Álgebra Linear – Semana 12

28 de Junho de 2017

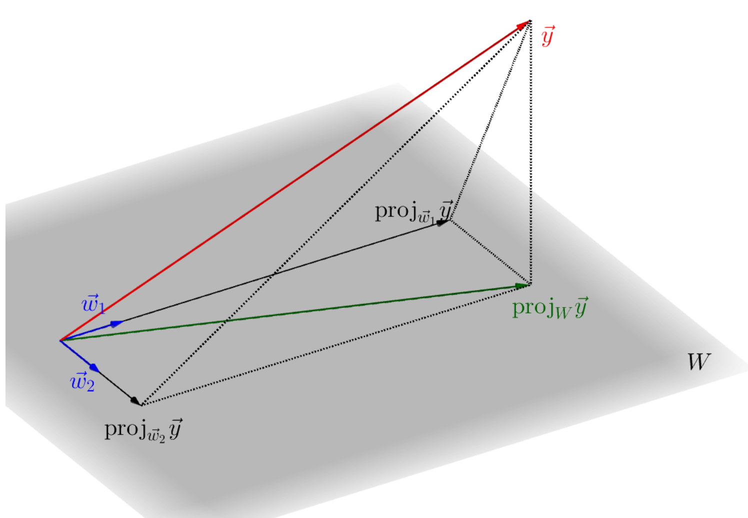
Conteúdo

1 Projeção ortogonais sobre subespaços

Na última semana, estudamos projeções ortogonais e vimos que a projeção ortogonal de $\vec{y}$ na direção de um vetor $\vec{w}$ pode ser calculada por

$$\mathrm{proj}_{\vec{w}}\, \vec{y} = \frac{\vec{y}\cdot\vec{w}}{\vec{w}\cdot\vec{w}}\ \vec{w}$$

Gostaríamos de endereçar agora a seguinte questão: como obter a projeção de um vetor $\vec{y} \in \mathbb{R}^3$ sobre um plano W?



Vamos trabalhar geométricamente: suponhamos que o plano

$$W = \operatorname{Span}\{\vec{w}_1, \vec{w}_2\}$$

e que os vetores $\vec{w}_1$ e $\vec{w}_2$ são ortogonais. Em outras palavras, $\{\vec{w}_1, \vec{w}_2\}$ é uma base ortogonal de W. Como veremos abaixo, o Processo de Gram-Schmidt fornece um algoritmo para obter uma base ortogonal de um espaço vetorial a partir de uma outra base qualquer e, logo, sempre é possível considerar uma base ortogonal.

A **projeção ortogonal sobre o subespaço** W, denotada por $\operatorname{proj}_W \vec{y}$, é definida como

$$\operatorname{proj}_W \vec{y} = \operatorname{proj}_{\vec{w}_1} \vec{y} + \operatorname{proj}_{\vec{w}_2} \vec{y} = \frac{\vec{y} \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\, \vec{w}_1 + \frac{\vec{y} \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\, \vec{w}_2.$$

Notamos que este vetor pertence a W, já que é uma combinação linear dos elementos de uma base de W. Mas mais fundamental é o fato de a projeção ser feita em uma direção ortogonal ao plano: devemos verificar que

$$\vec{y} - \operatorname{proj}_W \vec{y} \text{ é ortogonal a } W.$$

Qualquer vetor de W pode ser escrito como

$$\vec{w} = \frac{\vec{w} \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\, \vec{w}_1 + \frac{\vec{w} \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\, \vec{w}_2.$$

Exercício 1. Utilize as propriedades do produto escalar e a ortogonalidade da base $\{\vec{w}_1, \vec{w}_2\}$ para verificar que

$$\vec{w} \cdot \big(\vec{y} - \operatorname{proj}_W \vec{y}\big) = 0.$$

Exemplo 1. Vamos calcular a projeção do vetor

$$\vec{y} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

sobre o plano gerado pelos vetores

$$\vec{w}_1 = \begin{bmatrix} 0 \\ 3 \\ -7 \end{bmatrix} \quad \text{e} \quad \vec{w}_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}.$$

Comlculamos

$$\begin{aligned} \operatorname{proj}_W \vec{y} &= \frac{\vec{y} \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\, \vec{w}_1 + \frac{\vec{y} \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\, \vec{w}_2 = \frac{3-7}{9+49}\, \vec{w}_1 + \frac{1+2}{1+4}\, \vec{w}_2 \\ &= -\frac{2}{29} \begin{bmatrix} 0 \\ 3 \\ -7 \end{bmatrix} + \frac{3}{5} \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 3/5 \\ -6/29 \\ 244/145 \end{bmatrix}. \end{aligned}$$

Desenvolvemos acima o conceito de projeção ortogonal sobre um subespaço do espaço $\mathbb{R}^3$ de dimensão três apenas pela conveniência da visualização geométrica. Na verdade, o conceito pode ser definido em qualquer dimensão.

Seja $W \subset \mathbb{R}^n$ um subespaço vetorial de dimensão k e $\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k\}$ uma base ortogonal de W. Dado um vetor $\vec{y} \in \mathbb{R}^n$ qualquer, definimos a **projeção ortogonal de $\vec{y}$ sobre** W como

$$\operatorname{proj}_W \vec{y} = \frac{\vec{y} \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\, \vec{w}_1 + \frac{\vec{y} \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\, \vec{w}_2 + \cdots + \frac{\vec{y} \cdot \vec{w}_k}{\vec{w}_k \cdot \vec{w}_k}\, \vec{w}_k$$

Pode-se verificar, como no Exercício acima, que

$$\vec{w} \cdot \big(\vec{y} - \operatorname{proj}_W \vec{y}\big) = 0, \text{ para qualquer } \vec{w} \in W,$$

isto é, a projeção $\operatorname{proj}_W \vec{y}$ pertence a W e $\vec{y} - \operatorname{proj}_W \vec{y}$ é ortogonal a todo elemento de W, assim como nossa intuição esperaria.

Exemplo 2. Calcular a projeção do vetor

$$\vec{y} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ -3 \\ 2 \end{bmatrix}$$

sobre o espaço tridimensional gerado pelos vetores

$$\vec{w}_1 = \begin{bmatrix} 0 \\ 4 \\ 1 \\ 6 \\ 0 \end{bmatrix}, \quad \vec{w}_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \\ -1 \end{bmatrix} \quad \text{e} \quad \vec{w}_3 = \begin{bmatrix} 0 \\ 0 \\ 2 \\ 4 \\ 1 \end{bmatrix}.$$

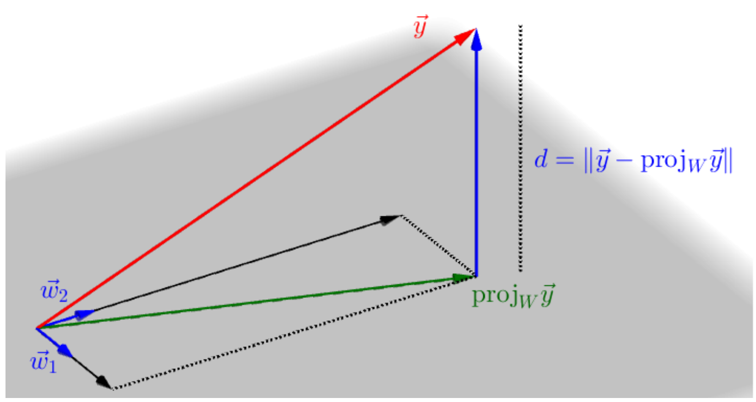
Devemos calcular

$$\begin{aligned} \operatorname{proj}_W \vec{y} &= \frac{\vec{y} \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\, \vec{w}_1 + \frac{\vec{y} \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\, \vec{w}_2 + \frac{\vec{y} \cdot \vec{w}_3}{\vec{w}_3 \cdot \vec{w}_3}\, \vec{w}_3 \\ &= \frac{-1-18}{16+1+36}\, \vec{w}_1 + \frac{1-2-2}{1+4+1}\, \vec{w}_2 + \frac{-2-12+2}{4+16+1}\, \vec{w}_3 \\ &= -\frac{19}{53} \begin{bmatrix} 0 \\ 4 \\ 1 \\ 6 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \\ -1 \end{bmatrix} - \frac{12}{21} \begin{bmatrix} 0 \\ 0 \\ 2 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/2 \\ -76/53 \\ -2784/1113 \\ -4938/1113 \\ -1/12 \end{bmatrix}. \end{aligned}$$

Confira estas contas com uma calculadora (os coeficientes dos vetores nem sempre são bonitinhos!)

2 Distância a subespaços e a melhor aproximação

A projeção ortogonal pode ser utilizada para calcular distâncias entre o pontos e subespaços.



Dado um ponto $P = (x_1, x_2, \ldots, x_n)$, formamos o vetor

$$\vec{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

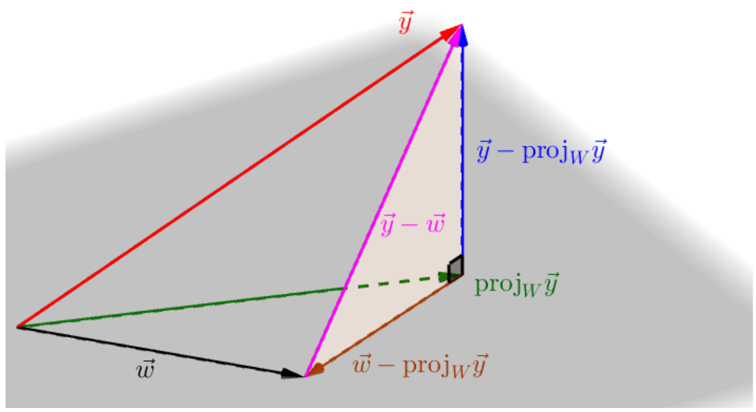
cujas componentes são as coordenadas de P. Como vimos na seção anterior, o vetor $\vec{y} - \mathrm{proj}_W \vec{y}$ é ortogonal ao subespaço W. Sendo a distância de P até W denotada por $\mathrm{dist}(P, W)$ e definida como a menor distância entre P e elementos de W, temos que

$$\mathrm{dist}(P, W) = \|\vec{y} - \mathrm{proj}_W \vec{y}\|.$$

Justificativa desta afirmação. Dado qualquer outro vetor $\vec{w} \in W$, o Teorema de Pitágoras implica que (ver figura)

$$\|\vec{y} - \vec{w}\|^2 = \|\vec{w} - \mathrm{proj}_W \vec{y}\|^2 + \|\vec{y} - \mathrm{proj}_W \vec{y}\|^2 > \|\vec{y} - \mathrm{proj}_W \vec{y}\|^2.$$

Em outras palavras, $\|\vec{y} - \mathrm{proj}_W \vec{y}\|$ é de fato a menor distância entre P e elementos de W. $\square$



Exemplo 3. Calcular a distância do ponto $P = (1, 2, -1, 0, 7)$ até o plano $W = \mathrm{Span}\{\vec{u}, \vec{v}\}$ (contido em $\mathbb{R}^5$) gerado pelos vetores

$$\vec{u} = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \\ 5 \end{bmatrix} \quad \text{e} \quad \vec{v} = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \\ 3 \end{bmatrix}.$$

Para isto, consideramos o vetor

$$\vec{y} = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \\ 7 \end{bmatrix}$$

e vamos calcular $\|\vec{y} - \mathrm{proj}_W \vec{y}\|$:

$$\vec{y} - \mathrm{proj}_W \vec{y} = \vec{y} - \frac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} \vec{u} - \frac{\vec{y} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} \vec{v} = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \\ 7 \end{bmatrix} - \frac{34}{30} \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \\ 5 \end{bmatrix} - \frac{18}{15} \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \\ 7 \end{bmatrix} - \frac{17}{15} \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \\ 5 \end{bmatrix} - \frac{18}{15} \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} -4/3 \\ 16/5 \\ -17/3 \\ 0 \\ -34/15 \end{bmatrix}$$

Desta forma, a distância de P a W é

$$\begin{aligned}\operatorname{dist}(P, W) = \|\vec{y} - \operatorname{proj}_W \vec{y}\| &= \sqrt{\frac{16}{9} + \frac{256}{25} + \frac{289}{9} + \frac{1156}{225}} = \sqrt{\frac{16 \cdot 25 + 256 \cdot 9 + 289 \cdot 25 + 1156}{225}} \\ &= \frac{\sqrt{11085}}{15} \simeq 7,02.\end{aligned}$$

3 O Processo de ortogonalização de Gram–Schmidt

O **Processo de Gram–Schmidt** é um algoritmo para obter uma base ortogonal (ou ortonormal) a partir de uma base qualquer. De maneira mais geral, o método permite transformar um conjunto de vetores linearmente independentes em um conjunto ortogonal que gera o mesmo espaço vetorial.

Vamos começar com um exemplo.

Exemplo 4. Consideramos os vetores

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{e} \quad \vec{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

O espaço gerado por $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ é todo o espaço $\mathbb{R}^3$, já que temos três vetores linearmente independentes em um espaço de dimensão três.

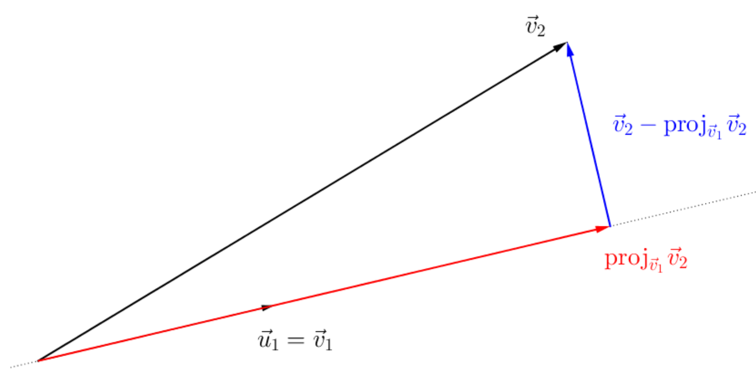
O processo consiste em fixar qualquer um dos vetores como o primeiro dos vetores do conjunto que virá a ser ortogonal. Por exemplo, fixamos

$$\vec{u}_1 = \vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Em seguida, já vimos que, ao fazer a projeção de $\vec{v}_2$ sobre $\vec{v}_1$, o vetor

$$\vec{v}_2 - \operatorname{proj}_{\vec{v}_1} \vec{v}_2$$

é ortogonal a $\vec{v}_1$, como ilustra a figura.



Assim sendo, definimos o segundo vetor do nosso conjunto que será ortogonal como o vetor

$$\vec{u}_2 = \vec{v}_2 - \operatorname{proj}_{\vec{u}_1} \vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1 \\ -1/2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}.$$

Temos que $\vec{u}_1$ e $\vec{u}_2$ são ortogonais e que estão no mesmo plano, de modo que também temos

$$\operatorname{Span}\{\vec{u}_1, \vec{u}_2\} = \operatorname{Span}\{\vec{v}_1, \vec{v}_2\}.$$

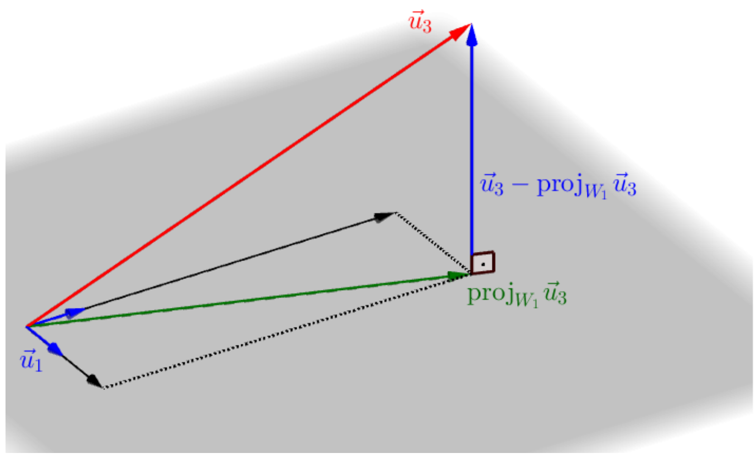
Vamos escrever momentariamente $W = \mathrm{Span}\{\vec{u}_1, \vec{u}_2\}$.

No próximo passo do processo, o terceiro vetor pode ser obtido como

$$\vec{u}_3 = \vec{v}_3 - \mathrm{proj}_W \vec{v}_3,$$

pois, desta forma, $\vec{u}_3$ é ortogonal a todos os vetores de W; em particular, é ortogonal a ambos $\vec{u}_1$ e $\vec{u}_2$. Além disso, como $\vec{u}_1$ e $\vec{u}_2$ já são vetores ortogonais, podemos calcular a projeção sobre W como de costume:

$$\vec{u}_3 = \vec{v}_3 - \mathrm{proj}_W \vec{v}_3 = \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\, \vec{u}_1 - \frac{\vec{v}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2}\, \vec{u}_2.$$



Calculando, temos (observe como dois fatores $1/2$ cancelam no último termo)

$$\vec{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{2-1}{1+4+1} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{6} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -2/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \frac{2}{3} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

Concluimos assim que o conjunto

$$\{\vec{u}_1, \vec{u}_2, \vec{u}_3\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1/2 \\ 1 \\ -1/2 \end{bmatrix}, \begin{bmatrix} -2/3 \\ 2/3 \\ 2/3 \end{bmatrix} \right\}$$

é ortogonal e gera o mesmo espaço que $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$.

Na realidade, se tivéssemos considerado múltiplos dos vetores acima, não perderíamos a propriedade de ortogonalidade, e temos

$$\mathrm{Span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\} = \mathrm{Span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

"Colocar em evidência" os fatores comuns desta forma e "desconsiderá-los" pode facilitar as contas em alguns casos.

Observamos também que, fazendo a normalização dos vetores, podemos obter uma base ortonormal:

$$\begin{cases} \|\vec{u}\| = \sqrt{1+0+1} = \sqrt{2} \\ \|\vec{v}\| = \sqrt{1+4+1} = \sqrt{6} \\ \|\vec{w}\| = \sqrt{1+1+1} = \sqrt{3} \end{cases}$$

e assim,

$$\frac{\vec{u}_1}{\|\vec{u}_1\|} = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\0\\1\end{bmatrix} = \begin{bmatrix}1/\sqrt{2}\\0\\1/\sqrt{2}\end{bmatrix}, \quad \frac{\vec{u}_2}{\|\vec{u}_2\|} = \frac{1}{\sqrt{6}}\begin{bmatrix}1\\2\\-1\end{bmatrix} = \begin{bmatrix}1/\sqrt{6}\\2/\sqrt{6}\\-1/\sqrt{6}\end{bmatrix} \text{ e } \frac{\vec{u}_3}{\|\vec{u}_3\|} = \frac{1}{\sqrt{3}}\begin{bmatrix}-1\\1\\1\end{bmatrix} = \begin{bmatrix}-1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\end{bmatrix}$$

forma um conjunto ortonormal que gera o mesmo espaço vetorial que $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. $\triangleleft$

Em seguida, mostramos como funciona o processo de maneira geral (um bom entendimento do exemplo acima deve deixar claro o porquê deste processo funcionar): sejam k vetores linearmente independentes $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k$ pertencentes ao espaço n-dimensional $\mathbb{R}^n$. Vamos denotar

$$W = \operatorname{Span}\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k\}$$

e construir uma base ortogonal (e em seguida ortonormal) para W. Começamos com

$$\boxed{\vec{u}_1 = \vec{v}_1}$$

e definimos

$$\boxed{\vec{u}_2 = \vec{v}_2 - \operatorname{proj}_{\vec{u}_1} \vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\, \vec{u}_1.}$$

Temos que $\vec{u}_1$ e $\vec{u}_2$ são ortogonais e que estão no mesmo plano, de modo que também temos

$$\operatorname{Span}\{\vec{u}_1, \vec{u}_2\} = \operatorname{Span}\{\vec{v}_1, \vec{v}_2\} = W_2 \;\leftrightsquigarrow\; \text{notação.}$$

Em seguida, definimos

$$\boxed{\vec{u}_3 = \vec{v}_3 - \operatorname{proj}_{W_2} \vec{v}_3 = \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\, \vec{u}_1 - \frac{\vec{v}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2}\, \vec{u}_2.}$$

Temos que $\vec{u}_1, \vec{u}_2, \vec{u}_3$ ortogonais e

$$\operatorname{Span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\} = \operatorname{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = W_3 \;\leftrightsquigarrow\; \text{notação.}$$

Em seguida

$$\boxed{\vec{u}_4 = \vec{v}_4 - \operatorname{proj}_{W_3} \vec{v}_4 = \vec{v}_4 - \frac{\vec{v}_4 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\, \vec{u}_1 - \frac{\vec{v}_4 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2}\, \vec{u}_2 - \frac{\vec{v}_4 \cdot \vec{u}_3}{\vec{u}_3 \cdot \vec{u}_3}\, \vec{u}_3.}$$

Temos que $\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4$ ortogonais e

$$\operatorname{Span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4\} = \operatorname{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{u}_4\} = W_4 \;\leftrightsquigarrow\; \text{notação.}$$

E assim seguimos fazendo as contas até chegar ao último vetor: sendo

$$\operatorname{Span}\{\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_{k-1}\} = \operatorname{Span}\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_{k-1}\} = W_{k-1} \;\leftrightsquigarrow\; \text{notação,}$$

escrevemos

$$\boxed{\vec{u}_k = \vec{v}_k - \operatorname{proj}_{W_{k-1}} \vec{v}_k = \vec{v}_k - \frac{\vec{v}_k \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\, \vec{u}_1 - \frac{\vec{v}_k \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2}\, \vec{u}_2 - \cdots - \frac{\vec{v}_k \cdot \vec{u}_{k-1}}{\vec{u}_{k-1} \cdot \vec{u}_{k-1}}\, \vec{u}_{k-1}.}$$

Temos que $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ é um conjunto ortogonal e

$$\operatorname{Span}\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\} = \operatorname{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{u}_k\} = W \;\leftrightsquigarrow\; \text{que é o espaço inicial,}$$

como queríamos.

Este é a parte mais "pesada" do método. Por exemplo, se agora quiséssemos uma base ortonormal para W, basta considerar o conjunto

$$\left\{ \frac{\vec{u}_1}{\|\vec{u}_1\|}, \frac{\vec{u}_2}{\|\vec{u}_2\|}, \ldots, \frac{\vec{u}_k}{\|\vec{u}_k\|} \right\},$$

que continua sendo ortogonal, mas agora formado por vetores unitários.

Exemplo 5 (Retirado de Wikipedia.org). Vamos obter uma base ortonormal para o espaço coluna $\operatorname{Col} A$ da matriz

$$A = \begin{bmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{bmatrix}.$$

Lembrando que $\operatorname{Col} A$ é o espaço gerado pelas colunas de A, vamos aplicar o processo de Gram–Schmidt para os vetores

$$\vec{v}_1 = \begin{bmatrix} 12 \\ 6 \\ -4 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -51 \\ 167 \\ 24 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 4 \\ -68 \\ -41 \end{bmatrix}.$$

O Processo começa como:

$$\vec{u}_1 = \vec{v}_1 = \begin{bmatrix} 12 \\ 6 \\ -4 \end{bmatrix},$$

$$\vec{u}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 = \begin{bmatrix} -51 \\ 167 \\ 24 \end{bmatrix} - \frac{-612 + 1002 - 96}{144 + 36 + 16} \begin{bmatrix} 12 \\ 6 \\ -4 \end{bmatrix} = \begin{bmatrix} -51 \\ 167 \\ 24 \end{bmatrix} - \frac{296}{196} \begin{bmatrix} 12 \\ 6 \\ -4 \end{bmatrix} = \begin{bmatrix} -69 \\ 158 \\ 30 \end{bmatrix}$$

Em seguida (conferir estas contas com o auxílio de uma calculadora)

$$\vec{u}_3 = \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 - \frac{\vec{v}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2 = \begin{bmatrix} 4 \\ -68 \\ -41 \end{bmatrix} - \frac{48 - 408 + 164}{196} \begin{bmatrix} 12 \\ 6 \\ -4 \end{bmatrix} - \frac{-276 - 10744 - 1230}{4761 + 24694 + 900} \begin{bmatrix} -69 \\ 158 \\ 30 \end{bmatrix}$$
$$= \begin{bmatrix} 4 \\ -68 \\ -41 \end{bmatrix} + \frac{19\cancel{6}}{\cancel{1}96} \begin{bmatrix} 12 \\ 6 \\ -4 \end{bmatrix} + \frac{12250}{30355} \begin{bmatrix} -69 \\ 158 \\ 30 \end{bmatrix} = \begin{bmatrix} -58/5 \\ 6/5 \\ -33 \end{bmatrix}.$$

Para obter uma base ortonormal para o espaço coluna, vamos ainda considerar

$$\frac{\vec{u}_1}{\|\vec{u}_1\|}, \frac{\vec{u}_2}{\|\vec{u}_2\|}, \frac{\vec{u}_3}{\|\vec{u}_3\|}.$$

Calculamos

$$\begin{cases} \|\vec{u}_1\| = \sqrt{196} = 14 \\ \|\vec{u}_2\| = \sqrt{30625} = 175 \\ \|\vec{u}_3\| = \dfrac{\sqrt{30625}}{5} = 35 \end{cases}$$

Portanto, uma base ortonormal para $\operatorname{Col} A$ é

$$\vec{n}_1 = \frac{\vec{u}_1}{\|\vec{u}_1\|} = \begin{bmatrix} 6/7 \\ 3/7 \\ -2/7 \end{bmatrix}, \quad \vec{n}_2 = \frac{\vec{u}_2}{\|\vec{u}_2\|} = \begin{bmatrix} -69/175 \\ 158/175 \\ 6/35 \end{bmatrix} \quad \text{e} \quad \vec{n}_3 = \frac{\vec{u}_3}{\|\vec{u}_3\|} = \begin{bmatrix} -58/175 \\ 6/175 \\ -33/35 \end{bmatrix}. \triangleleft$$

4 Fatoração QR e matrizes ortogonais

O método de Gram-Schmidt nos permite fatorar uma matriz A cujas colunas são LI em um produto

$$A = QR,$$

onde Q é uma matriz cujas colunas formam um conjunto ortonormal e R é uma matriz triangular superior. Tal fatoração é muito importante do ponto de vista computacional, como veremos em breve, quando usarmos a fatoração QR como uma alternativa para resolução de sistemas lineares por mínimos quadrados. Também é a base de um eficiente método numérico para encontrar autovalores de uma matriz quadrada.

Matrizes cujas colunas formam um conjunto ortonormal são muito úteis e por isso recebem um nome especial:são chamadas de matrizes ortogonais. Aqui cabe uma observação relevante: este é um exemplo de uma situação onde a terminologia clássica está longe de ser satisfatória: matrizes cujas colunas formam um conjunto *ortonormal* são chamadas de matrizes *ortogonais*.

Faremos agora uma exposição da fatoração QR usando uma matriz inicial A que tem 3 colunas, observando que a construção abaixo é facilmente generalizada para uma matriz de qualquer dimensão (desde que tenha colunas LI).

Dada uma matriz $A = [\vec{v}_1\vec{v}_2\vec{v}_3]$ cujas colunas são LI e são denotadas por $\vec{v}_1$, $\vec{v}_2$, e $\vec{v}_3$, o processo de Gram-Schmidt começa com a construção de uma base ortogonal formada pelos vetores $\vec{u}_1$, $\vec{u}_2$, e $\vec{u}_3$, para o espaço-coluna da matriz A. Para isso, definimos

$$\begin{cases} \vec{u}_1 = \vec{v}_1 \\ \vec{u}_2 = \vec{v}_2 - \operatorname{proj}_{\vec{u}_1} \vec{v}_2 \\ \vec{u}_3 = \vec{v}_3 - \operatorname{proj}_{\vec{u}_1,\vec{u}_2} \vec{v}_3 \end{cases}$$

Portanto existem coeficientes b_1, c_1 e c_2 tais que

$$\begin{cases} \vec{u}_1 = \vec{v}_1 \\ \vec{u}_2 = \vec{v}_2 - b_1\vec{u}_1 \\ \vec{u}_3 = \vec{v}_3 - c_1\vec{u}_1 - c_2\vec{u}_2 \end{cases}$$

Em um segundo momento, cada um dos três vetores $\vec{u}_1$, $\vec{u}_2$, e $\vec{u}_3$ é normalizado, ou seja, dividimos cada um deles pelo seu comprimento. Assim, $\vec{u}_1$, $\vec{u}_2$, e $\vec{u}_3$ formam uma base ortonormal para o espaço coluna de A, e existem coeficientes α_1, β_1, β_2, γ_1, γ_2 e γ_3 tais que

$$\begin{cases} \vec{u}_1 = \alpha_1\vec{v}_1 \\ \vec{u}_2 = \beta_2\vec{v}_2 - \beta_1\vec{u}_1 \\ \vec{u}_3 = \gamma_3\vec{v}_3 - \gamma_1\vec{u}_1 - \gamma_2\vec{u}_2 \end{cases}$$

Se colocarmos as colunas da matriz A em função dos vetores da base ortonormal que acaba de ser obtida, temos

$$\begin{cases} \vec{v}_1 = \alpha_1^{-1}\vec{u}_1 \\ \vec{v}_2 = \beta_2^{-1}\vec{u}_2 + \beta_2^{-1}\beta_1\vec{u}_1 \\ \vec{v}_3 = \gamma_3^{-1}\vec{u}_3 + \gamma_3^{-1}\gamma_1\vec{u}_1 + \gamma_3^{-1}\gamma_2\vec{u}_2 \end{cases}$$

ou trocando a ordem das parcelas no lado direito,

$$\begin{cases} \vec{v}_1 = \alpha_1^{-1}\vec{u}_1 \\ \vec{v}_2 = \beta_2^{-1}\beta_1\vec{u}_1 + \beta_2^{-1}\vec{u}_2 \\ \vec{v}_3 = \gamma_3^{-1}\gamma_1\vec{u}_1 + \gamma_3^{-1}\gamma_2\vec{u}_2 + \gamma_3^{-1}\vec{u}_3 \end{cases}$$

Assim, ao definirmos a matriz $Q = [\vec{u}_1\vec{u}_2\vec{u}_3]$, obtemos

$$A = QR$$

onde

$$R = \begin{bmatrix} \alpha_1^{-1} & \beta_2^{-1}\beta_1 & \gamma_3^{-1}\gamma_1 \\ 0 & \beta_2^{-1} & \gamma_3^{-1}\gamma_2 \\ 0 & 0 & \gamma_3^{-1} \end{bmatrix}$$

visto que cada coluna do produto QR, é definido como a combinação linear das colunas da matriz Q com pesos dados pela respectiva coluna da matriz R.

Note ainda que R é invertível, visto que sua diagonal principal contém apenas termos não-nulos.

A construção acima é facilmente generalizada para matrizes quaisquer e nos permite enunciar o seguinte fato (lembre que uma matriz ortogonal é uma matriz cujas colunas formam um conjunto ortonormal):

Teorema 6. *Qualquer matriz A cujas colunas são linearmente independentes pode ser fatorada na forma $A = QR$, onde Q é uma matriz ortogonal e R é uma matriz triangular superior invertível. Ainda, o espaço coluna das matrizes A e Q é o mesmo, ou seja, as colunas de Q formam uma base ortonormal para o subespaço gerado pelas colunas de A.*

Note que a fatoração QR está para o processo de Gram-Schmidt assim como a fatoração LU está para o método de linha-redução a forma escada!

Bem, vimos que a matriz Q tem como colunas a base ortonormal gerada pelo processo de Gram-Schmidt. E como podemos calcular a matriz R no caso geral ?

Usando o seguinte fato surpreendente:

Teorema 7. *uma matriz ortogonal é sempre invertível e sua inversa é dada pela sua transposta, ou seja $Q^{-1} = Q^T$.*

A prova deste teorema é simples:

$$Q^T Q = \begin{bmatrix} — & \vec{u}_1 & — \\ — & \vec{u}_2 & — \\ — & \vec{u}_3 & — \end{bmatrix} \begin{bmatrix} | & | & | \\ \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} \vec{u}_1 \cdot \vec{u}_1 & \vec{u}_1 \cdot \vec{u}_2 & \vec{u}_1 \cdot \vec{u}_2 \\ \vec{u}_2 \cdot \vec{u}_1 & \vec{u}_2 \cdot \vec{u}_2 & \vec{u}_2 \cdot \vec{u}_3 \\ \vec{u}_3 \cdot \vec{u}_1 & \vec{u}_3 \cdot \vec{u}_2 & \vec{u}_3 \cdot \vec{u}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I.$$

Assim, se $A = QR$, basta multiplicar ambos os lados desta equação por Q^T pela esquerda para obtermos:

$$Q^T A = Q^T QR \Longrightarrow R = Q^T A.$$

No último exemplo da seção anterior, obtemos através do método de Gram-Schmidt uma base ortonormal para $\operatorname{Col} A$ onde

$$A = \begin{bmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{bmatrix}.$$

Vamos definir a matriz Q como a matriz formada pelos vetores desta base ortonormal:

$$Q = \begin{bmatrix} 6/7 & -69/175 & -58/175 \\ 3/7 & 158/175 & 6/175 \\ -2/7 & 6/35 & -33/35 \end{bmatrix}$$

Vimos que $R = Q^T A$ e portanto

$$R = Q^T A = \begin{bmatrix} 6/7 & 3/7 & -2/7 \\ -69/175 & 158/175 & 6/35 \\ -58/175 & 6/175 & -33/35 \end{bmatrix} \begin{bmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{bmatrix} = \begin{bmatrix} 14 & 21 & -14 \\ 0 & 175 & -70 \\ 0 & 0 & 35 \end{bmatrix}.$$

Finalmente podemos escrever

$$A = \begin{bmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{bmatrix} = \begin{bmatrix} 6/7 & -69/175 & -58/175 \\ 3/7 & 158/175 & 6/175 \\ -2/7 & 6/35 & -33/35 \end{bmatrix} \begin{bmatrix} 14 & 21 & -14 \\ 0 & 175 & -70 \\ 0 & 0 & 35 \end{bmatrix} = QR. \triangleleft$$

5 Matrizes Ortogonais preservam comprimento e ângulo

Vamos aproveitar para falar um pouco mais sobre matrizes ortogonais. Um bom exemplo de matriz ortogonal 2×2 é dada pela matriz de rotação, ou seja, pela matriz

$$Q = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix},$$

onde $0 \leq \theta \leq 2\pi$. Podemos ver que esta matriz define uma transformação linear no plano que rotaciona vetores em um ângulo θ no sentido anti-horário. Assim, tal transformação preserva comprimento e ângulo entre vetores. Estas propriedades não são particulares da matriz de rotação: todas as matrizes ortogonais se comportam da mesma forma. O teorema a seguir traz estas informações e mais algumas.

Teorema 8. *Seja $Q = [\vec{u}_1, ..., \vec{u}_n]$ uma matriz ortogonal $n \times n$. Então, para quaisquer vetores $\vec{v}$ e $\vec{w}$ em $\mathbb{R}^n$ temos*

i. $Q\vec{v} \cdot Q\vec{w} = \vec{v} \cdot \vec{w}$

ii. $\|Q\vec{v}\| = \|\vec{v}\|$

iii. o ângulo entre $Q\vec{v}$ e $Q\vec{w}$ é igual ao ângulo entre $\vec{v}$ e $\vec{w}$

A prova do primeiro item é consequência direta da linearidade do produto interno e da hipótese das colunas formarem uma base ortonormal: se $\vec{v} = (\alpha_1, ..., \alpha_n)$ e $\vec{w} = (\beta_1, ..., \beta_n)$, então

$$Q\vec{v} \cdot Q\vec{w} = \left(\sum_{i=1}^{n} \alpha_i \vec{u}_i\right) \cdot \left(\sum_{j=1}^{n} \beta_j \vec{u}_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \beta_j \; \vec{u}_i \cdot \vec{u}_j$$

devido a linearidade do produto interno (ou seja a propriedade distributiva aplicada tanto na primeira quanto na segunda componente do produto interno). Agora, lembrando que $\vec{u}_i \cdot \vec{u}_j = 0$ se $i \neq j$, a ultima expressão iguala-se a

$$\sum_{i=1}^{n} \alpha_i \beta_i \; \vec{u}_i \cdot \vec{u}_i.$$

Finalmente, lembrando que cada coluna $\vec{u}_i$ de Q tem comprimento unitário, concluimos que

$$Q\vec{v} \cdot Q\vec{w} = \sum_{i=1}^{n} \alpha_i \beta_i = \vec{v} \cdot \vec{w}.$$

Os próximos dois itens seguem diretamente do primeiro item, visto que

$$\|Q\vec{v}\| = \sqrt{Q\vec{v} \cdot Q\vec{v}} = \sqrt{\vec{v} \cdot \vec{v}} = \|\vec{v}\|$$

e o ângulo entre Qv e Qw é definido como

$$\arccos\left(\frac{Qv \cdot Qw}{\|Qv\|\|Qw\|}\right) = \arccos\left(\frac{v \cdot w}{\|v\|\|w\|}\right),$$

que por definição é o ângulo entre v e w. Assim finalizamos a prova do teorema.

Um outro exemplo de matriz ortogonal é dado pelas matrizes de reflexão. Um caso particular é dado pela reflexão em torno do eixo horizontal:

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