ é uma transformação por meio de matriz (e por isso uma transformação linear), com matriz canônica $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

(b) Considere que a reta $\ell$ tem vetor diretor $\mathbf{d}$ e que $\mathbf{v}$ é um vetor arbitrário. Então, $P_\ell$ é dada por $\text{proj}_{\mathbf{d}}(\mathbf{v})$, a projeção de $\mathbf{v}$ em $\mathbf{d}$, da qual você se lembrará da Seção 1.3, e cuja fórmula é

$$\text{proj}_{\mathbf{d}}(\mathbf{v}) = \left(\frac{\mathbf{d}\cdot\mathbf{v}}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d}$$

Assim, para mostrar que $P_\ell$ é linear, procedemos da seguinte forma:

$$\begin{aligned} P_\ell(\mathbf{u} + \mathbf{v}) &= \left(\frac{\mathbf{d}\cdot(\mathbf{u}+\mathbf{v})}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} \\ &= \left(\frac{\mathbf{d}\cdot\mathbf{u} + \mathbf{d}\cdot\mathbf{v}}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} \\ &= \left(\frac{\mathbf{d}\cdot\mathbf{u}}{\mathbf{d}\cdot\mathbf{d}} + \frac{\mathbf{d}\cdot\mathbf{v}}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} \\ &= \left(\frac{\mathbf{d}\cdot\mathbf{u}}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} + \left(\frac{\mathbf{d}\cdot\mathbf{v}}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} = P_\ell(\mathbf{u}) + P_\ell(\mathbf{v}) \end{aligned}$$

Analogamente, $P_\ell(c\mathbf{v}) = cP_\ell(\mathbf{v})$ para todo escalar $c$ (Exercício 52). Por isso, $P_\ell$ é uma transformação linear.

Para encontrar a matriz canônica de $P_\ell$, aplicamos o Teorema 2. Sendo $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$,

$$P_\ell(\mathbf{e}_1) = \left(\frac{\mathbf{d}\cdot\mathbf{e}_1}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} = \frac{d_1}{d_1^2 + d_2^2}\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \frac{1}{d_1^2 + d_2^2}\begin{bmatrix} d_1^2 \\ d_1 d_2 \end{bmatrix}$$

e

$$P_\ell(\mathbf{e}_2) = \left(\frac{\mathbf{d}\cdot\mathbf{e}_2}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} = \frac{d_2}{d_1^2 + d_2^2}\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \frac{1}{d_1^2 + d_2^2}\begin{bmatrix} d_1 d_2 \\ d_2^2 \end{bmatrix}$$

Dessa maneira, a matriz canônica da projeção é

$$A = \frac{1}{d_1^2 + d_2^2}\begin{bmatrix} d_1^2 & d_1d_2 \\ d_1d_2 & d_2^2 \end{bmatrix} = \begin{bmatrix} d_1^2/(d_1^2 + d_2^2) & d_1d_2/(d_1^2 + d_2^2) \\ d_1d_2/(d_1^2 + d_2^2) & d_2^2/(d_1^2 + d_2^2) \end{bmatrix}$$

Como verificação, note que, na parte (a), poderíamos considerar $\mathbf{d} = \mathbf{e}_1$ como um vetor diretor para o eixo $x$. Dessa forma, $d_1 = 1$ e $d_2 = 0$, e obteríamos $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, como antes.

## *Novas Transformações Lineares a partir de Antigas*

Se $T : \mathbb{R}^m \to \mathbb{R}^n$ e $S : \mathbb{R}^n \to \mathbb{R}^p$ são transformações lineares, podemos aplicar $T$ e depois $S$ para formar a ***composta*** das duas transformações, que será denotada por $S \circ T$. Note que, para que $S \circ T$ faça sentido, o contradomínio de $T$ e o domínio de $S$ devem ser o mesmo (neste caso, ambos são $\mathbb{R}^n$), e a transformação composta resultante $S \circ T$ vai do domínio de $T$ ao contradomínio de $S$ (neste caso, $S \circ T : \mathbb{R}^m \to \mathbb{R}^p$). A Figura 9 mostra esquematicamente como essa composta funciona. A definição formal da composta de transformações pode ser tirada diretamente dessa figura, e é a mesma que a definição correspondente da composta de funções ordinárias:

$$(S \circ T)(\mathbf{v}) = S(T(\mathbf{v}))$$

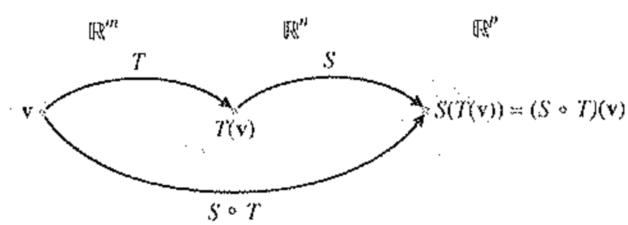



**Figura 9** A composta de transformações

Obviamente, gostaríamos que $S \circ T$ também fosse uma transformação linear, e descobrimos que ela é. Podemos demonstrar isso verificando que $S \circ T$ satisfaz a definição de transformação linear (o que faremos no Capítulo 6), mas, como no momento assumimos que transformações lineares e transformações por meio de matriz são a mesma coisa, é suficiente mostrar que $S \circ T$ é uma transformação por meio de matriz. Usaremos a notação $[T]$ para a matriz canônica de uma transformação linear $T$.

### ◆ TEOREMA 3

Sejam $T : \mathbb{R}^m \to \mathbb{R}^n$ e $S : \mathbb{R}^n \to \mathbb{R}^p$ transformações lineares. Então, $S \circ T : \mathbb{R}^m \to \mathbb{R}^p$ é uma transformação linear. Além disso, suas matrizes canônicas são relacionadas por

$$[S \circ T] = [S][T]$$

DEMONSTRAÇÃO: Sejam $[S] = A$ e $[T] = B$. (Note que $A$ é $p \times n$ e $B$ é $n \times m$.) Se $\mathbf{v}$ é um vetor em $\mathbb{R}^m$, simplesmente calculamos

$$(S \circ T)(\mathbf{v}) = S(T(\mathbf{v})) = S(B\mathbf{v}) = A(B\mathbf{v}) = (AB)\mathbf{v}$$

(Perceba que as dimensões de $A$ e $B$ garantem que o produto $AB$ faz sentido.) Dessa forma, vemos que o efeito de $S \circ T$ é a multiplicação de vetores por $AB$, de onde segue imediatamente que $S \circ T$ é uma transformação por meio de matriz (logo, linear) com $[S \circ T] = [S][T]$. ◈

Esse não é um grande resultado? Diga em palavras: "A matriz da composta é o produto das matrizes". Que fórmula encantadora!

**EXEMPLO 6** Considere a transformação linear $T: \mathbb{R}^2 \to \mathbb{R}^3$ do Exemplo 1, definida por

$$T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 2x_1 - x_2 \\ 3x_1 + 4x_2 \end{bmatrix}$$

e a transformação linear $S : \mathbb{R}^3 \to \mathbb{R}^4$, definida por

$$S\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2y_1 + y_3 \\ 3y_2 - y_3 \\ y_1 - y_2 \\ y_1 + y_2 + y_3 \end{bmatrix}$$

Ache $S \circ T : \mathbb{R}^2 \to \mathbb{R}^4$.

SOLUÇÃO: Podemos ver que as matrizes canônicas são

$$[S] = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & -1 \\ 1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{e} \quad [T] = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix}$$

por isso, pelo Teorema 3,

$$[S \circ T] = [S][T] = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & -1 \\ 1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 3 & -7 \\ -1 & 1 \\ 6 & 3 \end{bmatrix}$$

Conseqüentemente,

$$(S \circ T)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 3 & -7 \\ -1 & 1 \\ 6 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5x_1 + 4x_2 \\ 3x_1 - 7x_2 \\ -x_1 + x_2 \\ 6x_1 + 3x_2 \end{bmatrix}$$

(No Exercício 29, você verificará esse resultado calculando:

e substituindo esses valores na definição de $S$, desse modo calculando $(S \circ T)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ diretamente.)

**EXEMPLO 7** Encontre a matriz canônica da transformação que primeiro faz a rotação de 90° no sentido anti-horário em torno da origem de um ponto e que depois reflete o resultado no eixo $x$.

SOLUÇÃO: A rotação $R$ e a reflexão $F$ foram discutidas nos Exemplos 3 e 2, respectivamente, nos quais achamos suas matrizes canônicas $[R] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ e $[F] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Portanto, a função composta $F \circ R$ tem por matriz

$$[F \circ R] = [F][R] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

(Verifique que esse resultado está correto, considerando o efeito de $F \circ R$ nos vetores $\mathbf{e}_1$ e $\mathbf{e}_2$ da base canônica. Note a importância da *ordem* nas transformações: $R$ é efetuada antes de $F$, mas escrevemos $F \circ R$. Neste caso, $R \circ F$ também faz sentido. Vale a igualdade $R \circ F = F \circ R$?)

## *Inversas de Transformações Lineares*

Considere o efeito de uma rotação de 90° no sentido anti-horário em torno da origem, seguida por uma rotação de 90° no sentido horário em torno da origem. Sem dúvida, isso deixa inalterado qualquer ponto de $\mathbb{R}^2$. Se denotarmos essas transformações por $R_{90}$ e $R_{-90}$ (lembre-se de que um ângulo negativo corresponde a uma rotação no sentido horário), poderemos expressar isso por $(R_{90} \circ R_{-90})(\mathbf{v}) = \mathbf{v}$ para todo $\mathbf{v}$ em $\mathbb{R}^2$. Note que, neste caso, se fizermos a transformação na ordem inversa, obteremos o mesmo resultado final: $(R_{-90} \circ R_{90})(\mathbf{v}) = \mathbf{v}$ para todo $\mathbf{v}$ em $\mathbb{R}^2$.

Dessa maneira, $R_{90} \circ R_{-90}$ (e $R_{-90} \circ R_{90}$ também) é uma transformação linear que deixa todo vetor de $\mathbb{R}^2$ inalterado. Essa transformação é chamada de ***transformação identidade***. Geralmente, temos uma transformação desse tipo para cada $\mathbb{R}^n$ — a saber, $I: \mathbb{R}^n \to \mathbb{R}^n$ tal que $I(\mathbf{v}) = \mathbf{v}$ para todo vetor em $\mathbb{R}^n$. (Se for importante manter a idéia de dimensão, podemos escrever $I_n$ para uma maior clareza.)

Com essa notação, temos $R_{90} \circ R_{-90} = I = R_{-90} \circ R_{90}$. Duas transformações que estão relacionadas desse modo são chamadas de ***transformações inversas***.

**Definição** Considere $S$ e $T$ como transformações lineares de $\mathbb{R}^n$ em $\mathbb{R}^n$. Então, $S$ e $T$ são ***transformações inversas*** se $S \circ T = I_n$ e $T \circ S = I_n$.

Observação:
- Já que nessa definição há uma simetria em relação a $S$ e $T$, dizemos que, quando essa situação ocorre, $S$ é a inversa de $T$ e $T$ é a inversa de $S$. Além disso, dizemos que $S$ e $T$ são ***invertíveis***.

Em termos de matrizes, vemos imediatamente que, se $S$ e $T$ são transformações lineares inversas, $[S][T] = [S \circ T] = [I] = I$, sendo o último $I$ a *matriz* identidade. (Por que a matriz canônica da transformação identidade é a matriz identidade?) Precisamos ter também $[T][S] = [T \circ S] = [I] = I$. Isso mostra que $[S]$ e $[T]$ são matrizes inversas. E mostra algo mais: se uma transformação linear $T$ é invertível, sua matriz canônica $[T]$ tem que ser invertível, e, já que matrizes inversas são únicas, isso significa que a inversa de $T$ também é única. Portanto, podemos usar sem ambigüidade a notação $T^{-1}$ para nos referirmos *à* inversa de $T$. Dessa forma, podemos reescrever as equações anteriores como $[T][T^{-1}] = I = [T^{-1}][T]$, a qual mostra que a matriz de $T^{-1}$ é a inversa de $[T]$. Acabamos de provar o teorema a seguir.

### ◆ TEOREMA 4

Seja $T : \mathbb{R}^n \to \mathbb{R}^n$ uma transformação linear invertível. Então, sua matriz canônica $[T]$ é uma matriz invertível, e

$$[T^{-1}] = [T]^{-1}$$

Observação: Diga também essa expressão em palavras: "A matriz da inversa é a inversa da matriz". Fabuloso!

**EXEMPLO 8** Encontre a matriz canônica de uma rotação de ângulo de 60° no sentido horário, em torno da origem em $\mathbb{R}^2$.

SOLUÇÃO: Vimos que a matriz da rotação de ângulo 60° no sentido anti-horário em torno da origem é

$$[R_{60}] = \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}$$

Considerando que uma rotação de 60° no sentido horário é a inversa de uma rotação de 60° anti-horária, podemos aplicar o Teorema 4 para obter

$$[R_{-60}] = [(R_{60})^{-1}] = \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}^{-1} = \begin{bmatrix} 1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & 1/2 \end{bmatrix}$$

(Verifique o cálculo da matriz inversa. O caminho mais rápido, neste caso de matriz 2×2, é usar o Teorema 3 da Seção 3.4. Além disso, desenhe um diagrama para verificar que a matriz resultante tem o efeito desejado na base canônica de $\mathbb{R}^2$.)

**EXEMPLO 9** Determine se a projeção sobre o eixo $x$ é uma transformação invertível. Caso seja, encontre sua inversa.

SOLUÇÃO: A matriz canônica dessa projeção, $P$, é $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, a qual não é invertível, já que seu determinante é 0. Portanto, $P$ também não é invertível.

Observação:

- A Figura 10 nos dá alguma idéia de por que a transformação $P$ do Exemplo 9 não é invertível. A projeção "esmaga" $\mathbb{R}^2$ no eixo $x$. Para que $P$ fosse invertível, precisaríamos de um modo de "desfazer" isso e recuperar o ponto $(a, b)$ com o qual começamos. Entretanto, existem infinitos candidatos para serem a imagem de $(a, 0)$ sob tal "inversa" hipotética. Qual deles deveríamos usar? Não poderíamos simplesmente dizer que $P^{-1}$ precisa levar $(a, 0)$ até $(a, b)$: isso não seria uma *definição*, pois não teríamos meios de saber o que $b$ deve ser. (Veja o Exercício 42.)

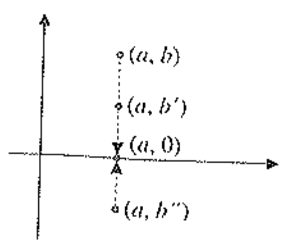



**Figura 10** Projeções não são invertíveis

## Associatividade

O Teorema 2(a) da Seção 3.3 mostrou a propriedade associativa da multiplicação de matrizes: $A(BC) = (AB)C$. (Se você não tentou provar isso, tente agora. Mesmo que seja apenas com matrizes $2\times2$, você terá a noção da complexidade envolvida em uma demonstração "elemento a elemento"; isso o fará apreciar a demonstração que daremos.)

Nossa abordagem à prova será via transformações lineares. Vimos que toda matriz $A$ $m\times n$ produz uma transformação linear $T_A: \mathbb{R}^n \to \mathbb{R}^m$; reciprocamente, toda transformação linear $T: \mathbb{R}^n \to \mathbb{R}^m$ tem uma matriz $[T]$ $m\times n$ correspondente. As duas correspondências são inversamente relacionadas; dado $A$, $[T_A] = A$, e dado $T$, $T_{[T]} = T$.

Considere $R = T_A$, $S = T_B$ e $T = T_C$. Então, pelo Teorema 3,

$$A(BC) = (AB)C \text{ se, e somente se, } R \circ (S \circ T) = (R \circ S) \circ T$$

Provaremos agora a igualdade das transformações. Considere $\mathbf{x}$ no domínio de $T$ (e por isso no domínio de $R \circ (S \circ T)$ e $(R \circ S) \circ T$ — por quê?). Para provar que $R \circ (S \circ T) = (R \circ S) \circ T$, é suficiente mostrar que essas transformações têm o mesmo efeito em $\mathbf{x}$. Aplicando-se repetidas vezes a definição de compostas, temos

$$\begin{aligned}(R \circ (S \circ T))(\mathbf{x}) &= R((S \circ T)(\mathbf{x})) \\ &= R(S(T(\mathbf{x}))) \\ &= (R \circ S)(T(\mathbf{x})) = ((R \circ S) \circ T)(\mathbf{x})\end{aligned}$$

como desejado. (Verifique cuidadosamente como a definição de compostas foi utilizada quatro vezes.)

Esta seção serviu como introdução às transformações lineares. No Capítulo 6, daremos uma olhada mais detalhada e mais geral nessas transformações. Os exercícios a seguir contêm algumas explorações adicionais desse importante conceito.

## ◈ EXERCÍCIOS 3.6 ◈

**1.** Seja $T_A: \mathbb{R}^2 \to \mathbb{R}^2$ a transformação por meio de matriz correspondente a $A = \begin{bmatrix} 2 & -1 \\ 3 & 4 \end{bmatrix}$. Ache $T_A(\mathbf{u})$ e $T_A(\mathbf{v})$, em que $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ e $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$.

**2.** Seja $T_A: \mathbb{R}^3 \to \mathbb{R}^2$ a transformação por meio de matriz correspondente a $A = \begin{bmatrix} 4 & 0 & -1 \\ -2 & 1 & 3 \end{bmatrix}$. Ache $T_A(\mathbf{u})$ e $T_A(\mathbf{v})$, em que $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$ e $\mathbf{v} = \begin{bmatrix} 0 \\ 5 \\ -1 \end{bmatrix}$.

*Nos Exercícios de 3 a 6, prove que a transformação dada é linear, usando a definição (ou a Observação) que aparece depois do Exemplo 1.*

**3.** $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + y \\ x - y \end{bmatrix}$

**4.** $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ x + 2y \\ 3x - 4y \end{bmatrix}$

**5.** $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - y + z \\ 2x + y - 3z \end{bmatrix}$

**6.** $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + z \\ y + z \\ x + y \end{bmatrix}$

*Nos Exercícios de 7 a 10, dê um contra-exemplo para mostrar que a transformação dada não é linear.*

**7.** $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x^2 \end{bmatrix}$

**8.** $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} |x| \\ |y| \end{bmatrix}$

**9.** $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} xy \\ x + y \end{bmatrix}$

**10.** $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 1 \\ y - 1 \end{bmatrix}$

*Nos Exercícios de 11 a 14, encontre a matriz canônica da transformação linear no exercício dado.*

**11.** Exercício 3 **12.** Exercício 4

**13.** Exercício 5 **14.** Exercício 6

*Nos Exercícios de 15 a 18, mostre que a transformação dada de $\mathbb{R}^2$ em $\mathbb{R}^2$ é linear, mostrando que ela é uma transformação por meio de matriz.*

**15.** $F$ reflete um vetor em torno do eixo $y$.

**16.** $R$ faz a rotação de 45° no sentido anti-horário em torno da origem de um vetor.

**17.** $D$ estende um vetor por um fator de 2 na componente $x$ e por um fator de 3 na componente $y$.

**18.** $P$ projeta um vetor na reta $y = x$.

**19.** Os três tipos de matrizes elementares produzem cinco tipos de matrizes 2×2 com uma das seguintes formas:

$$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix} \text{ou} \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \text{ou} \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$$

Descreva geometricamente o efeito da transformação por meio de matriz de $\mathbb{R}^2$ em $\mathbb{R}^2$ que corresponde a cada uma dessas matrizes elementares. Faça desenhos para ilustrar suas respostas.

*Nos Exercícios de 20 a 25, encontre as matrizes canônicas das transformações lineares dadas de $\mathbb{R}^2$ em $\mathbb{R}^2$.*

**20.** Rotação de 120° no sentido anti-horário em torno da origem.

**21.** Rotação de 30° no sentido horário em torno da origem.

**22.** Projeção sobre a reta $y = 2x$.

**23.** Projeção sobre a reta $y = -x$.

**24.** Reflexão em torno da reta $y = x$.

**25.** Reflexão em torno da reta $y = -x$.

**26.** Sejam $\ell$ uma reta que passa pela origem em $\mathbb{R}^2$, $P_\ell$ a transformação linear que projeta um vetor sobre $\ell$ e $F_\ell$ a transformação que reflete um vetor em torno de $\ell$.

(a) Desenhe diagramas para mostrar que $F_\ell$ é linear.

(b) A Figura 11 sugere um modo de achar a matriz de $F_\ell$, usando o fato de que as diagonais de um paralelogramo se interceptam no ponto médio. Prove que $F_\ell(\mathbf{x}) = 2P_\ell(\mathbf{x}) - \mathbf{x}$ e use esse resultado para mostrar que a matriz canônica de $F_\ell$ é

$$\frac{1}{d_1^2 + d_2^2}\begin{bmatrix} d_1^2 - d_2^2 & 2d_1d_2 \\ 2d_1d_2 & -d_1^2 + d_2^2 \end{bmatrix}$$

(em que o vetor diretor de $\ell$ é $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$).

(c) Se o ângulo entre $\ell$ e o semi-eixo positivo $x$ é $\theta$, mostre que a matriz de $F_\ell$ é

$$\begin{bmatrix} \cos 2\theta & \operatorname{sen} 2\theta \\ \operatorname{sen} 2\theta & -\cos 2\theta \end{bmatrix}$$

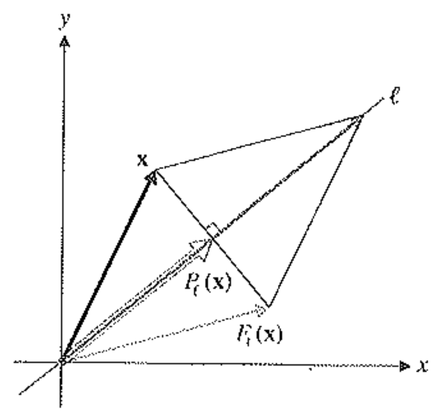



**Figura 11**

*Nos Exercícios 27 e 28, aplique as partes (b) ou (c) do Exercício 26 para encontrar a matriz canônica da transformação.*

**27.** Reflexão em torno da reta $y = 2x$.

**28.** Reflexão em torno da reta $y = \sqrt{3}x$

**29.** Verifique a fórmula para $S \circ T$ no Exemplo 6, fazendo diretamente a substituição sugerida.

*Nos Exercícios de 30 a 35, verifique o Teorema 3 achando a matriz de $S \circ T$ por (a) substituição direta e (b) multiplicação pela matriz $[S][T]$.*

**30.** $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 \\ x_1 + x_2 \end{bmatrix}, S\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2y_1 \\ -y_2 \end{bmatrix}$

**31.** $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ -3x_1 + x_2 \end{bmatrix}, S\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_1 + 3y_2 \\ y_1 - y_2 \end{bmatrix}$

32. $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -x_1 \end{bmatrix}, S\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_1 + 3y_2 \\ 2y_1 + y_2 \\ y_1 - y_2 \end{bmatrix}$

33. $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 - x_3 \\ 2x_1 - x_2 + x_3 \end{bmatrix}, S\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 4y_1 - 2y_2 \\ -y_1 + y_2 \end{bmatrix}$

34. $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ 2x_2 - x_3 \end{bmatrix}, S\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_1 - y_2 \\ y_1 + y_2 \\ -y_1 + y_2 \end{bmatrix}$

35. $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \\ x_1 + x_3 \end{bmatrix}, S\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} y_1 - y_2 \\ y_2 - y_3 \\ -y_1 + y_3 \end{bmatrix}$

*Nos Exercícios de 36 a 39, encontre a matriz canônica da transformação composta de $\mathbb{R}^2$ em $\mathbb{R}^2$.*

**36.** Rotação de 60° no sentido anti-horário em torno da origem, seguida por uma reflexão em torno da reta $y = x$.

**37.** Reflexão em torno do eixo $y$, seguida por uma rotação de 30° no sentido horário em torno da origem.

**38.** Rotação de 45° no sentido horário em torno da origem, seguida por uma projeção sobre o eixo $y$ e depois por uma rotação de 45° no sentido horário em torno da origem.

**39.** Reflexão em torno da reta $y = x$, seguida por uma rotação de 30° no sentido anti-horário em torno da origem e depois por uma reflexão em torno da reta $y = -x$.

*Nos Exercícios de 40 a 43, use matrizes para provar as afirmações sobre transformações de $\mathbb{R}^2$ em $\mathbb{R}^2$.*

**40.** Se $R_\theta$ denota uma rotação (em torno da origem) de um ângulo $\theta$, então $R_\alpha \circ R_\beta = R_{\alpha+\beta}$.

**41.** Se $\theta$ é o ângulo entre as retas $\ell$ e $m$ (que passam pela origem), então $F_m \circ F_\ell = R_{+2\theta}$. (Veja o Exercício 26.)

**42.** (a) Se $P$ é uma projeção, então $P \circ P = P$.

(b) A matriz de uma projeção nunca pode ser invertível.

**43.** Se $\ell$, $m$ e $n$ são três retas que passam pela origem, então $F_n \circ F_m \circ F_\ell$ é também uma reflexão em uma reta que passa pela origem.

**44.** Seja $T$ uma transformação linear de $\mathbb{R}^2$ em $\mathbb{R}^2$ (ou de $\mathbb{R}^3$ em $\mathbb{R}^3$). Prove que $T$ leva uma reta ou em outra reta ou em um ponto. (*Sugestão*: use a equação vetorial de uma reta.)

**45.** Seja $T$ uma transformação linear de $\mathbb{R}^2$ em $\mathbb{R}^2$ (ou de $\mathbb{R}^3$ em $\mathbb{R}^3$). Prove que $T$ leva retas paralelas ou em retas paralelas ou em uma reta única, ou em um par de pontos ou um ponto único.

*Nos Exercícios de 46 a 51, considere ABCD como um quadrado com vértices $(-1, 1)$, $(1, 1)$, $(1, -1)$ e $(-1, -1)$. Use os resultados dos Exercícios 44 e 45 para encontrar e desenhar a imagem de ABCD sob a transformação dada.*

**46.** $T$ do Exercício 3

**47.** $D$ do Exercício 17

**48.** $P$ do Exercício 18

**49.** A projeção do Exercício 22

**50.** $T$ do Exercício 31

**51.** A transformação no Exercício 37

**52.** Prove que $P_\ell(c\mathbf{v}) = cP_\ell(\mathbf{v})$ para todo escalar $c$ [Exemplo 5(b)].

**53.** Prove que $T: \mathbb{R}^n \to \mathbb{R}^m$ é uma transformação linear se, e somente se,

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2)$$

para todos $\mathbf{v}_1$, $\mathbf{v}_2$ em $\mathbb{R}^n$ e escalares $c_1$, $c_2$.

**54.** Prove que (como observado no início desta seção) a imagem de uma transformação linear $T: \mathbb{R}^n \to \mathbb{R}^m$ é o espaço coluna de sua matriz $[T]$.

**55.** Se $A$ é uma matriz invertível 2×2, o que o Teorema Fundamental (Teorema 9 da Seção 3.5) afirma sobre a transformação linear correspondente $T_A$? (Veja o Exercício 19.)

# 3.7 Aplicações

## *Cadeias de Markov*

Uma equipe de pesquisa de mercado está conduzindo um levantamento para determinar a preferência das pessoas em relação a pastas de dentes. A amostra consiste em 200 pessoas; cada uma experimenta duas marcas durante vários meses. Baseada nas respostas do levantamento, a equipe de pesquisa compila a seguinte estatística sobre as preferências quanto a pastas de dentes: