5

Ortogonalidade

"... e aquele escocês dos escoceses,
o esperto Douglas,
que sobe colinas perpendiculares acima
montado a cavalo..."

– William Shakespeare
Henrique IV, Parte I
Ato II, Cena IV
(tradução: Bárbara Heliodora,
Editora Nova Aguilar S.A.)

5.1 Introdução: Sombras em uma Parede

Neste capítulo, estenderemos a noção de projeção ortogonal que encontramos primeiro no Capítulo 1 e depois no Capítulo 3. Até agora, discutimos apenas a projeção ortogonal sobre um único vetor (ou, equivalentemente, o subespaço unidimensional gerado por esse vetor). Nesta seção, veremos a possibilidade de achar fórmulas análogas para a projeção sobre um plano de $\mathbb{R}^3$. A Figura 1 mostra o que acontece, por exemplo, quando raios de luz paralelos criam uma sombra em uma parede. Um processo semelhante ocorre quando um objeto tridimensional é exibido em uma tela bidimensional, tal como o monitor de um computador. Mais tarde, neste capítulo, consideraremos essas idéias em total generalidade.

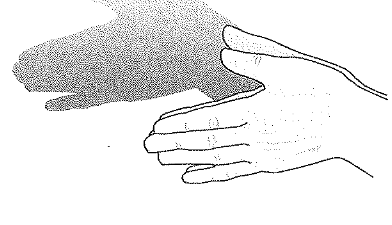

Figura 1 Sombras em uma parede são projeções

Para começar, vamos rever o que já sabemos sobre projeções. Na Seção 3.6, mostramos que, em $\mathbb{R}^2$, a matriz canônica de uma projeção sobre uma reta que passa pela origem e com vetor $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$ é

$$P = \frac{1}{d_1^2 + d_2^2}\begin{bmatrix} d_1^2 & d_1d_2 \\ d_1d_2 & d_2^2 \end{bmatrix} = \begin{bmatrix} d_1^2/(d_1^2 + d_2^2) & d_1d_2/(d_1^2 + d_2^2) \\ d_1d_2/(d_1^2 + d_2^2) & d_2^2/(d_1^2 + d_2^2) \end{bmatrix}$$

Assim, a projeção do vetor $\mathbf{v}$ sobre essa reta é exatamente $P\mathbf{v}$.

Problema 1: Mostre que P pode ser escrita na forma equivalente

$$P = \begin{bmatrix} \cos^2\theta & \cos\theta \operatorname{sen}\theta \\ \cos\theta \operatorname{sen}\theta & \operatorname{sen}^2\theta \end{bmatrix}$$

(O que θ representa aqui?)

Problema 2: Mostre que P também pode ser escrita na forma $P = \mathbf{u}\mathbf{u}^T$, onde $\mathbf{u}$ é um vetor *unitário* na direção de $\mathbf{d}$.

Problema 3: Usando o Problema 2, ache P e a projeção de $\mathbf{v} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$ sobre as retas com os seguintes vetores direção unitários:

$$\text{(a) } \mathbf{u} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \qquad \text{(b) } \mathbf{u} = \begin{bmatrix} \frac{4}{5} \\ \frac{3}{5} \end{bmatrix} \qquad \text{(c) } \mathbf{u} = \begin{bmatrix} -\frac{3}{5} \\ \frac{4}{5} \end{bmatrix}$$

Problema 4: Usando a forma $P = \mathbf{u}\mathbf{u}^T$, mostre que (a) $P^T = P$ (isto é, P é simétrica) e (b) $P^2 = P$ (isto é, P é idempotente).

Problema 5: Quando P é uma matriz de projeção 2×2, a reta sobre a qual ela projeta vetores é o espaço coluna de P. Explique por quê.

Agora vamos mudar para $\mathbb{R}^3$ e considerar projeções sobre planos que passam pela origem. Investigaremos várias abordagens.

A Figura 2 mostra uma maneira de se proceder. Se $\mathscr{P}$ é um plano que passa pela origem em $\mathbb{R}^3$ com vetor normal $\mathbf{n}$ e $\mathbf{v}$ é um vetor em $\mathbb{R}^3$, então $\mathbf{p} = \operatorname{proj}_{\mathscr{P}}(\mathbf{v})$ é um vetor em $\mathscr{P}$ tal que $\mathbf{v} - c\mathbf{n} = \mathbf{p}$ para algum escalar c.

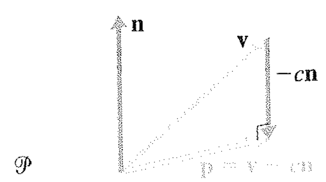

Figura 2 Projeção sobre um plano

Problema 6: Usando o fato de $\mathbf{n}$ ser ortogonal a todo vetor em $\mathscr{P}$, encontre c tal que $\mathbf{v} - c\mathbf{n} = \mathbf{p}$ para obter uma expressão para $\mathbf{p}$ em termos de $\mathbf{v}$ e $\mathbf{n}$.

Problema 7: Use o método do Problema 6 para achar a projeção de

$$\mathbf{v} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$$

sobre os planos com as seguintes equações:

(a) $x + y + z = 0$ (b) $x - 2z = 0$ (c) $2x - 3y + z = 0$

Outra abordagem para o problema de encontrar a projeção de um vetor sobre um plano é a sugerida pela Figura 3: podemos decompor a projeção de **v** sobre $\mathscr{P}$ na *soma* de suas projeções sobre os vetores diretores de $\mathscr{P}$. Isso dá certo apenas quando os vetores diretores são ortogonais e unitários.

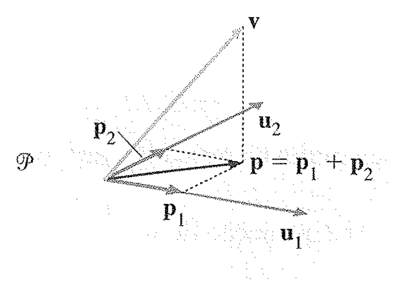

Figura 3

Assim, sejam $\mathbf{u}_1$ e $\mathbf{u}_2$ vetores diretores de $\mathscr{P}$ com a propriedade

$$\|\mathbf{u}_1\| = \|\mathbf{u}_2\| = 1 \quad \text{e} \quad \mathbf{u}_1 \cdot \mathbf{u}_2 = 0$$

Pelo Problema 2, as projeções de **v** sobre $\mathbf{u}_1$ e $\mathbf{u}_2$ são

$$\mathbf{p}_1 = \mathbf{u}_1\mathbf{u}_1^T\mathbf{v} \quad \text{e} \quad \mathbf{p}_2 = \mathbf{u}_2\mathbf{u}_2^T\mathbf{v}$$

respectivamente. Para mostrar que $\mathbf{p}_1 + \mathbf{p}_2$ é a projeção de **v** sobre $\mathscr{P}$, precisamos provar que $\mathbf{v} - (\mathbf{p}_1 + \mathbf{p}_2)$ é ortogonal a $\mathscr{P}$. É suficiente provar que $\mathbf{v} - (\mathbf{p}_1 + \mathbf{p}_2)$ é ortogonal a $\mathbf{u}_1$ e a $\mathbf{u}_2$. (Por quê?)

Problema 8: Mostre que $\mathbf{u}_1 \cdot (\mathbf{v} - (\mathbf{p}_1 + \mathbf{p}_2)) = 0$ e $\mathbf{u}_2 \cdot (\mathbf{v} - (\mathbf{p}_1 + \mathbf{p}_2)) = 0$. (*Sugestão*: use a forma alternativa do produto escalar $\mathbf{x}^T\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ junto com o fato de $\mathbf{u}_1$ e $\mathbf{u}_2$ serem vetores unitários ortogonais.)

De acordo com o Problema 8 e os comentários que o precedem, a matriz da projeção sobre o subespaço $\mathscr{P}$ de $\mathbb{R}^3$ gerado pelos vetores ortogonais $\mathbf{u}_1$ e $\mathbf{u}_2$ é

$$P = \mathbf{u}_1\mathbf{u}_1^T + \mathbf{u}_2\mathbf{u}_2^T \tag{1}$$

Problema 9: Refaça o Problema 7 usando a fórmula dada pela equação (1). Use o mesmo **v** e use $\mathbf{u}_1$ e $\mathbf{u}_2$, como indicado a seguir. (Primeiro verifique que $\mathbf{u}_1$ e $\mathbf{u}_2$ são vetores unitários ortogonais no plano dado.)

(a) $x + y + z = 0$ com $\mathbf{u}_1 = \begin{bmatrix} -2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$ e $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$

$$\text{(b)}\quad x - 2z = 0 \quad \text{com} \quad \mathbf{u}_1 = \begin{bmatrix} 2/\sqrt{5} \\ 0 \\ 1/\sqrt{5} \end{bmatrix} \quad \text{e} \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\text{(c)}\quad 2x - 3y + z = 0 \quad \text{com} \quad \mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \quad \text{e} \quad \mathbf{u}_2 = \begin{bmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix}$$

Problema 10: Mostre que a matriz de projeção dada pela equação (1) satisfaz as propriedades (a) e (b) do Problema 4.

Problema 11: Mostre que a matriz P de uma projeção sobre um plano em $\mathbb{R}^3$ pode ser expressa como

$$P = AA^T$$

onde A é uma matriz 3×2. (*Sugestão*: mostre que a equação (1) é uma expansão em produto externo.)

Problema 12: Mostre que, se P é a matriz de uma projeção sobre um plano em $\mathbb{R}^3$, posto $(P) = 2$.

Neste capítulo, veremos os conceitos de ortogonalidade e projeção ortogonal com grandes detalhes. Veremos que as idéias introduzidas nesta seção podem ser generalizadas e têm muitas aplicações importantes.

5.2 Ortogonalidade em $\mathbb{R}^n$

Nesta seção, generalizaremos a noção de ortogonalidade de dois vetores em $\mathbb{R}^n$ para um conjunto de vetores. Ao fazer isso, veremos que duas propriedades tornam a base canônica $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ de $\mathbb{R}^n$ fácil de ser manipulada: primeiro, dois vetores distintos quaisquer são ortogonais; segundo, cada vetor no conjunto é unitário. Essas duas propriedades nos levam às noções de bases ortogonais e bases ortonormais — conceitos que poderemos usar em uma grande variedade de aplicações.

Conjuntos Ortogonais e Ortonormais de Vetores

Definição Um conjunto de vetores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ em $\mathbb{R}^n$ é chamado de ***conjunto ortogonal*** quando todos os pares de vetores distintos no conjunto são ortogonais — isto é, quando

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0 \text{ sempre que } i \neq j \quad \text{para } i, j = 1, 2, \ldots, k$$

A base canônica $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ de $\mathbb{R}^n$ é um conjunto ortogonal assim como todos os seus subconjuntos. Como o primeiro exemplo ilustra, existem muitas outras possibilidades.

EXEMPLO 1 Mostre que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ é um subconjunto ortogonal de $\mathbb{R}^3$ se

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

SOLUÇÃO: Precisamos mostrar que todo par de vetores desse conjunto é ortogonal. Isso é verdade, pois

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = 2(0) + 1(1) + (-1)(1) = 0$$
$$\mathbf{v}_2 \cdot \mathbf{v}_3 = 0(1) + 1(-1) + (1)(1) = 0$$
$$\mathbf{v}_1 \cdot \mathbf{v}_3 = 2(1) + 1(-1) + (-1)(1) = 0$$

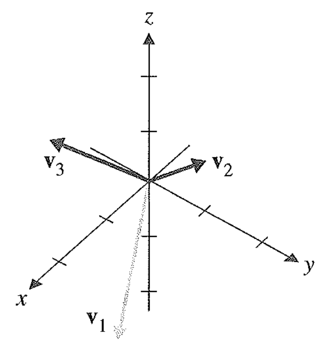

Figura 1 Um conjunto ortogonal de vetores

Geometricamente, os vetores no Exemplo 1 são mutuamente perpendiculares, como mostra a Figura 1. ◆

Uma das vantagens de trabalhar com conjuntos ortogonais de vetores é que eles são linearmente independentes, como mostra o Teorema 1.

◆ TEOREMA 1

Se $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ é um conjunto ortogonal de vetores não nulos em $\mathbb{R}^n$, esses vetores são linearmente independentes.

DEMONSTRAÇÃO: Se $c_1, \ldots, c_k$ são escalares tais que $c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k = \mathbf{0}$, então

$$(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) \cdot \mathbf{v}_i = \mathbf{0} \cdot \mathbf{v}_i = 0$$

ou, equivalentemente,

$$c_1(\mathbf{v}_1 \cdot \mathbf{v}_i) + \cdots + c_i(\mathbf{v}_i \cdot \mathbf{v}_i) + \cdots + c_k(\mathbf{v}_k \cdot \mathbf{v}_i) = 0 \quad (1)$$

Como $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ é um conjunto ortogonal, todos os produtos escalares na equação (1) são iguais a zero, exceto $\mathbf{v}_i \cdot \mathbf{v}_i$. Assim, a equação (1) se reduz a

$$c_i(\mathbf{v}_i \cdot \mathbf{v}_i) = 0$$

Agora, $\mathbf{v}_i \cdot \mathbf{v}_i \neq 0$ pois, por hipótese, $\mathbf{v}_i \neq 0$. Logo, precisamos ter $c_i = 0$. O fato de isso ser verdadeiro para todo $i = 1, \ldots, k$ implica que $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ é um conjunto linearmente independente. ◆

Observação: ◆ Graças ao Teorema 1, sabemos que, se um conjunto de vetores não nulos é ortogonal, ele é, automaticamente, linearmente independente. Por exemplo, podemos deduzir imediatamente que os três vetores no Exemplo 1 são linearmente independentes. Compare essa abordagem com o trabalho que precisávamos ter para provar a independência linear deles diretamente!

Definição Uma ***base ortogonal*** para um subespaço W de $\mathbb{R}^n$ é uma base de W que é um conjunto ortogonal.

EXEMPLO 2 Os vetores

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

do Exemplo 1 são ortogonais, e, portanto, linearmente independentes. Como três vetores quaisquer linearmente independentes em $\mathbb{R}^3$ formam uma base de $\mathbb{R}^3$, pelo Teorema Fundamental das Matrizes Invertíveis, segue que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ é uma base ortogonal de $\mathbb{R}^3$.

EXEMPLO 3 Encontre uma base ortogonal para o subespaço W de $\mathbb{R}^3$ dado por

$$W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : x - y + 2z = 0 \right\}$$

SOLUÇÃO: A Seção 5.4 dá um procedimento geral para resolver esse tipo de problema. Por ora, encontraremos uma base ortogonal na força bruta. O subespaço W é um plano que passa pela origem em $\mathbb{R}^3$. Da equação do plano, temos que $x = y - 2z$, e, assim, W consiste nos vetores da forma

$$\begin{bmatrix} y - 2z \\ y \\ z \end{bmatrix} = y\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + z\begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

Segue que $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ e $\mathbf{v} = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$ formam uma base para W, mas eles *não* são ortogonais. É suficiente encontrar outro vetor em W que seja ortogonal a qualquer um desses.

Suponha que $\mathbf{w} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ seja um vetor em W ortogonal a $\mathbf{u}$. Então, $x - y + 2z = 0$, já que $\mathbf{w}$ está no plano W. Como $\mathbf{u} \cdot \mathbf{w} = 0$, também temos que $x + y = 0$. Resolvendo o sistema linear

$$\begin{aligned} x - y + 2z &= 0 \\ x + y &= 0 \end{aligned}$$

encontramos $x = -z$ e $y = z$. (Verifique isso.) Então, qualquer vetor não nulo $\mathbf{w}$ da forma

$$\mathbf{w} = \begin{bmatrix} -z \\ z \\ z \end{bmatrix}$$

servirá. Podemos tomar, por exemplo, $\mathbf{w} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$. É fácil verificar que $\{\mathbf{u}, \mathbf{w}\}$ é um conjunto ortogonal em W, e, conseqüentemente, uma base ortogonal para W, já que $\dim W = 2$.

Outra vantagem de trabalhar com uma base ortogonal é que as coordenadas de um vetor em relação a tal base são fáceis de calcular. De fato, existe uma fórmula para essas coordenadas, como o próximo teorema estabelece.

◆ TEOREMA 2

Seja $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ uma base ortogonal de um subespaço W de $\mathbb{R}^n$ e $\mathbf{w}$, um vetor em W. Então, os únicos escalares $c_1, \ldots, c_k$ tais que

$$\mathbf{w} = c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k$$

são dados por

$$c_i = \frac{\mathbf{w}\cdot\mathbf{v}_i}{\mathbf{v}_i\cdot\mathbf{v}_i} \quad \text{para } i = 1, \ldots, k$$

DEMONSTRAÇÃO: Como $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ é uma base de W, existem escalares $c_1, \ldots, c_k$, unicamente determinados, tais que $\mathbf{w} = c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k$ (do Teorema 11 da Seção 3.5). Para estabelecer a fórmula para c_i, efetuamos o produto escalar da combinação linear anterior por $\mathbf{v}_i$ para obter

$$\begin{aligned}\mathbf{w}\cdot\mathbf{v}_i &= (c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k)\cdot\mathbf{v}_i \\ &= c_1(\mathbf{v}_1\cdot\mathbf{v}_i) + \cdots + c_i(\mathbf{v}_i\cdot\mathbf{v}_i) + \cdots + c_k(\mathbf{v}_k\cdot\mathbf{v}_i) \\ &= c_i(\mathbf{v}_i\cdot\mathbf{v}_i)\end{aligned}$$

já que $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ se $i \neq j$. Como $\mathbf{v}_i \neq 0$, $\mathbf{v}_i \cdot \mathbf{v}_i \neq 0$. Dividindo por $\mathbf{v}_i \cdot \mathbf{v}_i$, obtemos o resultado desejado. ◆

EXEMPLO 4 Encontre as coordenadas do vetor $\mathbf{w} = \begin{bmatrix}1\\2\\3\end{bmatrix}$ com respeito à base ortogonal $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ dos Exemplos 1 e 2.

SOLUÇÃO: Usando o Teorema 2, calculamos

$$c_1 = \frac{\mathbf{w}\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1} = \frac{2+2-3}{4+1+1} = \frac{1}{6}$$

$$c_2 = \frac{\mathbf{w}\cdot\mathbf{v}_2}{\mathbf{v}_2\cdot\mathbf{v}_2} = \frac{0+2+3}{0+1+1} = \frac{5}{2}$$

$$c_3 = \frac{\mathbf{w}\cdot\mathbf{v}_3}{\mathbf{v}_3\cdot\mathbf{v}_3} = \frac{1-2+3}{1+1+1} = \frac{2}{3}$$

Então,

$$\mathbf{w} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \tfrac{1}{6}\mathbf{v}_1 + \tfrac{5}{2}\mathbf{v}_2 + \tfrac{2}{3}\mathbf{v}_3$$

(Verifique isso.) Com as notações da Seção 3.5, podemos também escrever a equação anterior como

$$[\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix}\frac{1}{6}\\ \frac{5}{2}\\ \frac{2}{3}\end{bmatrix}$$

Compare o procedimento no Exemplo 1 com o trabalho necessário para achar essas coordenadas diretamente. Você começará a apreciar a utilidade das bases ortogonais.

Como observado no início desta seção, a outra propriedade da base canônica de $\mathbb{R}^n$ é que cada vetor pertencente a ela é unitário. Combinando essa propriedade com a ortogonalidade, temos a seguinte definição:

Definição Um conjunto de vetores em $\mathbb{R}^n$ será um ***conjunto ortonormal*** se for um conjunto ortogonal de vetores unitários. Uma ***base ortonormal*** para um subespaço W de $\mathbb{R}^n$ é uma base de W que é um conjunto ortonormal.

Observação: ◆ Se $S = \{\mathbf{q}_1, \ldots, \mathbf{q}_k\}$ é um conjunto ortonormal de vetores, então $\mathbf{q}_i \cdot \mathbf{q}_j = 0$ para $i \neq j$ e $\|\mathbf{q}_i\| = 1$. O fato de cada $\mathbf{q}_i$ ser um vetor unitário é equivalente a $\mathbf{q}_i \cdot \mathbf{q}_i = 1$. Conseqüentemente, podemos sumarizar a afirmação de que S é ortonormal como

$$\mathbf{q}_i \cdot \mathbf{q}_j = \begin{cases} 0 & \text{se} \quad i \neq j \\ 1 & \text{se} \quad i = j \end{cases}$$

EXEMPLO 5 Mostre que $S = \{\mathbf{q}_1, \mathbf{q}_2\}$ é um subconjunto ortonormal de $\mathbb{R}^3$ se

$$\mathbf{q}_1 = \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \quad \text{e} \quad \mathbf{q}_2 = \begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$$

SOLUÇÃO: Verificamos que

$$\begin{aligned} \mathbf{q}_1 \cdot \mathbf{q}_2 &= 1/\sqrt{18} - 2/\sqrt{18} + 1/\sqrt{18} = 0 \\ \mathbf{q}_1 \cdot \mathbf{q}_1 &= 1/3 + 1/3 + 1/3 = 1 \\ \mathbf{q}_2 \cdot \mathbf{q}_2 &= 1/6 + 4/6 + 1/6 = 1 \end{aligned}$$

Quando temos um conjunto ortogonal, podemos obter facilmente dele um conjunto ortonormal: simplesmente normalizamos cada vetor.

EXEMPLO 6 Construa uma base ortonormal para $\mathbb{R}^3$ usando os vetores do Exemplo 1.

SOLUÇÃO: Como já sabemos que $\mathbf{v}_1$, $\mathbf{v}_2$ e $\mathbf{v}_3$ formam uma base ortogonal, nós os normalizamos para obter

$$\mathbf{q}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1 = \frac{1}{\sqrt{6}}\begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix}$$

$$\mathbf{q}_2 = \frac{1}{\|\mathbf{v}_2\|}\mathbf{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$\mathbf{q}_3 = \frac{1}{\|\mathbf{v}_3\|}\mathbf{v}_3 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

Então, $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$ é uma base ortonormal para $\mathbb{R}^3$.

Como todo conjunto ortonormal de vetores é, em particular, ortogonal, ele é linearmente independente pelo Teorema 1. Tendo uma base ortonormal, o Teorema 2 fica mais simples.

TEOREMA 3

Seja $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_k\}$ uma base ortonormal para um subespaço W de $\mathbb{R}^n$ e $\mathbf{w}$, um vetor em W. Então

$$\mathbf{w} = (\mathbf{w} \cdot \mathbf{q}_1)\mathbf{q}_1 + (\mathbf{w} \cdot \mathbf{q}_2)\mathbf{q}_2 + \ldots + (\mathbf{w} \cdot \mathbf{q}_k)\mathbf{q}_k$$

e essa representação é única.

DEMONSTRAÇÃO: Aplique o Teorema 2 e use o fato de que $\mathbf{q}_i \cdot \mathbf{q}_i = 1$ para $i = 1, \ldots, k$.

Matrizes Ortogonais

Matrizes cujas colunas formam um conjunto ortonormal aparecem freqüentemente em aplicações, como você verá na Seção 5.6. Essas matrizes têm várias propriedades atraentes, que vamos agora examinar.

TEOREMA 4

Seja Q uma matriz $m \times n$. Suas colunas formam um conjunto ortonormal se, e somente se, $Q^TQ = I_n$.

DEMONSTRAÇÃO: Precisamos provar que

$$(Q^TQ)_{ij} = \begin{cases} 0 & \text{se} \quad i \neq j \\ 1 & \text{se} \quad i = j \end{cases}$$

Denotemos por $\mathbf{q}_i$ a i-ésima coluna de Q (e então a i-ésima linha de Q^T). Como o elemento (i, j) de Q^TQ é o produto escalar da i-ésima linha de Q^T pela j-ésima coluna de Q, segue que

$$(Q^TQ)_{ij} = \mathbf{q}_i \cdot \mathbf{q}_j \tag{2}$$

pela definição de multiplicação de matrizes. Agora, as colunas de Q formam um conjunto ortonormal se e somente se

$$\mathbf{q}_i \cdot \mathbf{q}_j = \begin{cases} 0 & \text{se} \quad i \neq j \\ 1 & \text{se} \quad i = j \end{cases}$$

o que, pela equação (2), vale se e somente se

$$(Q^TQ)_{ij} = \begin{cases} 0 & \text{se} \quad i \neq j \\ 1 & \text{se} \quad i = j \end{cases}$$

Isso completa a demonstração.

Se a matriz Q no Teorema 4 é uma matriz *quadrada*, ela recebe um nome especial.

Definição Se Q é uma matriz $n \times n$ cujas colunas formam um conjunto ortonormal, então ela é chamada ***matriz ortogonal***.[1]

O fato mais importante a respeito de matrizes ortogonais é dado pelo teorema a seguir.

[1] Essa não é uma terminologia muito adequada. "Matriz ortonormal" seria certamente um termo melhor, mas não é usado. Além disso, não há nenhum termo para designar uma matriz não quadrada com colunas ortonormais.

◆ TEOREMA 5

Uma matriz quadrada Q é ortogonal se e somente se $Q^{-1} = Q^T$.

DEMONSTRAÇÃO: Pelo Teorema 4, Q é ortogonal se, e somente se, $Q^TQ = I$. Isso é verdade se, e somente se, Q é invertível e $Q^{-1} = Q^T$, pelo Teorema 8 da Seção 3.4 ◆

EXEMPLO 7 Mostre que as matrizes seguintes são ortogonais e encontre suas inversas.

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

SOLUÇÃO: As colunas de A são exatamente os vetores da base canônica de $\mathbb{R}^3$, que é claramente ortonormal. Logo, A é ortogonal e

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

Para B, verificamos diretamente que

$$\begin{aligned} B^TB &= \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos^2\theta + \operatorname{sen}^2\theta & -\cos\theta\ \operatorname{sen}\theta + \operatorname{sen}\theta\ \cos\theta \\ -\operatorname{sen}\theta\ \cos\theta + \cos\theta\operatorname{sen}\theta & \operatorname{sen}^2\theta + \cos^2\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \end{aligned}$$

Assim, B é ortogonal, e, pelo Teorema 5, temos que

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

Observação: ◆ A matriz A do Exemplo 7 é um exemplo de matriz de permutação — uma matriz obtida pela permutação de colunas de uma matriz-identidade. Em geral, qualquer matriz de permutação $n \times n$ é ortogonal (veja o Exercício 25). A matriz B é a matriz de uma rotação de ângulo θ em $\mathbb{R}^2$. Qualquer rotação tem a propriedade de ser uma transformação que *preserva comprimentos* (em geometria, conhecida como ***isometria***). O próximo teorema mostra que toda matriz de transformação ortogonal é uma isometria. Matrizes ortogonais também preservam produtos escalares. Na verdade, elas são caracterizadas por qualquer uma dessas duas propriedades.

A palavra *isometria* literalmente significa "medidas iguais", já que é derivada das raízes gregas *isos* ("igual") e *metron* ("medida").

◆ TEOREMA 6

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

a. Q é ortogonal.

b. $\|Q\mathbf{x}\| = \|\mathbf{x}\|$ para todo $\mathbf{x}$ em $\mathbb{R}^n$.

c. $Q\mathbf{x} \cdot Q\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ para todo $\mathbf{x}$ e $\mathbf{y}$ em $\mathbb{R}^n$.

DEMONSTRAÇÃO: Provaremos que (a) $\Rightarrow$ (c) $\Rightarrow$ (b) $\Rightarrow$ (a). Para isso, precisaremos usar o fato de que, se $\mathbf{x}$ e $\mathbf{y}$ são vetores (coluna) em $\mathbb{R}^n$, então $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T\mathbf{y}$.
(a) $\Rightarrow$ (c) Assuma que Q é ortogonal. Então, $Q^TQ = I$, e temos

$$Q\mathbf{x} \cdot Q\mathbf{y} = (Q\mathbf{x})^TQ\mathbf{y} = \mathbf{x}^TQ^TQ\mathbf{y} = \mathbf{x}^TI\,\mathbf{y} = \mathbf{x}^T\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$$

(c) $\Rightarrow$ (b) Assuma que $Q\mathbf{x} \cdot Q\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ para todo $\mathbf{x}$ e $\mathbf{y}$ em $\mathbb{R}^n$. Então, tomando $\mathbf{y} = \mathbf{x}$, temos $Q\mathbf{x} \cdot Q\mathbf{x} = \mathbf{x} \cdot \mathbf{x}$, por isso $\|Q\mathbf{x}\| = \sqrt{Q\mathbf{x} \cdot Q\mathbf{x}} = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \|\mathbf{x}\|$.
(b) $\Rightarrow$ (a) Considere válida a propriedade (b) e que $\mathbf{q}_i$ denota a i-ésima coluna de Q. Usando o Exercício 49 da Seção 1.3 e a propriedade (b), temos

$$\begin{aligned}
\mathbf{x} \cdot \mathbf{y} &= \tfrac{1}{4}(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2) \\
&= \tfrac{1}{4}(\|Q(\mathbf{x} + \mathbf{y})\|^2 - \|Q(\mathbf{x} - \mathbf{y})\|^2) \\
&= \tfrac{1}{4}(\|Q\mathbf{x} + Q\mathbf{y}\|^2 - \|Q\mathbf{x} - Q\mathbf{y}\|^2) \\
&= Q\mathbf{x} \cdot Q\mathbf{y}
\end{aligned}$$

para todo $\mathbf{x}$ e $\mathbf{y}$ em $\mathbb{R}^n$. [Isso mostra que (b) $\Rightarrow$ (c).]

Agora, se $\mathbf{e}_i$ é o i-ésimo vetor da base canônica, $\mathbf{q}_i = Q\mathbf{e}_i$. Conseqüentemente,

$$\mathbf{q}_i \cdot \mathbf{q}_j = Q\mathbf{e}_i \cdot Q\mathbf{e}_j = \mathbf{e}_i \cdot \mathbf{e}_j = \begin{cases} 0 & \text{se} \quad i \neq j \\ 1 & \text{se} \quad i = j \end{cases}$$

As colunas de Q formam, portanto, um conjunto ortonormal, e Q é uma matriz ortogonal. ◈

Olhando as matrizes A e B do Exemplo 7, você pode notar que não apenas suas colunas formam conjuntos ortonormais, mas também suas *linhas*. Na verdade, toda matriz ortogonal tem essa propriedade, como mostra o próximo teorema.

◆ TEOREMA 7

Se Q é uma matriz ortogonal, então suas linhas formam um conjunto ortonormal.

DEMONSTRAÇÃO: Pelo Teorema 5, sabemos que $Q^{-1} = Q^T$. Logo,

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

e assim Q^T é uma matriz ortogonal. Então, as colunas de Q^T — que são exatamente as linhas de Q — formam um conjunto ortonormal. ◈

O teorema final desta seção lista algumas outras propriedades de matrizes ortogonais.

◆ TEOREMA 8

Seja Q uma matriz ortogonal.
a. Q^{-1} é ortogonal.
b. $\det Q = \pm 1$
c. Se λ é um autovalor de Q, $|\lambda| = 1$.
d. Se Q_1 e Q_2 são matrizes ortogonais $n \times n$, Q_1Q_2 também é.

DEMONSTRAÇÃO: Provaremos a propriedade (c) e deixaremos as demonstrações das outras como exercícios. (c) Seja λ um autovalor de Q com um autovetor correspondente $\mathbf{v}$. Então, $Q\mathbf{v} = \lambda\mathbf{v}$, e, usando o Teorema 6 (b), temos

$$\|\mathbf{v}\| = \|Q\mathbf{v}\| = \|\lambda\mathbf{v}\| = |\lambda|\,\|\mathbf{v}\|$$

Como $\|\mathbf{v}\| \neq 0$, $\|\lambda\mathbf{v}\| = 1$. ◆

a + bi Observação: ◆ A propriedade (c) continua verdadeira para autovalores complexos. A matriz $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ é ortogonal com autovalores i e $-i$, ambos com módulo igual a 1.

◆ EXERCÍCIOS 5.2 ◆

Nos Exercícios de 1 a 6, determine quais conjuntos de vetores são ortogonais.

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

3. $\begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 4 \end{bmatrix}$ **4.** $\begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}$

5. $\begin{bmatrix} 2 \\ 3 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ -6 \\ 2 \\ 7 \end{bmatrix}$

6. $\begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$

Nos Exercícios de 7 a 10, mostre que os vetores dados formam uma base ortogonal para $\mathbb{R}^2$ ou $\mathbb{R}^3$. Então, use o Teorema 2 para expressar $\mathbf{w}$ como combinação linear dos vetores dessa base. Dê o vetor de coordenadas $[\mathbf{w}]_{\mathcal{B}}$ de $\mathbf{w}$ em relação à base $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ de $\mathbb{R}^2$ ou $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ de $\mathbb{R}^3$.

7. $\mathbf{v}_1 = \begin{bmatrix} 4 \\ -2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}; \mathbf{w} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$

8. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -6 \\ 2 \end{bmatrix}; \mathbf{w} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

9. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}; \mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

10. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}; \mathbf{w} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

Nos Exercícios de 11 a 15, determine se o conjunto ortogonal de vetores dado é ortonormal. Caso não seja, normalize os vetores para formar um conjunto ortonormal.

11. $\begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix}, \begin{bmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{bmatrix}$ **12.** $\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$

13. $\begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}, \begin{bmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -\frac{5}{2} \end{bmatrix}$ **14.** $\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{1}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{6} \\ \frac{1}{6} \\ -\frac{1}{6} \end{bmatrix}$

15. $\begin{bmatrix} 1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{bmatrix}, \begin{bmatrix} 0 \\ \sqrt{6}/3 \\ 1/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix}, \begin{bmatrix} \sqrt{3}/2 \\ -\sqrt{3}/6 \\ \sqrt{3}/6 \\ -\sqrt{3}/6 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

Nos Exercícios de 16 a 21, determine se a matriz dada é ortogonal. Caso seja, ache sua inversa.

16. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ **17.** $\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

18. $\begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{2}{5} \\ \frac{1}{2} & -\frac{1}{3} & \frac{2}{5} \\ -\frac{1}{2} & 0 & \frac{4}{5} \end{bmatrix}$

19. $\begin{bmatrix} \cos\theta\,\mathrm{sen}\,\theta & -\cos\theta & -\mathrm{sen}^2\,\theta \\ \cos^2\theta & \mathrm{sen}\,\theta & -\cos\theta\,\mathrm{sen}\,\theta \\ \mathrm{sen}\,\theta & 0 & \cos\theta \end{bmatrix}$

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

21. $\begin{bmatrix} 1 & 0 & 0 & 1/\sqrt{6} \\ 0 & 2/3 & 1/\sqrt{2} & 1/\sqrt{6} \\ 0 & -2/3 & 1/\sqrt{2} & -1/\sqrt{6} \\ 0 & 1/3 & 0 & 1/\sqrt{2} \end{bmatrix}$

22. Prove o Teorema 8(a).

23. Prove o Teorema 8(b).

24. Prove o Teorema 8(d).

25. Prove que toda matriz de permutação é ortogonal.

26. Se Q é uma matriz ortogonal, mostre que toda matriz obtida de Q por rearranjo de suas linhas também é ortogonal.

27. Seja Q uma matriz 2×2 ortogonal e sejam $\mathbf{x}$ e $\mathbf{y}$ vetores em $\mathbb{R}^2$. Se θ é o ângulo entre $\mathbf{x}$ e $\mathbf{y}$, mostre que o ângulo entre $Q\mathbf{x}$ e $Q\mathbf{y}$ também é θ. (Isso prova que transformações lineares definidas por matrizes ortogonais *preservam ângulos* em $\mathbb{R}^2$, um fato que é verdadeiro em geral.)

28. (a) Mostre que uma matriz ortogonal 2×2 tem necessariamente a forma

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \quad \text{ou} \quad \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$$

onde $\begin{bmatrix} a \\ b \end{bmatrix}$ é um vetor unitário.

(b) Usando a parte (a), mostre que toda matriz ortogonal 2×2 tem a forma

$$\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \quad \text{ou} \quad \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{bmatrix}$$

onde $0 < \theta \leq 2\pi$.

(c) Mostre que toda matriz 2×2 ortogonal corresponde a uma rotação ou a uma reflexão em $\mathbb{R}^2$.

(d) Mostre que uma matriz 2×2 ortogonal Q corresponde a uma rotação em $\mathbb{R}^2$ se det $Q = 1$, e a uma reflexão em $\mathbb{R}^2$ se det $Q = -1$.

Nos Exercícios de 29 a 32, use o Exercício 28 para determinar se a matriz dada representa uma rotação ou uma reflexão. Caso seja uma rotação, determine o ângulo de rotação; caso seja uma reflexão, dê a reta de reflexão.

29. $\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

30. $\begin{bmatrix} -1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & -1/2 \end{bmatrix}$

31. $\begin{bmatrix} -1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}$

32. $\begin{bmatrix} -\frac{3}{5} & -\frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{bmatrix}$

33. Sejam A e B matrizes ortogonais $n \times n$.

(a) Prove que $A(A^T + B^T)B = A + B$.

(b) Use a parte (a) para provar que, se det A + det $B = 0$, então $A + B$ não é invertível.

34. Seja $\mathbf{x}$ um vetor unitário em $\mathbb{R}^n$. Decomponha $\mathbf{x}$ como

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \hdashline x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ \mathbf{y} \end{bmatrix}$$

Seja

$$Q = \left[\begin{array}{c:c} x_1 & \mathbf{y}^T \\ \hdashline \mathbf{y} & I - \left(\dfrac{1}{1 - x_1}\right)\mathbf{y}\mathbf{y}^T \end{array}\right]$$

Prove que Q é ortogonal. (Esse procedimento fornece um método rápido para achar uma base ortonormal de $\mathbb{R}^n$ com um primeiro vetor prescrito $\mathbf{x}$, uma construção que é freqüentemente útil em aplicações.)

5.3 Complementos e Projeções Ortogonais

Nesta seção, vamos generalizar dois conceitos que encontramos no Capítulo 1. A noção de um vetor ser normal a um plano será estendida para complementos ortogonais, e a projeção de um vetor sobre outro dará origem ao conceito de projeção sobre um subespaço.

Complementos Ortogonais

Um vetor normal $\mathbf{n}$ a um plano é ortogonal a todo vetor nesse plano. Se o plano passa pela origem, ele é um subespaço W de $\mathbb{R}^3$, assim como ger($\mathbf{n}$). Temos então dois subespaços de $\mathbb{R}^3$, com a propriedade de que todo vetor em um deles é ortogonal a todo vetor no outro. Essa é a idéia por trás da seguinte definição.

Definição Seja W um subespaço de $\mathbb{R}^n$. Dizemos que um vetor $\mathbf{v}$ em $\mathbb{R}^n$ é ***ortogonal a*** W se $\mathbf{v}$ for ortogonal a todo vetor em W. O conjunto de todos os vetores ortogonais a W é chamado de ***complemento ortogonal de*** W, denotado por $W^{\perp}$.[2] Isto é,

$$W^{\perp} = \{\mathbf{v} \text{ em } \mathbb{R}^n : \mathbf{v} \cdot \mathbf{w} = 0 \text{ para todo } \mathbf{w} \text{ em } W\}$$

EXEMPLO 1 Se W é um plano que passa pela origem em $\mathbb{R}^3$ e se ℓ é a reta que passa pela origem e é perpendicular a W (isto é, paralela ao vetor normal a W), então todo vetor $\mathbf{v}$ em ℓ é ortogonal a todo vetor $\mathbf{w}$ em W; portanto, $\ell = W^{\perp}$. Além disso, W consiste *precisamente* nos vetores $\mathbf{w}$ que são ortogonais a todo vetor $\mathbf{v}$ em ℓ; por isso temos também que $W = \ell^{\perp}$. A Figura 1 ilustra essa situação.

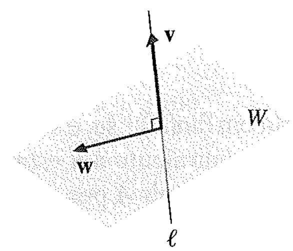

Figura 1 $\ell = W^{\perp}$ e $W = \ell^{\perp}$

No Exemplo 1, o complemento ortogonal de um subespaço é outro subespaço. Além disso, o complemento do complemento de um subespaço é o subespaço original. Essas propriedades são válidas em geral e são as propriedades (a) e (b) demonstradas no Teorema 1. As propriedades (c) e (d) também serão úteis. (Lembre-se de que a *intersecção* $A \cap B$ dos conjuntos A e B é o conjunto formado pelos seus elementos comuns. Veja o Apêndice A.)

TEOREMA 1 Seja W um subespaço de $\mathbb{R}^n$.

a. $W^{\perp}$ é um subespaço de $\mathbb{R}^n$.
b. $(W^{\perp})^{\perp} = W$
c. $W \cap W^{\perp} = \{\mathbf{0}\}$
d. Se $W = \text{ger}(\mathbf{w}_1, \ldots, \mathbf{w}_k)$, então $\mathbf{v}$ está em W se, e somente se, $\mathbf{v} \cdot \mathbf{w}_i = 0$ para todo $i = 1, \ldots, k$.

DEMONSTRAÇÃO: (a) Como $\mathbf{0} \cdot \mathbf{w} = 0$ para todo $\mathbf{w}$ em W, $\mathbf{0}$ está em $W^{\perp}$. Sejam $\mathbf{u}$ e $\mathbf{v}$ em $W^{\perp}$ e seja c um escalar. Então

$$\mathbf{u} \cdot \mathbf{w} = \mathbf{v} \cdot \mathbf{w} = 0 \quad \text{para todo} \quad \mathbf{w} \quad \text{em} \quad W$$

Logo,

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

e $\mathbf{u} + \mathbf{w}$ está em $W^{\perp}$.

Também temos que

$$(c\mathbf{u}) \cdot \mathbf{w} = c(\mathbf{u} \cdot \mathbf{w}) = c(0) = 0$$

onde vemos que $c\mathbf{u}$ está em $W^{\perp}$. Conseqüentemente, $W^{\perp}$ é um subespaço de $\mathbb{R}^n$.

[2] $W^{\perp}$ é pronunciado "W ortogonal".

(b) Provaremos essa propriedade como o Corolário 4.
(c) Pediremos que você prove essa propriedade no Exercício 23.
(d) Pediremos que você prove essa propriedade no Exercício 24. ◆

Podemos agora expressar algumas relações fundamentais envolvendo os subespaços associados a matrizes $m \times n$.

◆ TEOREMA 2

Seja A uma matriz $m \times n$. O complemento ortogonal do espaço linha de A é o espaço anulado por A, e o complemento ortogonal do espaço coluna de A é o espaço anulado por A^T:

$$(\text{lin}(A))^{\perp} = \text{anul}(A) \quad \text{e} \quad (\text{col}(A))^{\perp} = \text{anul}(A^T)$$

DEMONSTRAÇÃO: Se $\mathbf{x}$ é um vetor em $\mathbb{R}^n$, $\mathbf{x}$ está em $(\text{lin}(A))^{\perp}$ se, e somente se, $\mathbf{x}$ é ortogonal a toda linha de A. Mas isso é verdade se, e somente se, $A\mathbf{x} = \mathbf{0}$, o que é equivalente a $\mathbf{x}$ estar em anul(A), e, assim, provamos a primeira igualdade. Para provar a segunda igualdade, basta substituir A por A^T e usar o fato de que $\text{lin}(A^T) = \text{col}(A)$. ◆

Assim, matrizes $m \times n$ têm quatro subespaços: lin(A), anul(A), col(A) e anul(A^T). Os dois primeiros são complementares em $\mathbb{R}^n$, e os dois últimos, complementares em $\mathbb{R}^m$. A matriz A define uma transformação linear de $\mathbb{R}^n$ em $\mathbb{R}^m$ cuja imagem é col(A). Além disso, essa transformação linear manda anul(A) a $\mathbf{0}$ em $\mathbb{R}^m$. A Figura 2 é um esquema para ilustrar essa idéia. Esses quatro subespaços são chamados ***subespaços fundamentais*** da matriz A $m \times n$.

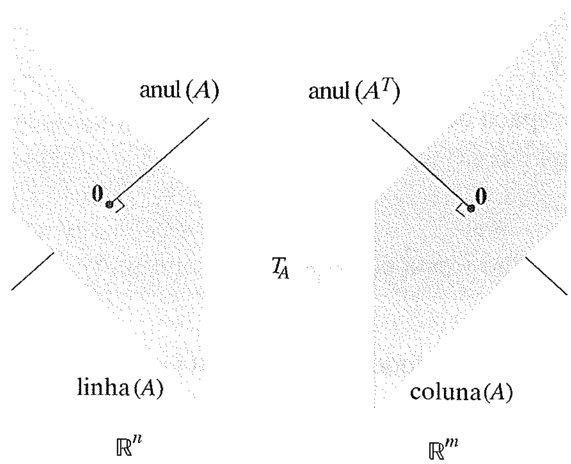

Figura 2 Os quatro subespaços fundamentais

EXEMPLO 2 Encontre bases para os quatro espaços fundamentais da matriz

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

e verifique o Teorema 2.

SOLUÇÃO: Nos Exemplos 9, 11, e 12 da Seção 3.5, calculamos bases para o espaço linha, o espaço coluna e o espaço anulado por A. Vimos que lin(A) = ger($\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$), onde

$$\mathbf{u}_1 = [1 \quad 0 \quad 1 \quad 0 \quad -1], \; \mathbf{u}_2 = [0 \quad 1 \quad 2 \quad 0 \quad 3], \; \mathbf{u}_3 = [0 \quad 0 \quad 1 \quad 4]$$

Além disso, anul(A) = ger($\mathbf{x}_1, \mathbf{x}_2$), onde

$$\mathbf{x}_1 = \begin{bmatrix} -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ -3 \\ 0 \\ -4 \\ 1 \end{bmatrix}$$

Para mostrar que $(\text{lin}(A))^\perp = \text{anul}(A)$, é suficiente provar que cada $\mathbf{u}_i$ é ortogonal a cada $\mathbf{x}_j$, o que é um exercício fácil. (Por que isso é suficiente?)
O espaço coluna de A é col(A) = ger($\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$), onde

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 1 \\ 1 \\ -2 \\ 1 \end{bmatrix}$$

Ainda precisamos calcular o espaço anulado por A^T. O escalonamento por linhas fornece

$$[A^T \mid \mathbf{0}] = \left[\begin{array}{cccc|c} 1 & 2 & -3 & 4 & 0 \\ 1 & -1 & 2 & 1 & 0 \\ 3 & 0 & 1 & 6 & 0 \\ 1 & 1 & -2 & 1 & 0 \\ 6 & -1 & 1 & 3 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 6 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Assim, se $\mathbf{y}$ está no espaço anulado por A^T, então $y_1 = -y_4$, $y_2 = -6y_4$ e $y_3 = -3y_4$. Conseqüentemente,

$$\text{anul}(A^T) = \left\{ \begin{bmatrix} -y_4 \\ -6y_4 \\ -3y_4 \\ y_4 \end{bmatrix} \right\} = \text{ger}\left(\begin{bmatrix} -1 \\ -6 \\ -3 \\ 1 \end{bmatrix} \right)$$

e é fácil verificar que esse vetor é ortogonal a $\mathbf{a}_1$, $\mathbf{a}_2$ e $\mathbf{a}_3$.

O método do Exemplo 2 é facilmente adaptado a outras situações.

EXEMPLO 3 Seja W o subespaço de $\mathbb{R}^5$ gerado por

$$\mathbf{w}_1 = \begin{bmatrix} 1 \\ -3 \\ 5 \\ 0 \\ 5 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} -1 \\ 1 \\ 2 \\ -2 \\ 3 \end{bmatrix}, \quad \mathbf{w}_3 = \begin{bmatrix} 0 \\ -1 \\ 4 \\ -1 \\ 5 \end{bmatrix}$$

Encontre uma base para $W^\perp$.

SOLUÇÃO: O subespaço W gerado por $\mathbf{w}_1$, $\mathbf{w}_2$ e $\mathbf{w}_3$ é o mesmo que o espaço coluna de

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

Logo, pelo Teorema 2, $W^{\perp}(\text{col}(A))^{\perp} = \text{anul}(A^T)$, e podemos proceder como no exemplo anterior. Calculamos

$$[A^T \mid \mathbf{0}] = \left[\begin{array}{rrrrr|r} 1 & -3 & 5 & 0 & 5 & 0 \\ -1 & 1 & 2 & -2 & 3 & 0 \\ 0 & -1 & 4 & -1 & 5 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{rrrrr|r} 1 & 0 & 0 & 3 & 4 & 0 \\ 0 & 1 & 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 \end{array}\right]$$

Logo, $\mathbf{y}$ está em $W^{\perp}$ se, e somente se, $y_1 = -3y_4 - 4y_5$, $y_2 = -y_4 - 3y_5$ e $y_3 = -2y_5$. Em conseqüência,

$$W^{\perp} = \left\{ \begin{bmatrix} -3y_4 - 4y_5 \\ -y_4 - 3y_5 \\ -2y_5 \\ y_4 \\ y_5 \end{bmatrix} \right\} = \text{ger}\left(\begin{bmatrix} -3 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ -3 \\ -2 \\ 0 \\ 1 \end{bmatrix} \right)$$

e esses dois vetores formam uma base para W.

Projeções Ortogonais

Lembre-se de que, em $\mathbb{R}^2$, a projeção ortogonal de um vetor $\mathbf{v}$ sobre um vetor não nulo $\mathbf{u}$ é dada por

$$\text{proj}_{\mathbf{u}}(\mathbf{v}) = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{u} \cdot \mathbf{u}}\right)\mathbf{u}$$

Além disso, o vetor $\text{perp}_{\mathbf{u}}(\mathbf{v}) = \mathbf{v} - \text{proj}_{\mathbf{u}}(\mathbf{v})$ é ortogonal a $\text{proj}_{\mathbf{u}}(\mathbf{v})$, e podemos decompor $\mathbf{v}$ como

$$\mathbf{v} = \text{proj}_{\mathbf{u}}(\mathbf{v}) + \text{perp}_{\mathbf{u}}(\mathbf{v})$$

como é mostrado na Figura 3.

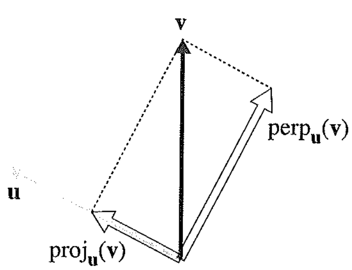

Figura 3 $\mathbf{v} = \text{proj}_{\mathbf{u}}(\mathbf{v}) + \text{perp}_{\mathbf{u}}(\mathbf{v})$

Se fizermos $W = \text{ger}(\mathbf{u})$, então $\mathbf{w} = \text{proj}_{\mathbf{u}}(\mathbf{v})$ está em W, e $\mathbf{w}^{\perp} = \text{perp}_{\mathbf{u}}(\mathbf{v})$ está em $W^{\perp}$. Temos então um modo de "decompor" $\mathbf{v}$ como soma de dois vetores, com um deles em W e o outro ortogonal a W — precisamente, $\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$. Generalizaremos essa idéia para $\mathbb{R}^n$.

Definição Seja W um subespaço de $\mathbb{R}^n$ e seja $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ uma base ortogonal de W. Para um vetor $\mathbf{v}$ em $\mathbb{R}^n$, a ***projeção ortogonal de* v *em* W** é definida por

$$\mathrm{proj}_W(\mathbf{v}) = \left(\frac{\mathbf{u}_1 \cdot \mathbf{v}}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right)\mathbf{u}_1 + \cdots + \left(\frac{\mathbf{u}_k \cdot \mathbf{v}}{\mathbf{u}_k \cdot \mathbf{u}_k}\right)\mathbf{u}_k$$

A ***componente ortogonal de* v *em relação a* W** é o vetor

$$\mathrm{perp}_W(\mathbf{v}) = \mathbf{v} - \mathrm{proj}_W(\mathbf{v})$$

Cada somando na definição de $\mathrm{proj}_W(\mathbf{v})$ é também uma projeção sobre um único vetor (ou, equivalentemente, sobre o espaço unidimensional gerado por ele — em nossa primeira definição). Assim, usando a notação de nossa definição anterior, podemos escrever

$$\mathrm{proj}_W(\mathbf{v}) = \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}) + \cdots + \mathrm{proj}_{\mathbf{u}_k}(\mathbf{v})$$

Como os vetores $\mathbf{u}_i$ são ortogonais, a projeção de $\mathbf{v}$ sobre W é a soma de suas projeções ortogonais sobre subespaços unidimensionais que são mutuamente ortogonais. A Figura 4 ilustra essa situação com $W = ger(\mathbf{u}_1, \mathbf{u}_2)$, $\mathbf{p} = \mathrm{proj}_W(\mathbf{v})$, $\mathbf{p}_1 = \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v})$ e $\mathbf{p}_2 = \mathrm{proj}_{\mathbf{u}_2}(\mathbf{v})$.

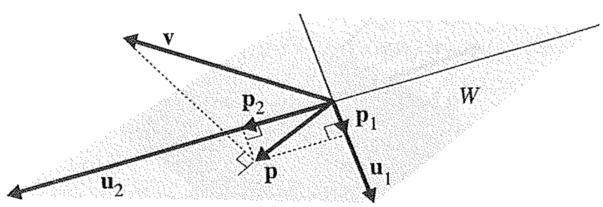

Figura 4 $\mathbf{p} = \mathbf{p}_1 + \mathbf{p}_2$

Como um caso especial da definição de $\mathrm{proj}_W(\mathbf{v})$, também obtemos agora uma boa interpretação do Teorema 2 da Seção 5.2. Em termos da nossa presente notação e terminologia, aquele teorema afirma que, se $\mathbf{w}$ está no subespaço W de $\mathbb{R}^n$, cuja base ortogonal é $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$, então

$$\begin{aligned}\mathbf{w} &= \left(\frac{\mathbf{w} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 + \cdots + \left(\frac{\mathbf{w} \cdot \mathbf{v}_k}{\mathbf{v}_k \cdot \mathbf{v}_k}\right)\mathbf{v}_k \\ &= \mathrm{proj}_{\mathbf{v}_1}(\mathbf{w}) + \cdots + \mathrm{proj}_{\mathbf{v}_k}(\mathbf{w})\end{aligned}$$

Logo, $\mathbf{w}$ é decomposto em uma soma de projeções ortogonais sobre subespaços mutuamente ortogonais e unidimensionais de W.

A definição anterior parece depender da escolha da base ortogonal, isto é, uma base ortogonal diferente $\{\mathbf{u}'_1, \ldots, \mathbf{u}'_k\}$ para W parece fornecer uma $\mathrm{proj}_W(\mathbf{v})$ e uma $\mathrm{perp}_W(\mathbf{v})$ "diferentes". Por sorte não é esse o caso, como provaremos logo. Por enquanto, vamos nos contentar com um exemplo.

EXEMPLO 4 Seja W um plano em $\mathbb{R}^3$ de equação $x - y + 2z = 0$, e seja $\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$. Encontre a projeção ortogonal de $\mathbf{v}$ sobre W e a componente ortogonal de $\mathbf{v}$ em relação a W.

SOLUÇÃO: No Exemplo 3 da Seção 5.2, encontramos uma base ortogonal para W. Tomando

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{e} \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

temos

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = 2 \quad \mathbf{u}_2 \cdot \mathbf{v} = -2$$
$$\mathbf{u}_1 \cdot \mathbf{u}_1 = 2 \quad \mathbf{u}_2 \cdot \mathbf{u}_2 = 3$$

Logo,

$$\begin{aligned} \text{proj}_W(\mathbf{v}) &= \left(\frac{\mathbf{u}_1 \cdot \mathbf{v}}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right)\mathbf{u}_1 + \left(\frac{\mathbf{u}_2 \cdot \mathbf{v}}{\mathbf{u}_2 \cdot \mathbf{u}_2}\right)\mathbf{u}_2 \\ &= \tfrac{2}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \tfrac{2}{3}\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{5}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \end{bmatrix} \end{aligned}$$

e

$$\text{perp}_W(\mathbf{v}) = \mathbf{v} - \text{proj}_W(\mathbf{v}) = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} - \begin{bmatrix} \frac{5}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{4}{3} \\ -\frac{4}{3} \\ \frac{8}{3} \end{bmatrix}$$

É fácil ver que $\text{proj}_W(\mathbf{v})$ está em W, pois satisfaz a equação do plano. É igualmente fácil ver que $\text{perp}_W(\mathbf{v})$ é ortogonal a W, já que é um múltiplo escalar de seu vetor normal $\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$ para W. (Veja a Figura 5.)

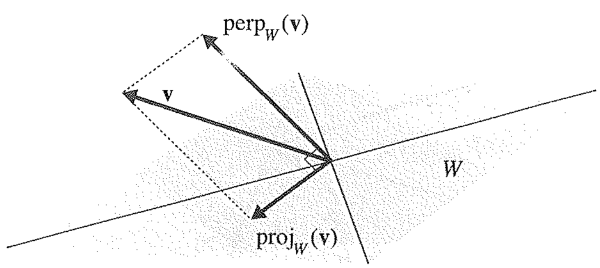

Figura 5 $\mathbf{v} = \text{proj}_W(\mathbf{v}) + \text{perp}_W(\mathbf{v})$

O próximo teorema mostra que sempre podemos decompor um vetor em relação a um subespaço e a seu complemento ortogonal.

◆ TEOREMA 3 O Teorema da Decomposição Ortogonal

Seja W um subespaço de $\mathbb{R}^n$ e seja $\mathbf{v}$ um vetor em $\mathbb{R}^n$. Existem vetores únicos $\mathbf{w}$ em W e $\mathbf{w}^\perp$ em $W^\perp$ tais que

$$\mathbf{v} = \mathbf{w} + \mathbf{w}^\perp$$

DEMONSTRAÇÃO: Precisamos provar duas coisas: que essa decomposição *existe* e que é *única*.

Para demonstrar a existência, escolhemos uma base ortogonal $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ de W. Sejam $\mathbf{w} = \text{proj}_W(\mathbf{v})$ e $\mathbf{w}^\perp = \text{perp}_W(\mathbf{v})$. Então:

$$\mathbf{w} + \mathbf{w}^\perp = \text{proj}_W(\mathbf{v}) + \text{perp}_W(\mathbf{v}) = \text{proj}_W(\mathbf{v}) + (\mathbf{v} - \text{proj}_W(\mathbf{v})) = \mathbf{v}$$

É claro que $\text{proj}_W(\mathbf{v})$ está em W, pois é uma combinação linear dos vetores $\mathbf{u}_1, \ldots, \mathbf{u}_k$ da base de W. Para mostrar que $\mathbf{w}^\perp$ está em $W^\perp$, é suficiente mostrar que $\mathbf{w}^\perp$ é ortogonal a cada vetor $\mathbf{u}_i$, pelo Teorema 1(d). Calculamos

$$
\begin{aligned}
\mathbf{u}_i \cdot \mathbf{w}^\perp &= \mathbf{u}_i \cdot \text{perp}_W(\mathbf{v}) \\
&= \mathbf{u}_i \cdot (\mathbf{v} - \text{proj}_W(\mathbf{v})) \\
&= \mathbf{u}_i \cdot \left(\mathbf{v} - \left(\frac{\mathbf{u}_1 \cdot \mathbf{v}}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right)\mathbf{u}_1 - \cdots - \left(\frac{\mathbf{u}_k \cdot \mathbf{v}}{\mathbf{u}_k \cdot \mathbf{u}_k}\right)\mathbf{u}_k \right) \\
&= \mathbf{u}_i \cdot \mathbf{v} - \left(\frac{\mathbf{u}_1 \cdot \mathbf{v}}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right)(\mathbf{u}_i \cdot \mathbf{u}_1) - \cdots - \left(\frac{\mathbf{u}_i \cdot \mathbf{v}}{\mathbf{u}_i \cdot \mathbf{u}_i}\right)(\mathbf{u}_i \cdot \mathbf{u}_i) - \cdots - \left(\frac{\mathbf{u}_k \cdot \mathbf{v}}{\mathbf{u}_k \cdot \mathbf{u}_k}\right)(\mathbf{u}_i \cdot \mathbf{u}_k) \\
&= \mathbf{u}_i \cdot \mathbf{v} - 0 - \cdots - \left(\frac{\mathbf{u}_i \cdot \mathbf{v}}{\mathbf{u}_i \cdot \mathbf{u}_i}\right)(\mathbf{u}_i \cdot \mathbf{u}_i) - \cdots - 0 \\
&= \mathbf{u}_i \cdot \mathbf{v} - \mathbf{u}_i \cdot \mathbf{v} = 0
\end{aligned}
$$

pois $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ para $j \neq i$. Isso prova que $\mathbf{w}^\perp$ está em $W^\perp$ e completa esta parte da demonstração.

Para provar a unicidade da decomposição, vamos considerar outra decomposição: $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_1^\perp$, onde $\mathbf{w}_1$ está em W e $\mathbf{w}_1^\perp$ está em $W^\perp$. Então, $\mathbf{w} + \mathbf{w}^\perp = \mathbf{w}_1 + \mathbf{w}_1^\perp$, de modo que

$$\mathbf{w} - \mathbf{w}_1 = \mathbf{w}_1^\perp - \mathbf{w}^\perp$$

No entanto, como $\mathbf{w} - \mathbf{w}_1$ está em W e $\mathbf{w}_1^\perp - \mathbf{w}^\perp$ está em $W^\perp$ (pois eles são subespaços), sabemos que esse vetor comum está em $W \cap W^\perp = \{\mathbf{0}\}$ [usando o Teorema 1(c)]. Logo,

$$\mathbf{w} - \mathbf{w}_1 = \mathbf{w}_1^\perp - \mathbf{w}^\perp = \mathbf{0}$$

e temos $\mathbf{w} = \mathbf{w}_1$ e $\mathbf{w}_1^\perp = \mathbf{w}^\perp$. ◆

O Exemplo 4 ilustrou o Teorema da Decomposição Ortogonal. Quando W é o subespaço de $\mathbb{R}^3$ dado pelo plano de equação $x - y + 2z = 0$, a decomposição ortogonal de $\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$, em relação a W, é $\mathbf{v} = \mathbf{w} + \mathbf{w}^\perp$, onde

$$\mathbf{w} = \text{proj}_W(\mathbf{v}) = \begin{bmatrix} \frac{5}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \end{bmatrix} \quad \text{e} \quad \mathbf{w}^\perp = \text{perp}_W(\mathbf{v}) = \begin{bmatrix} \frac{4}{3} \\ -\frac{4}{3} \\ \frac{8}{3} \end{bmatrix}$$

A unicidade da decomposição ortogonal garante que as definições de $\text{proj}_W(\mathbf{v})$ e $\text{perp}_W(\mathbf{v})$ não dependem da escolha da base ortogonal. O Teorema da Decomposição Ortogonal também nos permite demonstrar a propriedade (b) do Teorema 1. Vamos enunciar essa propriedade como um corolário do Teorema da Decomposição Ortogonal.

A palavra *corolário* vem da palavra latina *corollarium*, que se refere a uma grinalda dada como recompensa. Um corolário, portanto, é uma pequena recompensa extra que segue de um teorema.

◆ COROLÁRIO 4

Se W é um subespaço de $\mathbb{R}^n$, então

$$(W^\perp)^\perp = W$$

DEMONSTRAÇÃO: Se $\mathbf{w}$ está em W e $\mathbf{x}$ está em $W^\perp$, $\mathbf{w} \cdot \mathbf{x} = 0$. Mas isso implica que $\mathbf{w}$ está em $(W^\perp)^\perp$. Logo, $W \subseteq (W^\perp)^\perp$. Para provar que realmente a igualdade é válida aqui, suponha o contrário. Existe, então, um vetor $\mathbf{v}$ em $(W^\perp)^\perp$ que não está em W. Pelo Teorema 3, podemos escrever $\mathbf{v} = \mathbf{w} + \mathbf{w}^\perp$ para vetores (determinados de modo único) $\mathbf{w}$ em W e $\mathbf{w}^\perp$ em $W^\perp$. Mas agora

$$0 = \mathbf{v} \cdot \mathbf{w}^\perp = (\mathbf{w} + \mathbf{w}^\perp) \cdot \mathbf{w}^\perp = \mathbf{w} \cdot \mathbf{w}^\perp + \mathbf{w}^\perp \cdot \mathbf{w}^\perp = 0 + \mathbf{w}^\perp \cdot \mathbf{w}^\perp = \mathbf{w}^\perp \cdot \mathbf{w}^\perp$$

de modo que $\mathbf{w}^\perp = \mathbf{0}$. Logo, $\mathbf{v} = \mathbf{w} + \mathbf{w}^\perp = \mathbf{w}$, e, portanto, $\mathbf{v}$ *está* em W — o que é uma contradição. Concluímos que $(W^\perp)^\perp = W$, como queríamos. ◈

Há também uma boa relação entre as dimensões de W e $W^\perp$, expressa no Teorema 5.

◆ TEOREMA 5

Se W é um subespaço de $\mathbb{R}^n$,

$$\dim W + \dim W^\perp = n$$

DEMONSTRAÇÃO: Seja $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ uma base ortogonal de W e seja $\{\mathbf{v}_1, \ldots, \mathbf{v}_l\}$ uma base ortogonal de $W^\perp$. Então, $\dim W = k$ e $\dim W^\perp = l$. Seja $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_l\}$. Afirmamos que $\mathcal{B}$ é uma base ortogonal de $\mathbb{R}^n$.

Primeiro, notamos que, como cada $\mathbf{u}_i$ está em W e cada $\mathbf{v}_j$ está em $W^\perp$,

$$\mathbf{u}_i \cdot \mathbf{v}_j = 0 \quad \text{para todo} \quad i = 1, \ldots, k \text{ e } j = 1, \ldots, l$$

Assim, $\mathcal{B}$ é um conjunto ortogonal e, portanto, linearmente independente, de acordo com o Teorema 1 da Seção 5.2. A seguir, se $\mathbf{v}$ é um vetor em $\mathbb{R}^n$, o Teorema da Decomposição Ortogonal nos diz que $\mathbf{v} = \mathbf{w} + \mathbf{w}^\perp$ para algum $\mathbf{w}$ em W e $\mathbf{w}^\perp$ em $W^\perp$. Já que $\mathbf{w}$ pode ser escrito como uma combinação linear dos vetores $\mathbf{u}_i$ e $\mathbf{w}^\perp$ pode ser escrito como uma combinação linear dos vetores $\mathbf{v}_j$, $\mathbf{v}$ pode ser escrito como uma combinação linear dos vetores de $\mathcal{B}$. Logo, $\mathcal{B}$ também gera $\mathbb{R}^n$ e, portanto, é uma base de $\mathbb{R}^n$. Segue que $k + l = n = \dim \mathbb{R}^n$, ou

$$\dim W + \dim W^\perp = n \quad ◈$$

Aplicando esse resultado aos subespaços fundamentais de uma matriz, obtemos como bônus uma demonstração rápida do Teorema do Posto (Teorema 8 da Seção 3.5), enunciado aqui como Corolário 6.

◆ COROLÁRIO 6

Se A é uma matriz $m \times n$,

$$\text{posto}(A) + \text{nulidade}(A) = n$$

DEMONSTRAÇÃO: No Teorema 5, considere $W = \text{lin}(A)$. Então, $W^\perp = \text{anul}(A)$, pelo Teorema 2, de modo que $\dim W = \text{posto}(A)$ e $\dim W^\perp = \text{nulidade}(A)$. O resultado vem em seguida. ◈

Note que obtemos a identidade simétrica tomando $W = \text{col}(A)$ [e então $W^\perp = \text{anul}(A^T)$]:

$$\text{posto}(A) + \text{nulidade}(A^T) = m$$

As Seções 5.2 e 5.3 ilustraram algumas das vantagens de trabalhar com bases ortogonais. Entretanto, não demonstramos ainda que todo subespaço *tem* uma base ortogonal, nem mostramos um modo de construir essa base (exceto em alguns exemplos particulares, como o Exemplo 3 da Seção 5.2). Esses problemas serão o assunto da próxima seção.

◆ EXERCÍCIOS 5.3 ◆

Nos Exercícios de 1 a 6, encontre o complemento ortogonal $W^{\perp}$ de *W e dê uma base de* $W^{\perp}$.

1. $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : 2x - y = 0 \right\}$

2. $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : 3x + 4y = 0 \right\}$

3. $W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : x + y - z = 0 \right\}$

4. $W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : 2x - y + 3z = 0 \right\}$

5. $W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : x = t, y = -t, z = 3t \right\}$

6. $W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : x = 2t, y = 2t, z = -t \right\}$

Nos Exercícios 7 e 8, ache bases para o espaço linha de A e para o espaço anulado por A. Verifique que cada vetor em lin(A) é ortogonal a todo vetor em anul(A).

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

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

Nos Exercícios 9 e 10, encontre bases para o espaço coluna de A e para o espaço anulado por A^T, *sendo A a matriz no exercício dado. Verifique que todo vetor em col(A) é ortogonal a todo vetor em anul*(A^T).

9. Exercício 7 **10.** Exercício 8

Nos Exercícios de 11 a 14, seja W o subespaço gerado pelos vetores dados. Encontre uma base de $W^{\perp}$.

11. $\mathbf{w}_1 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}$

12. $\mathbf{w}_1 = \begin{bmatrix} 1 \\ -1 \\ 3 \\ -2 \end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix} 0 \\ 1 \\ -2 \\ 1 \end{bmatrix}$

13. $\mathbf{w}_1 = \begin{bmatrix} 2 \\ -1 \\ 6 \\ 3 \end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix} -1 \\ 2 \\ -3 \\ -2 \end{bmatrix}, \mathbf{w}_3 = \begin{bmatrix} 2 \\ 5 \\ 6 \\ 1 \end{bmatrix}$

14. $\mathbf{w}_1 = \begin{bmatrix} 4 \\ 6 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \\ -3 \end{bmatrix}, \mathbf{w}_3 = \begin{bmatrix} 2 \\ 2 \\ 2 \\ -1 \\ 2 \end{bmatrix}$

Nos Exercícios de 15 a 18, encontre a projeção ortogonal de **v** *sobre o subespaço W gerado pelos vetores* $\mathbf{u}_i$. *(Você pode assumir que os vetores* $\mathbf{u}_i$ *são ortogonais.)*

15. $\mathbf{v} = \begin{bmatrix} 7 \\ -4 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

16. $\mathbf{v} = \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$

17. $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ 4 \end{bmatrix}$

18. $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \\ 4 \\ -3 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$

Nos Exercícios de 19 a 22, encontre a decomposição ortogonal de **v** *em relação a W.*

19. $\mathbf{v} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}, W = \text{ger}\left(\begin{bmatrix} 1 \\ 3 \end{bmatrix} \right)$

20. $\mathbf{v} = \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix}, W = \text{ger}\left(\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}\right)$

21. $\mathbf{v} = \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix}, W = \text{ger}\left(\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}\right)$

22. $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ 5 \\ 3 \end{bmatrix}, W = \text{ger}\left(\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}\right)$

23. Prove o Teorema 1(c).

24. Prove o Teorema 1(d).

25. Sejam W um subespaço de $\mathbb{R}^n$ e $\mathbf{v}$ um vetor em $\mathbb{R}^n$. Suponha que $\mathbf{w}$ e $\mathbf{w}'$ são vetores ortogonais com $\mathbf{w}$ em W, e que $\mathbf{v} = \mathbf{w} + \mathbf{w}'$. O vetor $\mathbf{w}'$ tem necessariamente que estar em $W^\perp$? Prove ou dê um contra-exemplo.

26. Sejam $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ uma base ortogonal de $\mathbb{R}^n$ e $W = \text{ger}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$. É necessariamente verdade que $W^\perp = \text{ger}(\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n)$? Prove ou dê um contra-exemplo.

5.4 O Processo de Gram-Schmidt e a Fatoração *QR*

Nesta seção, apresentaremos um método simples para construir uma base ortogonal (ou ortonormal) para qualquer subespaço de $\mathbb{R}^n$. Esse método nos levará a uma das mais úteis fatorações de matrizes.

O Processo de Gram-Schmidt

Gostaríamos de poder achar uma base ortogonal para um subespaço W de $\mathbb{R}^n$. A idéia é começar com uma base $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ de W e "ortogonalizar" um vetor de cada vez. Ilustraremos essa construção básica com o subespaço W do Exemplo 3 da Seção 5.2.

EXEMPLO 1 Seja $W = \text{ger}(\mathbf{x}_1, \mathbf{x}_2)$, onde

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{e} \quad \mathbf{x}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

Construa uma base ortogonal para W.

SOLUÇÃO: Começando por $\mathbf{x}_1$, obtemos um segundo vetor ortogonal a ele tomando a componente ortogonal de $\mathbf{x}_2$ em relação a $\mathbf{x}_1$ (Figura 1). Algebricamente, seja $\mathbf{v}_1 = \mathbf{x}_1$. Então

$$\begin{aligned} \mathbf{v}_2 &= \text{perp}_{\mathbf{x}_1}(\mathbf{x}_2) = \mathbf{x}_2 - \text{proj}_{\mathbf{x}_1}(\mathbf{x}_2) \\ &= \mathbf{x}_2 - \left(\frac{\mathbf{x}_1 \cdot \mathbf{x}_2}{\mathbf{x}_1 \cdot \mathbf{x}_1}\right)\mathbf{x}_1 \\ &= \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} - \left(\frac{-2}{2}\right)\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \end{aligned}$$

Então, $\{\mathbf{v}_1, \mathbf{v}_2\}$ é um conjunto ortogonal de vetores em W. Assim, $\{\mathbf{v}_1, \mathbf{v}_2\}$ é linearmente independente e, portanto, é uma base para W, já que $\dim W = 2$.

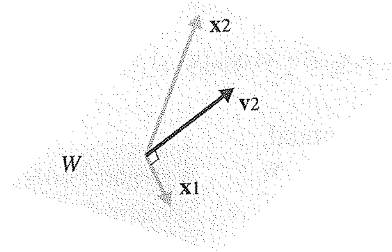

Figura 1 Construindo $\mathbf{v}_2$ ortogonal a $\mathbf{x}_1$

Observação: ◆ Note que esse método depende da *ordem* dos vetores na base original. No Exemplo 1, se tivéssemos tomado $\mathbf{x}_1 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$ e $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, teríamos obtido uma base ortogonal diferente para W. (Verifique isso.)

A generalização desse método para mais de dois vetores começa como no Exemplo 1. O processo é construir iterativamente os vetores subseqüentes fazendo-os ortogonais a todos os vetores que já foram construídos. Esse método é conhecido como ***Processo de Gram-Schmidt***.

◆ TEOREMA 1 O Processo de Gram-Schmidt

Seja $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ uma base para um subespaço W de $\mathbb{R}^n$ e defina o seguinte:

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1, & W_1 &= \text{ger}(\mathbf{x}_1)\\
\mathbf{v}_2 &= \mathbf{x}_2 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_2}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1, & W_2 &= \text{ger}(\mathbf{x}_1, \mathbf{x}_2)\\
\mathbf{v}_3 &= \mathbf{x}_3 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_3}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2 \cdot \mathbf{x}_3}{\mathbf{v}_2 \cdot \mathbf{v}_2}\right)\mathbf{v}_2, & W_3 &= \text{ger}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)\\
&\ \vdots\\
\mathbf{v}_k &= \mathbf{x}_k - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_k}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2 \cdot \mathbf{x}_k}{\mathbf{v}_2 \cdot \mathbf{v}_2}\right)\mathbf{v}_2 - \cdots\\
&\quad - \left(\frac{\mathbf{v}_{k-1} \cdot \mathbf{x}_k}{\mathbf{v}_{k-1} \cdot \mathbf{v}_{k-1}}\right)\mathbf{v}_{k-1}, & W_k &= \text{ger}(\mathbf{x}_1, \ldots, \mathbf{x}_k)
\end{aligned}$$

Então, para cada $i = 1, \ldots, k$, $\{\mathbf{v}_1, \ldots, \mathbf{v}_i\}$ é uma base ortogonal de W_i. Em particular, $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ é uma base de W.

Enunciado sucintamente, o Teorema 1 diz que todo subespaço de $\mathbb{R}^n$ tem uma base ortogonal e fornece um algoritmo para construir uma base dessas.

DEMONSTRAÇÃO: Provaremos, por indução, que, para cada $i = 1, \ldots, k$, $\{\mathbf{v}_1, \ldots, \mathbf{v}_i\}$ é uma base ortogonal de W_i.

Como $\mathbf{v}_1 = \mathbf{x}_1$, é claro que $\{\mathbf{v}_1\}$ é uma base (ortogonal) de $W_1 = \text{ger}(\mathbf{x}_1)$. Assuma agora que, para algum $i < k$, $\{\mathbf{v}_1, \ldots, \mathbf{v}_i\}$ é uma base ortogonal de W_i. Então:

$$\mathbf{v}_{i+1} = \mathbf{x}_{i+1} - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_{i+1}}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2 \cdot \mathbf{x}_{i+1}}{\mathbf{v}_2 \cdot \mathbf{v}_2}\right)\mathbf{v}_2 - \cdots - \left(\frac{\mathbf{v}_i \cdot \mathbf{x}_{i+1}}{\mathbf{v}_i \cdot \mathbf{v}_i}\right)\mathbf{v}_i$$

Pela hipótese de indução, $\{\mathbf{v}_1, \ldots, \mathbf{v}_i\}$ é uma base ortogonal para ger$(\mathbf{x}_1, \ldots, \mathbf{x}_i) = W_i$. Logo,

$$\mathbf{v}_{i+1} = \mathbf{x}_{i+1} - \text{proj}_{W_i}(\mathbf{x}_{i+1}) = \text{perp}_{W_i}(\mathbf{x}_{i+1})$$

Então, pelo Teorema da Decomposição Ortogonal, $\mathbf{v}_{i+1}$ é ortogonal a W_i. Por definição, $\mathbf{v}_1, \ldots, \mathbf{v}_i$ são combinações lineares de $\mathbf{x}_1, \ldots, \mathbf{x}_i$ e, portanto, estão em W_i. Conseqüentemente $\{\mathbf{v}_1, \ldots, \mathbf{v}_{i+1}\}$ é um subconjunto ortogonal de W_{i+1}.

Além disso, $\mathbf{v}_{i+1} \neq \mathbf{0}$, pois, caso contrário, $\mathbf{x}_{i+1} = \text{proj}_{W_i}(\mathbf{x}_{i+1})$, o que implica que $\mathbf{x}_{i+1}$ está em W_i. Mas isso é impossível, já que $W_i = \text{ger}(\mathbf{x}_1, \ldots, \mathbf{x}_i)$ e $\{\mathbf{x}_1, \ldots, \mathbf{x}_{i+1}\}$ é linearmente independente. (Por quê?) Concluímos que $\{\mathbf{v}_1, \ldots, \mathbf{v}_{i+1}\}$ é um conjunto com $i + 1$ vetores linearmente independentes em W_{i+1}. Conseqüentemente, $\{\mathbf{v}_1, \ldots, \mathbf{v}_{i+1}\}$ é uma base de W_{i+1}, pois dim $W_{i+1} = i + 1$. Isso completa a demonstração. ◆

Se quisermos uma base ortonormal de W, precisamos simplesmente normalizar os vetores produzidos pelo Processo de Gram-Schmidt. Para cada i, substituímos $\mathbf{v}_i$ pelo vetor unitário $\mathbf{q}_i = (1/\|\mathbf{v}_i\|)\mathbf{v}_i$.

Jörgen Pedersen Gram (1850–1916) foi um atuário dinamarquês (estatístico da área de seguros) que era interessado na ciência de medir. Ele foi o primeiro a publicar, em 1883, o processo que leva seu nome, em um artigo sobre os mínimos quadrados. Erhard Schmidt (1876–1959) foi um matemático alemão que estudou sob a orientação do grande David Hilbert e é considerado um dos fundadores de um ramo da matemática conhecido como análise funcional. Sua contribuição ao Processo de Gram-Schmidt veio em um artigo sobre equações integrais, em 1907, no qual ele escreveu os detalhes do método mais explicitamente do que Gram havia feito.

EXEMPLO 2 Aplique o Processo de Gram-Schmidt para construir uma base ortonormal para o subespaço $W = \text{ger}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ de $\mathbb{R}^4$, onde

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 2 \\ 2 \\ 1 \\ 2 \end{bmatrix}$$

SOLUÇÃO: Primeiro, note que o conjunto $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ é linearmente independente e, portanto, é uma base de W. Começamos pondo $\mathbf{v}_1 = \mathbf{x}_1$. Em seguida, calculamos a componente ortogonal de $\mathbf{x}_2$ em relação a $W_1 = \text{ger}(\mathbf{v}_1)$:

$$\begin{aligned}
\mathbf{v}_2 = \text{perp}_{W_1}(\mathbf{x}_2) &= \mathbf{x}_2 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_2}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 \\
&= \begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix} - \left(\tfrac{2}{4}\right)\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} \frac{3}{2} \\ \frac{3}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}
\end{aligned}$$

Para facilitar os cálculos, uma boa idéia, nesse ponto, é "simplificar" $\mathbf{v}_2$ para eliminar frações. No fim, podemos normalizar o conjunto ortogonal que estamos construindo para obter um conjunto ortonormal; assim, podemos substituir $\mathbf{v}_i$ por qualquer múltiplo escalar conveniente sem afetar o resultado final. De acordo com isso, substituímos $\mathbf{v}_2$ por

$$\mathbf{v}_2' = 2\mathbf{v}_2 = \begin{bmatrix} 3 \\ 3 \\ 1 \\ 1 \end{bmatrix}$$

Agora, achamos a componente ortogonal de $\mathbf{x}_3$ em relação a

$$W_2 = \text{ger}(\mathbf{x}_1, \mathbf{x}_2) = \text{ger}(\mathbf{v}_1, \mathbf{v}_2) = \text{ger}(\mathbf{v}_1, \mathbf{v}_2')$$

usando a base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2'\}$:

$$\begin{aligned}\mathbf{v}_3 = \text{perp}_{W_2}(\mathbf{x}_3) &= \mathbf{x}_3 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_3}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2' \cdot \mathbf{x}_3}{\mathbf{v}_2' \cdot \mathbf{v}_2'}\right)\mathbf{v}_2' \\ &= \begin{bmatrix} 2 \\ 2 \\ 1 \\ 2 \end{bmatrix} - \left(\tfrac{1}{4}\right)\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} - \left(\tfrac{15}{20}\right)\begin{bmatrix} 3 \\ 3 \\ 1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 1 \end{bmatrix}\end{aligned}$$

Multiplicamos agora pelo escalar 2 para obter $\mathbf{v}_3' = 2\mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$.

Temos agora uma base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_3'\}$ de W. (Verifique para ter certeza de que esses vetores são ortogonais.) Para obter uma base ortonormal, normalizamos cada vetor:

$$\mathbf{q}_1 = \left(\frac{1}{\|\mathbf{v}_1\|}\right)\mathbf{v}_1 = \left(\frac{1}{2}\right)\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{bmatrix}$$

$$\mathbf{q}_2 = \left(\frac{1}{\|\mathbf{v}_2'\|}\right)\mathbf{v}_2' = \left(\frac{1}{2\sqrt{5}}\right)\begin{bmatrix} 3 \\ 3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3/2\sqrt{5} \\ 3/2\sqrt{5} \\ 1/2\sqrt{5} \\ 1/2\sqrt{5} \end{bmatrix} = \begin{bmatrix} 3\sqrt{5}/10 \\ 3\sqrt{5}/10 \\ \sqrt{5}/10 \\ \sqrt{5}/10 \end{bmatrix}$$

$$\mathbf{q}_3 = \left(\frac{1}{\|\mathbf{v}_3'\|}\right)\mathbf{v}_3' = \left(\frac{1}{\sqrt{6}}\right)\begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{6} \\ 0 \\ 1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix} = \begin{bmatrix} -\sqrt{6}/6 \\ 0 \\ \sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix}$$

Então, $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$ é uma base ortogonal de W.

O Processo de Gram-Schmidt é muito útil para construir uma base ortogonal que contenha um vetor específico. O próximo exemplo ilustra essa aplicação.

EXEMPLO 3 Encontre uma base ortogonal de $\mathbb{R}^3$ que contenha o vetor

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

SOLUÇÃO: Primeiro, achamos uma base qualquer de $\mathbb{R}^3$ que contenha $\mathbf{v}_1$. Se tomarmos

$$\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{e} \quad \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

então $\{\mathbf{v}_1, \mathbf{x}_2, \mathbf{x}_3\}$ obviamente será uma base de $\mathbb{R}^3$. (Por quê?) Aplicamos agora o Processo de Gram-Schmidt a essa base para obter

$$\mathbf{v}_2 = \mathbf{x}_2 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_2}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \left(\tfrac{2}{14}\right)\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{7} \\ \frac{5}{7} \\ -\frac{3}{7} \end{bmatrix}, \quad \mathbf{v}_2' = \begin{bmatrix} -1 \\ 5 \\ -3 \end{bmatrix}$$

e, finalmente,

$$\mathbf{v}_3 = \mathbf{x}_3 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_3}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2' \cdot \mathbf{x}_3}{\mathbf{v}_2' \cdot \mathbf{v}_2'}\right)\mathbf{v}_2' = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \left(\tfrac{3}{14}\right)\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \left(\tfrac{-3}{35}\right)\begin{bmatrix} -1 \\ 5 \\ -3 \end{bmatrix} = \begin{bmatrix} \frac{-3}{10} \\ 0 \\ \frac{1}{10} \end{bmatrix},$$

$$\mathbf{v}_3' = \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

Então, $\{\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_3'\}$ é uma base ortonormal de $\mathbb{R}^3$ que contém $\mathbf{v}_1$.

Analogamente, dado um vetor unitário, podemos achar uma base ortonormal que o contenha pelo método agora descrito, e então normalizar os vetores ortogonais obtidos.

Observação: ◆ Quando o Processo de Gram-Schmidt é implementado em um computador, existe quase sempre algum erro de arredondamento, levando a uma perda da ortogonalidade dos vetores $\mathbf{q}_i$. Para evitar essa perda da ortogonalidade, geralmente são feitas algumas modificações. Em vez de normalizarmos os vetores $\mathbf{v}_i$ só no fim para obtermos os vetores $\mathbf{q}_i$, aqueles são normalizados logo depois de calculados, e, assim que cada $\mathbf{q}_i$ é calculado, os vetores restantes $\mathbf{x}_j$ são modificados para ficarem ortogonais a $\mathbf{q}_i$. Esse procedimento é conhecido por ***Processo de Gram-Schmidt Modificado***. Entretanto, na prática, a versão da fatoração QR é usada para calcular bases ortonormais.

A Fatoração QR

Se A é uma matriz $m \times n$ com colunas linearmente independentes (exigindo $m \geq n$), aplicar o Processo de Gram-Schmidt a essas colunas implica uma fatoração muito útil da matriz A em um produto de uma matriz ortogonal Q por uma matriz triangular superior R. Essa é a ***fatoração QR***, e tem aplicações na aproximação numérica de autovalores — o que investigaremos no fim desta seção — e no problema da aproximação pelo método dos mínimos quadrados, que discutiremos no Capítulo 7.

Para ver como surge a fatoração QR, considere $\mathbf{a}_1, \ldots, \mathbf{a}_n$ como as colunas (linearmente independentes) de A, e $\mathbf{q}_1, \ldots, \mathbf{q}_n$ como os vetores ortonormais obtidos pela aplicação do Processo de Gram-Schmidt à matriz A com normalizações. Pelo Teorema 1, sabemos que, para cada $i = 1, \ldots, n$,

$$W_i = \text{ger}(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \text{ger}(\mathbf{q}_1, \ldots, \mathbf{q}_i)$$

Assim, existem escalares $r_{1i}, r_{2i}, \ldots, r_{ii}$ tais que

$$\mathbf{a}_i = r_{1i}\mathbf{q}_1 + r_{2i}\mathbf{q}_2 + \cdots + r_{ii}\mathbf{q}_i \quad \text{para} \quad i = 1, \ldots, n$$

Isto é,

$$\begin{aligned}
\mathbf{a}_1 &= r_{11}\mathbf{q}_1 \\
\mathbf{a}_2 &= r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2 \\
&\vdots \\
\mathbf{a}_n &= r_{1n}\mathbf{q}_1 + r_{2n}\mathbf{q}_2 + \cdots + r_{nn}\mathbf{q}_n
\end{aligned}$$

o que pode ser escrito na forma de matriz como

$$A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n] = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \cdots \quad \mathbf{q}_n]\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix} = QR$$

É claro que a matriz Q tem colunas ortogonais. Também é verdade que os elementos na diagonal de R são não nulos. Para ver isso, observe que, se $r_{ii} = 0$, então $\mathbf{a}_i$ é uma combinação linear de $\mathbf{q}_1, \ldots, \mathbf{q}_{i-1}$, e, portanto, está em W_{i-1}. Mas então $\mathbf{a}_i$ seria uma combinação linear de $\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}$, o que seria impossível, já que $\mathbf{a}_1, \ldots, \mathbf{a}_i$ são linearmente independentes. Concluímos que $r_{ii} \neq 0$ para $i = 1, \ldots, n$. Como R é triangular superior, segue que ela tem que ser invertível. (Veja o Exercício 21.)

Provamos o seguinte teorema:

◆ TEOREMA 2 A Fatoração QR

Seja A uma matriz $m \times n$ com colunas linearmente independentes. Então, A pode ser escrita como $A = QR$, onde Q é uma matriz $m \times n$ com colunas ortonormais e R é uma matriz triangular superior invertível.

Observações:

- Também podemos assumir que os elementos da diagonal de R são *positivos*. Se $r_{ii} < 0$, simplesmente substituímos $\mathbf{q}_i$ por $-\mathbf{q}_i$ e r_{ii} por $-r_{ii}$.
- A exigência de que A tenha colunas linearmente independentes é uma condição necessária. Para provar isso, suponha que A seja uma matriz $m \times n$ com uma fatoração QR, como no Teorema 2. Então, como R é invertível, temos que $Q = AR^{-1}$. Logo, posto(Q) = posto(A), pelo Exercício 55 da Seção 3.5. Mas posto$(Q) = n$, já que suas colunas são ortonormais e, portanto, linearmente independentes. Logo, posto$(A) = n$ também, e, conseqüentemente, as colunas de A são linearmente independentes, de acordo com o Teorema Fundamental.
- A fatoração QR pode ser estendida para matrizes arbitrárias de um modo um pouco diferente. Se A é uma matriz $m \times n$, é possível achar uma seqüência de matrizes ortogonais $Q_1, \ldots, Q_{m-1}$ tais que $Q_{m-1} \cdots Q_2Q_1A$ sejam uma matriz $m \times n$ triangular superior R. Então, $A = QR$, onde $Q = (Q_{m-1} \cdots Q_2Q_1)^{-1}$ é uma matriz ortogonal. Examinaremos essa abordagem em Investigação: A Fatoração QR Modificada.

EXEMPLO 4 Encontre a fatoração QR de

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

SOLUÇÃO: As colunas de A são exatamente os vetores do Exemplo 2. Uma base ortonormal para col(A), produzida pelo Processo de Gram-Schmidt, é

$$\mathbf{q}_1 = \begin{bmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} 3\sqrt{5}/10 \\ 3\sqrt{5}/10 \\ \sqrt{5}/10 \\ \sqrt{5}/10 \end{bmatrix}, \quad \mathbf{q}_3 = \begin{bmatrix} -\sqrt{6}/6 \\ 0 \\ \sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix}$$

de modo que

$$Q = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3] = \begin{bmatrix} 1/2 & 3\sqrt{5}/10 & -\sqrt{6}/6 \\ -1/2 & 3\sqrt{5}/10 & 0 \\ -1/2 & \sqrt{5}/10 & \sqrt{6}/6 \\ 1/2 & \sqrt{5}/10 & \sqrt{6}/3 \end{bmatrix}.$$

Do Teorema 2, $A = QR$ para alguma matriz triangular superior R. Para encontrar R, usamos o fato de Q ter colunas ortonormais e, conseqüentemente, $Q^TQ = I$. Logo,

$$Q^TA = Q^TQR = IR = R$$

Calculamos

$$R = Q^TA = \begin{bmatrix} 1/2 & -1/2 & -1/2 & 1/2 \\ 3\sqrt{5}/10 & 3\sqrt{5}/10 & \sqrt{5}/10 & \sqrt{5}/10 \\ -\sqrt{6}/6 & 0 & \sqrt{6}/6 & \sqrt{6}/3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \\ -1 & 1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1/2 \\ 0 & \sqrt{5} & 3\sqrt{5}/2 \\ 0 & 0 & \sqrt{6}/2 \end{bmatrix}$$

◆ EXERCÍCIOS 5.4 ◆

Nos Exercícios de 1 a 4, são dados vetores que formam uma base para $\mathbb{R}^2$ ou $\mathbb{R}^3$. Aplique o Processo de Gram-Schmidt para obter uma base ortogonal. Normalize então a base encontrada para obter uma base ortonormal.

1. $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

2. $\mathbf{x}_1 = \begin{bmatrix} 3 \\ -3 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

3. $\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 0 \\ 3 \\ 3 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix}$

4. $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

Nos Exercícios 5 e 6, são dados vetores que formam uma base para um subespaço W de $\mathbb{R}^3$ ou $\mathbb{R}^4$. Aplique o Processo de Gram-Schmidt para obter uma base ortogonal para W.

5. $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$

6. $\mathbf{x}_1 = \begin{bmatrix} 2 \\ -1 \\ 1 \\ 2 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 3 \\ -1 \\ 0 \\ 4 \end{bmatrix}$

Nos Exercícios 7 e 8, encontre a decomposição ortogonal de $\mathbf{v}$ *em relação ao subespaço W.*

7. $\mathbf{v} = \begin{bmatrix} 4 \\ -4 \\ 3 \end{bmatrix}$, W como no Exercício 5.

8. $\mathbf{v} = \begin{bmatrix} 1 \\ 4 \\ 0 \\ 2 \end{bmatrix}$, W como no Exercício 6.

Use o Processo de Gram-Schmidt para achar uma base ortogonal para os espaços coluna das matrizes dos Execícios 9 e 10.

9. $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ **10.** $\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ -1 & 1 & 0 \\ 1 & 5 & 1 \end{bmatrix}$

11. Encontre uma base ortogonal de $\mathbb{R}^3$ que contenha o vetor

$$\begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix}$$

12. Encontre uma base ortogonal de $\mathbb{R}^4$ que contenha os vetores

$$\begin{bmatrix} 2 \\ 1 \\ 0 \\ -1 \end{bmatrix} \text{ e } \begin{bmatrix} 1 \\ 0 \\ 3 \\ 2 \end{bmatrix}$$

Nos Exercícios 13 e 14, encontre a fatoração QR da matriz no exercício dado.

13. Exercício 9 **14.** Exercício 10

Nos Exercícios 15 e 16, as colunas de Q foram obtidas pela aplicação do Processo de Gram-Schmidt às colunas da matriz A. Encontre a matriz triangular superior R de maneira que A = QR.

15. $A = \begin{bmatrix} 2 & 8 & 2 \\ 1 & 7 & -1 \\ -2 & -2 & 1 \end{bmatrix}, Q = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix}$

16. $A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \\ -1 & -1 \\ 0 & 1 \end{bmatrix}, Q = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{3} \\ 2/\sqrt{6} & 0 \\ -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 1/\sqrt{3} \end{bmatrix}$

17. Se A é uma matriz ortogonal, encontre a fatoração QR de A.

18. Prove que A é invertível se, e somente se, $A = QR$, onde Q é ortogonal e R é triangular superior com elementos não nulos na diagonal.

Nos Exercícios 19 e 20, use o método sugerido pelo Exercício 18 para calcular A^{-1} para a matriz A no exercício dado.

19. Exercício 9 **20.** Exercício 15

21. Seja A uma matriz $m \times n$ com colunas linearmente independentes. Dê uma demonstração alternativa do fato de que a matriz triangular superior R de uma fatoração QR tem que ser invertível, usando a propriedade (c) do Teorema Fundamental.

22. Seja A uma matriz $m \times n$ com colunas linearmente independentes, e seja $A = QR$ uma fatoração QR de A. Mostre que A e Q têm o mesmo espaço coluna.