Ângulo entre vetores e ortogonalidade

Junior R. Ribeiro
jrodrib@usp.br

25 de setembro de 2020

Conteúdo

1 Ângulo entre dois vetores em $\mathbb{R}^n$ 1

2 Ortogonalidade 2

3 Ortogonalidade para conjuntos de vetores 2

4 Dimensão 3

1 Ângulo entre dois vetores em $\mathbb{R}^n$

Vamos usar a lei dos cossenos

$$a^2 = b^2 + c^2 - 2bc\cos(\hat{A}),$$

e o seguinte raciocínio:

> Tome dois vetores u e v com um ângulo θ entre eles e tome o vetor diferença $u - v$. Faça o desenho se achar necessário. Esses três vetores formam um triângulo cujos lados medem $||u||$, $||v||$ e $||u - v||$. Vamos aplicar a lei dos cossenos e obter a fórmula para calcular o ângulo entre u e v. Considere $||u|| = b$ e $||v|| = c$ na fórmula da lei dos cossenos, e lembre que $||u||^2 = u_1^2 + ... + u_n^2 = u^T u$, considerando u um vetor coluna e u^T o seu transposto.

$$\begin{aligned}
||u - v||^2 &= ||u||^2 + ||v||^2 - 2||u||||v||\cos(\theta)\\
(u - v)^T(u - v) &= u^T u + v^T v - 2||u||||v||\cos(\theta)\\
u^T u - u^T v - v^T u + v^T v &= u^T u + v^T v - 2||u||||v||\cos(\theta)\\
-u^T v - v^T u &= -2||u||||v||\cos(\theta),
\end{aligned}$$

mas ambos $u^T v$ e $v^T u$ são escalares, de modo que podemos tomar seu transposto e fazer $v^T u = (v^T u)^T = u^T v$, obtendo

$$\begin{aligned} -u^T v - u^T v &= -2||u||||v|| \cos(\theta) \\ -2u^T v &= -2||u||||v|| \cos(\theta) \\ \cos(\theta) &= \frac{u^T v}{||u||||v||}. \end{aligned}$$

ou com notação de produto interno/produto escalar

$$\cos(\theta) = \frac{\langle u, v\rangle}{||u||||v||}.$$

Conclusão: o cosseno do ângulo entre dois vetores é o produto interno entre eles normalizado pelas normas deles.

2 Ortogonalidade

Quando dois vetores formam 90° entre si, são chamados de ortogonais. Mas perceba que $\cos(90°) = 0$. Assim, dois vetores são ortogonais quando

$$0 = \cos(\theta) = \frac{u^T v}{||u||||v||},$$

ou seja, quando

$$\begin{aligned} u^T v = 0 &\qquad \text{(notação matricial)}, \\ \langle u, v\rangle = 0 &\qquad \text{(notação algébrica)}. \end{aligned}$$

3 Ortogonalidade para conjuntos de vetores

Tome um conjunto de vetores de $\mathbb{R}^n$ não nulos $\{x_1, ..., x_n\}$, ortogonais dois a dois. Ou seja,

$$\begin{gathered} x_1^T x_2 = 0, \quad x_1^T x_3 = 0, \quad ..., \quad x_1^T x_n = 0, \\ x_2^T x_3 = 0, \quad x_2^T x_4 = 0, \quad ..., \quad x_2^T x_n = 0, \\ \vdots \\ x_{n-2}^T x_{n-1} = 0, \quad x_{n-2}^T x_n = 0, \\ x_{n-1}^T x_n = 0. \end{gathered}$$

Perceba que $x_1^T x_1 \neq 0$, $x_2^T x_2 \neq 0$, ..., $x_n^T x_n \neq 0$, pois

$$x_1^T x_1 = [x_1]_1^2 + [x_1]_2^2 + ... + [x_1]_n^2 = ||x_1||^2,$$

mas[1] x_1 é não nulo e com isso $||x_1|| \neq 0$. O mesmo fato se aplica aos demais vetores. Esse conjunto é uma base ortogonal de $\mathbb{R}^n$, pois são de fato linearmente independentes e geram todo o $\mathbb{R}^n$, além de serem ortogonais.

[1] Temos a r-ésima componente do vetor x_1 representada por $[x_1]_r$.

Vamos considerar os vetores $\{x_1, ..., x_n\}$ todos com norma 1, ou seja, $||x_i|| = 1$ para $i = 1, ..., n$. Perceba que estamos impondo uma nova propriedade para o conjunto $\{x_1, ..., x_n\}$. Quando esse conjunto é ortogonal e todos os vetores têm norma 1, chamamo-los de conjunto ortonormal.

Ora, se $||x|| = 1$, logicamente $||x||^2 = 1 = x^T x$.

Considere uma matriz Q cujas colunas são os tais vetores $\{x_1, ..., x_n\}$ ortonormais. Multiplique $Q^T Q$. O que isso resulta?

Fazer a multiplicação $Q^T Q$ é o mesmo que

$$\begin{bmatrix} x_1 & x_2 & \dots & x_n \\ x_1 & x_2 & \dots & x_n \\ \vdots & \dots & \dots & \vdots \\ x_1 & x_2 & \dots & x_n \end{bmatrix}^T \begin{bmatrix} x_1 & x_2 & \dots & x_n \\ x_1 & x_2 & \dots & x_n \\ \vdots & \dots & \dots & \vdots \\ x_1 & x_2 & \dots & x_n \end{bmatrix} = \begin{bmatrix} x_1 & x_1 & \dots & x_1 \\ x_2 & x_2 & \dots & x_2 \\ \vdots & \dots & \dots & \vdots \\ x_n & x_n & \dots & x_n \end{bmatrix} \begin{bmatrix} x_1 & x_2 & \dots & x_n \\ x_1 & x_2 & \dots & x_n \\ \vdots & \dots & \dots & \vdots \\ x_1 & x_2 & \dots & x_n \end{bmatrix} = \begin{bmatrix} x_1^T x_1 & x_1^T x_2 & \dots & x_1^T x_n \\ x_2^T x_1 & x_2^T x_2 & \dots & x_2^T x_n \\ \vdots & \dots & \dots & \vdots \\ x_n^T x_1 & x_n^T x_2 & \dots & x_n^T x_n \end{bmatrix}$$

Note que os elementos da diagonal são 1 e os demais são zero.

Agora, tome uma matriz Q retangular, com mais linhas que colunas. Considere suas colunas vetores ortonormais. Como fica o produto $Q^T Q$?

$$\begin{bmatrix} x_1 & x_2 \\ x_1 & x_2 \\ \vdots & \vdots \\ x_1 & x_2 \end{bmatrix}^T \begin{bmatrix} x_1 & x_2 \\ x_1 & x_2 \\ \vdots & \vdots \\ x_1 & x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_1 & \dots & x_1 \\ x_2 & x_2 & \dots & x_2 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \\ x_1 & x_2 \\ \vdots & \vdots \\ x_1 & x_2 \end{bmatrix} = \begin{bmatrix} x_1^T x_1 & x_1^T x_2 \\ x_2^T x_1 & x_2^T x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Por fim, tome duas matrizes Q e Z.

$$\begin{bmatrix} x_1 & \dots & x_p \\ x_1 & \dots & x_p \\ \vdots & \dots & \vdots \\ x_1 & \dots & x_p \end{bmatrix}^T \begin{bmatrix} z_1 & \dots & z_q \\ z_1 & \dots & z_q \\ \vdots & \dots & \vdots \\ z_1 & \dots & z_q \end{bmatrix} = \begin{bmatrix} x_1 & x_1 & \dots & x_1 \\ \vdots & \vdots & \dots & \vdots \\ x_2 & x_2 & \dots & x_2 \end{bmatrix} \begin{bmatrix} z_1 & \dots & z_q \\ z_1 & \dots & z_q \\ \vdots & \vdots & \vdots \\ z_1 & \dots & z_q \end{bmatrix} = \begin{bmatrix} x_1^T z_1 & x_1^T z_2 & \dots & x_1^T z_q \\ x_2^T z_1 & x_2^T z_2 & \dots & x_2^T z_q \\ \vdots & \vdots & \dots & \vdots \\ x_p^T z_1 & x_p^T z_2 & \dots & x_p^T z_q \end{bmatrix}$$

Se o conjunto $\{x_1, ..., x_p\}$ for ortogonal ao conjunto $\{z_1, ..., z_q\}$, então a matriz resultante é nula, pois $x_i^T z_j = 0$ para todo i e todo j.

4 Dimensão

Precisamos ter bem claro o que é um **espaço** vetorial [veja as propriedades aqui], um **subespaço** vetorial (é um espaço está contido em um espaço vetorial e tem as mesmas propriedades de espaço vetorial) e um **vetor**. As transformações lineares agem em subespaços vetoriais, que por sua vez estão contidos em espaços vetoriais.

Chamamos o espaço vetorial de "o espaço como um todo" e existem infinitos subespaços vetoriais contidos nele. A dimensão de um vetor é a dimensão do "espaço todo", ou seja, é exatamente o número de componentes. Se o espaço vetorial é $\mathbb{R}^2$, os vetores têm dimensão 2, têm a forma $[x_1, x_2]^T$. Em $\mathbb{R}^{12}$, os vetores têm dimensão 12 e têm a forma $[x_1, ..., x_{12}]^T$. Ambos, vetores e espaços vetoriais, têm a mesma dimensão, pois os vetores "moram" no espaço vetorial.

Agora, falando de subespaços, o conceito de dimensão é diferente! A dimensão de um subespaço vetorial é a **quantidade** de vetores (vetores do espaço vetorial) linearmente independentes necessários para gerá-lo. Um conjunto l.i. que gere o subespaço é uma **base** para esse subespaço. Um conjunto l.i. que gere o espaço é uma **base** para o espaço.

Um subespaço vetorial bem trivial é constituído de apenas um ponto: o zero. O subespaço $\{\vec{\mathbf{0}}\}$ tem dimensão zero por definição. Esse é um subespaço, pois satisfaz todas as propriedades da definição no link acima.

Qualquer **reta** em $\mathbb{R}^n$, $n \geq 1$, que passe pela origem do espaço é um subespaço vetorial de $\mathbb{R}^n$ e a sua dimensão é 1, pois para gerar uma reta é necessário apenas um único vetor, e um conjunto com um único vetor é l.i. por definição. Perceba que existem infinitos vetores que geram essa mesma reta, portanto existem infinitas bases para esse subespaço. Se $n > 1$, existem infinitas retas em $\mathbb{R}^n$ que são subespaços vetoriais (se $n = 1$, o subespaço vetorial coincide com o espaço todo, a reta real).

Qualquer **plano** em $\mathbb{R}^n$, $n \geq 2$, que passe pela origem é um subespaço vetorial e sua dimensão é 2, pos são necessários dois vetores l.i. para gerar um plano, e note que existem infinitos pares de vetores l.i que geram esse plano (infinitas bases para esse subespaço vetorial). E note que, se $n > 2$, existem infinitos planos em $\mathbb{R}^n$ que são subespaços vetoriais (se $n = 2$, o subespaço coincide com o espaço todo, o plano cartesiano).

etc.
etc.
etc.

Em $\mathbb{R}^6$, por exemplo, temos infinitas retas (subespaços de dimensão 1), infinitos planos (subespaços de dimensão 2), infinitos "espaços 3d" (subespaços de dimensão 3), infinitos "espaços 4d" (subespaços de dimensão 4), infinitos "espaços 5d" (subespaços de dimensão 5) e um único subespaço 6d (o espaço todo). Cada um desses subespaços é gerado por uma base com 1, 2, 3, 4, 5 e 6 vetores l.i. respectivamente. E as tais bases são infinitas. Tome por exemplo um par de vetores l.i. $\{u, v\}$ que gerem um certo plano em $\mathbb{R}^6$, e portanto uma base para esse subespaço. Então $\{2u, -v\}$ também gera esse plano, pois é um conjunto l.i. Vemos assim que existem infinitas bases para o mesmo subespaço vetorial.

A dimensão de um subespaço vetorial é portanto o número de vetores em sua base.

Às vezes, para esclarecer certas situações, dizemos *um subespaço próprio* de $\mathbb{R}^n$, no intuito de dizer que é um subespaço contido em $\mathbb{R}^n$ com dimensão menor que n. Se nada é mencionado, o espaço todo é considerado um subespaço de si mesmo.