

# Álgebra Linear e Geometria Analítica

## Distância, Ortogonalidade, Projeção Ortogonal

Departamento de Matemática  Universidade de Aveiro

# Produto interno em $\mathbb{R}^n$

Dados os vetores  $X = (x_1, \dots, x_n)$  e  $Y = (y_1, \dots, y_n) \in \mathbb{R}^n$

- o **produto interno** (ou **produto escalar**) de  $X$  e  $Y$  é o escalar real

$$\begin{aligned} X \cdot Y &= X^T Y = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \\ &= x_1 y_1 + \cdots + x_n y_n \end{aligned}$$

Nota: Pode também utilizar-se a notação  $X|Y$  ou  $\langle X, Y \rangle$ .

- o **comprimento** ou **norma** de  $X$  é

$$\|X\| = \sqrt{X \cdot X} = \sqrt{x_1^2 + \cdots + x_n^2}$$

# Propriedades do produto interno em $\mathbb{R}^n$

Dados  $X, Y, Z \in \mathbb{R}^n$  e  $\alpha \in \mathbb{R}$ ,

1.  $X \cdot X \geq 0$ ;

2.  $X \cdot X = 0 \iff X = 0$ ;

3.  $X \cdot Y = Y \cdot X$ ;

4. i.  $(X + Y) \cdot Z = X \cdot Z + Y \cdot Z$ ,

ii.  $X \cdot (Y + Z) = X \cdot Y + X \cdot Z$ ;

5.  $(\alpha X) \cdot Y = \alpha (X \cdot Y) = X \cdot (\alpha Y)$ ;

6.  $\|\alpha X\| = |\alpha| \|X\|$ .

# Desigualdade de Cauchy-Schwarz e desigualdade triangular

Teorema (Desigualdade de Cauchy-Schwarz)

Dados  $X, Y \in \mathbb{R}^n$ ,

$$|X \cdot Y| \leq \|X\| \|Y\|.$$

Teorema (Desigualdade Triangular)

Dados  $X, Y \in \mathbb{R}^n$ ,

$$\|X + Y\| \leq \|X\| + \|Y\|.$$

# Ângulo entre vetores

Em  $\mathbb{R}^2$ , sejam  $X = (\underline{x}, 0)$ ,  $\underline{x} > 0$

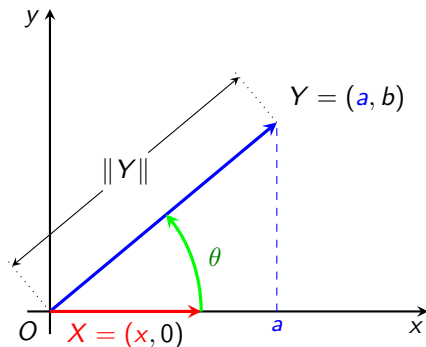
e  $Y = (\underline{a}, b) \neq (0, 0)$

vetores não nulos. Temos:

- $X \cdot Y = \underline{x}\underline{a}$  e  $\|X\| = x$

- $\frac{X \cdot Y}{\|X\|} = \underline{a} = \|Y\| \cos(\theta)$

Logo,  $\cos(\theta) = \frac{X \cdot Y}{\|X\| \|Y\|}$ ,  $\theta \in [0, \pi]$



Em geral, para  $X, Y \in \mathbb{R}^n$ ,  $X, Y \neq 0$ , o ângulo entre os vetores  $X$  e  $Y$  é

$$\theta = \angle(X, Y) = \arccos \frac{X \cdot Y}{\|X\| \|Y\|} = \arccos \left( \frac{X}{\|X\|} \cdot \frac{Y}{\|Y\|} \right).$$

Nota: pela desigualdade de Cauchy-Schwarz  $\left| \frac{X \cdot Y}{\|X\| \|Y\|} \right| \leq 1$  e  $\theta \in [0, \pi]$ .

# Vetores ortogonais, colineares, com mesmo sentido e unitários

- Dados os vetores  $X, Y \in \mathbb{R}^n$ ,  $X, Y \neq 0$ 
  - ▶  $X$  e  $Y$  são **ortogonais ou perpendiculares**,  $X \perp Y$ ,  
se  $\theta = \frac{\pi}{2}$ , i.e. se  $X \cdot Y = 0$ .
  - ▶  $X$  e  $Y$  são **colineares ou paralelos ou têm a mesma direção**,  
se  $\theta = 0$  ou  $\theta = \pi$ , i.e. se  $|X \cdot Y| = \|X\| \|Y\|$ .
    - ▶  $X$  e  $Y$  **têm o mesmo sentido**,  
se  $\theta = 0$ , i.e. se  $X \cdot Y = \|X\| \|Y\|$ .
    - ▶  $X$  e  $Y$  **têm sentido oposto ou contrário**,  
se  $\theta = \pi$ , i.e. se  $X \cdot Y = -\|X\| \|Y\|$ .

Por convenção, se  $X = 0$  ou  $Y = 0$ , então  $X$  e  $Y$  são colineares e ortogonais.

- Um vetor **unitário** é um vetor de norma igual a 1.

Se  $X \neq 0$ , o vetor

$$U = \frac{1}{\|X\|} X$$

é um vetor unitário com a mesma direção e sentido de  $X$ .

# Conjunto ortogonal e ortonormado em $\mathbb{R}^n$

Um conjunto  $\{X_1, \dots, X_k\}$  de vetores de  $\mathbb{R}^n$  diz-se **ortogonal** se

$$X_i \cdot X_j = 0, \quad i \neq j, \quad i, j = 1, \dots, k,$$

e diz-se **ortonormado (o.n.)** se é ortogonal e também se verifica

$$X_i \cdot X_i = 1, \quad i = 1, \dots, k.$$

- Exemplo:**
1.  $\{(1, 1, 0), (2, -2, 1)\}$  é ortogonal;
  2.  $\left\{\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0\right), \left(\frac{2}{3}, -\frac{2}{3}, \frac{1}{3}\right)\right\}$  é o.n.

**Teorema:** Todo o conjunto ortogonal de vetores não nulos é l.i.

**Corolário:** Todo o conjunto o.n. é l.i.

# Coordenadas de um vetor de $\mathbb{R}^n$ numa base o.n.

Uma **base ortogonal/o.n.** é uma base que é um conjunto ortogonal/o.n.

**Nota:** Todo o conjunto o.n. de  $n$  vetores de  $\mathbb{R}^n$  é uma **base** de  $\mathbb{R}^n$ .

**Teorema:** Seja  $X \in \mathbb{R}^n$  e  $\mathcal{B} = (X_1, \dots, X_n)$  uma **base o.n.** de  $\mathbb{R}^n$ . Então

$$[X]_{\mathcal{B}} = \begin{bmatrix} X \cdot X_1 \\ \vdots \\ X \cdot X_n \end{bmatrix},$$

isto é,  $X = a_1 X_1 + \dots + a_n X_n$ , sendo  $a_i = X \cdot X_i$ ,  $i = 1, \dots, n$ .

**Exemplo:** Determinar as coordenadas do vetor  $(1, 5)$  na base o.n. de  $\mathbb{R}^2$

$$\left( \left( \frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2} \right), \left( \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right) \right).$$



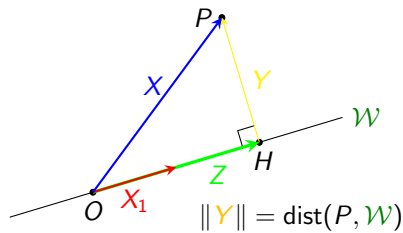
# Projeção ortogonal em $\mathbb{R}^n$

$Y \in \mathbb{R}^n$  é ortogonal ao subespaço  $\mathcal{W}$  de  $\mathbb{R}^n$  se  $Y \cdot Z = 0$  para cada  $Z \in \mathcal{W}$ .

**Teorema:** Seja  $Y \in \mathbb{R}^n$  e  $\mathcal{B}$  uma base de um subespaço  $\mathcal{W}$  de  $\mathbb{R}^n$ . Então,  
 $Y$  é ortogonal a  $\mathcal{W}$  se e só se  $Y$  é ortogonal a cada vetor de  $\mathcal{B}$ .

A **projeção ortogonal** de  $X \in \mathbb{R}^n$  sobre o subespaço  $\mathcal{W}$  de  $\mathbb{R}^n$  é o vetor  $Z = \text{proj}_{\mathcal{W}} X \in \mathcal{W}$  tal que  $X = Y + Z$ , sendo  $Y$  ortogonal a  $\mathcal{W}$ .

**Exemplo:** Sejam  $\mathcal{W} = \langle X_1 \rangle$  uma reta,  $\{X_1\}$  base o.n. de  $\mathcal{W}$  e  $X = \overrightarrow{OP}$ . Logo,  $Z = \text{proj}_{\mathcal{W}} X = \alpha X_1$  e  $Y \cdot X_1 = 0$ .  
Então, se  $X = Y + Z = Y + \alpha X_1$ ,  
 $X \cdot X_1 = Y \cdot X_1 + \alpha X_1 \cdot X_1 = \alpha$ .  
Portanto,  $\text{proj}_{\mathcal{W}} X = (X \cdot X_1) X_1$ .

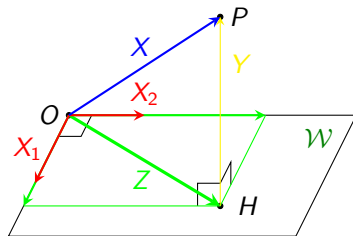


# Projeção ortogonal sobre um plano em $\mathbb{R}^3$

**Exemplo:** Sejam  $\mathcal{W}$  um plano gerado pela **base o.n.**  $\{X_1, X_2\}$  e  $X = \overrightarrow{OP} = Z + Y$ , com  $Z = \text{proj}_{\mathcal{W}} X = \alpha_1 X_1 + \alpha_2 X_2$  e  $Y \cdot X_1 = Y \cdot X_2 = 0$ . Então, sendo

$$X = Y + Z = Y + \alpha_1 X_1 + \alpha_2 X_2, \\ X \cdot X_1 = \alpha_1 \text{ e } X \cdot X_2 = \alpha_2.$$

$$\text{Logo, } \text{proj}_{\mathcal{W}} X = (X \cdot X_1)X_1 + (X \cdot X_2)X_2.$$



$$\|Y\| = \text{dist}(P, \mathcal{W})$$

**Teorema:** A projeção ortogonal de  $X \in \mathbb{R}^n$  sobre o subespaço  $\mathcal{W}$  de  $\mathbb{R}^n$  é

$$\text{proj}_{\mathcal{W}} X = (X \cdot X_1)X_1 + \cdots + (X \cdot X_k)X_k \in \mathcal{W},$$

em que  $\{X_1, \dots, X_k\}$  é uma **base o.n.** de  $\mathcal{W}$ .

**Nota:**  $Y = X - \text{proj}_{\mathcal{W}} X$  é ortogonal a todos os vetores de  $\mathcal{W}$ .

# Método de ortonormalização de Gram-Schmidt (opcional)

**Teorema:** Todo o subespaço  $\mathcal{W} \neq \{0\}$  de  $\mathbb{R}^n$  possui uma base o.n.

**Ideia da Demonstração:**

Dada  $\{X_1, \dots, X_m\}$  uma base de  $\mathcal{W}$ , sejam  $Y_1 = \frac{X_1}{\|X_1\|}$ , e

$$X'_k = X_k - \sum_{i=1}^{k-1} (X_k \cdot Y_i) Y_i, \quad Y_k = \frac{X'_k}{\|X'_k\|},$$

para  $k = 2, \dots, m$ . Pode verificar-se que  $\mathcal{B} = \{Y_1, \dots, Y_m\}$  é um conjunto o.n., logo l.i. em  $\mathcal{W}$ . Sendo  $\dim \mathcal{W} = m$ , conclui-se que  $\mathcal{B}$  é uma base o.n. de  $\mathcal{W}$ .

**Exemplo:**

Determinar uma base o.n. de  $\langle (1, 1, 1, 1), (1, 2, -1, 3), (2, 1, -2, 2) \rangle$ .

# Produto externo em $\mathbb{R}^3$

Dados os vetores  $X = (x_1, x_2, x_3)$  e  $Y = (y_1, y_2, y_3) \in \mathbb{R}^3$ ,

- o **produto externo** (ou **produto vetorial**) de  $X$  e  $Y$  é o vetor de  $\mathbb{R}^3$

$$X \times Y = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1).$$

**Nota:** Para determinar o produto externo pode utilizar-se COMO AUXILIAR DE CÁLCULO o seguinte “determinante simbólico”

$$X \times Y \longleftrightarrow \begin{vmatrix} i & j & k \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} \quad \text{com} \quad \begin{aligned} i &= (1, 0, 0) \\ j &= (0, 1, 0) \\ k &= (0, 0, 1) \end{aligned}$$

fazendo o seu desenvolvimento pela primeira linha.

# Propriedades do produto externo em $\mathbb{R}^3$

Dados  $X, Y, Z \in \mathbb{R}^3$ ,  $\alpha \in \mathbb{R}$ , e  $O$  o vetor nulo de  $\mathbb{R}^3$

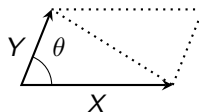
1.  $X \times Y = -(Y \times X)$ ;
2.
  - i.  $X \times (Y + Z) = X \times Y + X \times Z$ ,
  - ii.  $(X + Y) \times Z = X \times Z + Y \times Z$ ;
3.  $\alpha(X \times Y) = (\alpha X) \times Y = X \times (\alpha Y)$ ;
4.  $X \times X = O$ ;
5.  $X \times O = O \times X = O$ ;
6.  $X \times Y$  é ortogonal a  $X$  e a  $Y$ .
7.  $\|X \times Y\| = \|X\| \|Y\| \sin(\theta)$ , onde  $\theta$  é o ângulo entre  $X$  e  $Y$ .

# Uma aplicação do produto externo

Sejam  $X, Y, Z \in \mathbb{R}^3$ , então

- a **área do paralelogramo** com lados correspondentes aos vetores  $X, Y$  é

$$A_{\diamond} = \|X \times Y\|$$



- a **área do triângulo** com dois dos seus lados correspondentes aos vetores  $X, Y$  é

$$A_{\triangle} = \frac{\|X \times Y\|}{2}$$

## Aplicações do Produto Interno. Relembrar retas em $\mathbb{R}^3$

Dada uma reta  $\mathcal{R}$  em  $\mathbb{R}^3$  que passa pelo ponto  $P$  e tem vetor diretor  $v$ , temos

$$X \in \mathcal{R} \iff \exists \alpha \in \mathbb{R} : \overrightarrow{OX} = \overrightarrow{OP} + \alpha v.$$

Uma **equação vetorial** da reta  $\mathcal{R}$  é  $\overrightarrow{OX} = \overrightarrow{OP} + \alpha v$ ,  $\alpha \in \mathbb{R}$ , a partir da qual se obtêm as **equações paramétricas** de  $\mathcal{R}$ :

$$\begin{cases} x = x_0 + \alpha v_1 \\ y = y_0 + \alpha v_2 \\ z = z_0 + \alpha v_3 \end{cases}, \quad \alpha \in \mathbb{R},$$

sendo  $X(x, y, z)$ ,  $P(x_0, y_0, z_0)$  e  $v = (v_1, v_2, v_3)$ .

Eliminando o parâmetro  $\alpha$  do anterior sistema, obtém-se um sistema de grau 1 com 3 incógnitas e 2 equações, ditas as **equações cartesianas** de  $\mathcal{R}$ .

# Planos em $\mathbb{R}^3$ – Equações vetoriais e paramétricas

Dado um plano  $\mathcal{P}$  em  $\mathbb{R}^3$  que passa pelo ponto  $P$  e tem vetores diretores  $u$  e  $v$  (não colineares),

$$X \in \mathcal{P} \iff \exists \alpha, \beta \in \mathbb{R} : \overrightarrow{OX} = \overrightarrow{OP} + \alpha u + \beta v.$$

Uma **equação vetorial** do plano  $\mathcal{P}$  é

$$\overrightarrow{OX} = \overrightarrow{OP} + \alpha u + \beta v, \quad \alpha, \beta \in \mathbb{R},$$

a partir da qual se obtêm as **equações paramétricas** de  $\mathcal{P}$ :

$$\begin{cases} x = x_0 + \alpha u_1 + \beta v_1 \\ y = y_0 + \alpha u_2 + \beta v_2, \\ z = z_0 + \alpha u_3 + \beta v_3 \end{cases} \quad \alpha, \beta \in \mathbb{R},$$

com  $X(x, y, z)$ ,  $P(x_0, y_0, z_0)$ ,  $u = (u_1, u_2, u_3)$ ,  $v = (v_1, v_2, v_3)$ .



## Planos em $\mathbb{R}^3$ – Equações cartesianas

Eliminando os parâmetros  $\alpha$  e  $\beta$  do anterior sistema, obtém-se uma equação

$$ax + by + cz + d = 0,$$

dita equação (cartesiana) geral do plano  $\mathcal{P}$ .

Verifica-se que  $w = (a, b, c)$  é um vetor não nulo ortogonal a  $\mathcal{P}$ . De facto, dois pontos arbitrários deste plano,  $P_i(x_i, y_i, z_i)$ ,  $i = 0, 1$ , satisfazem

$$ax_i + by_i + cz_i + d = 0, \quad i = 0, 1,$$

donde

$$a(x_1 - x_0) + b(y_1 - y_0) + c(z_1 - z_0) = 0,$$

ou seja, para qualquer vetor  $\overrightarrow{P_0P_1}$  do plano  $\mathcal{P}$ , tem-se

$$w \cdot \overrightarrow{P_0P_1} = 0.$$

# Distâncias

A **distância entre dois pontos**  $P$  e  $Q$  de  $\mathbb{R}^n$  é

$$d(P, Q) = \|\overrightarrow{PQ}\|.$$

Em particular, para  $Q(x_1, \dots, x_n)$  e  $P(y_1, \dots, y_n)$ , tem-se

$$d(P, Q) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}.$$

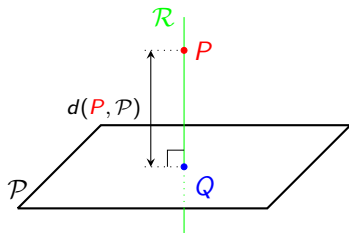
Dados um ponto, reta ou plano  $\mathcal{F}$  e um ponto, reta ou plano  $\mathcal{G}$  de  $\mathbb{R}^3$ , a **distância entre  $\mathcal{F}$  e  $\mathcal{G}$**  é

$$d(\mathcal{F}, \mathcal{G}) = \min \{ d(P, Q) : P \in \mathcal{F}, Q \in \mathcal{G} \}.$$

**Nota:** Se  $\mathcal{F} \cap \mathcal{G} \neq \emptyset$ , então  $d(\mathcal{F}, \mathcal{G}) = 0$ . Interessa-nos o caso em que  $\mathcal{F}$  é um ponto e  $\mathcal{G}$  é uma reta ou um plano, respetivamente.

# Distância de um ponto a um plano

Dados um plano  $\mathcal{P}$  e um ponto  $P \notin \mathcal{P}$ , existe uma única reta  $\mathcal{R}$  perpendicular ao plano  $\mathcal{P}$  e contendo o ponto  $P$ .



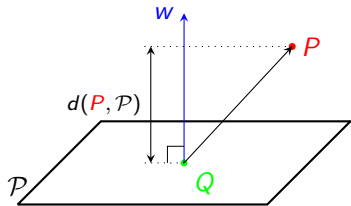
A distância do ponto  $P$  ao plano  $\mathcal{P}$  é

$$d(P, \mathcal{P}) = d(P, Q),$$

em que  $Q$  é o ponto de interseção da reta  $\mathcal{R}$  com o plano  $\mathcal{P}$ .

# Distância de um ponto a um plano (equação geral)

Dados um plano  $\mathcal{P}$  e um ponto  $P \notin \mathcal{P}$ , sejam  $Q \in \mathcal{P}$  e  $w$  um vetor não nulo ortogonal ao plano  $\mathcal{P}$ .



Então,

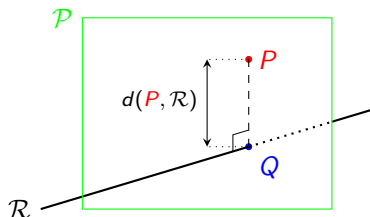
$$d(P, \mathcal{P}) = \|\text{proj}_{\langle w \rangle} \overrightarrow{QP}\| = \frac{|\overrightarrow{QP} \cdot w|}{\|w\|}.$$

Sendo  $P(x_0, y_0, z_0)$  e  $ax + by + cz + d = 0$  uma equação geral do plano  $\mathcal{P}$ , tem-se

$$d(P, \mathcal{P}) = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

# Distância de um ponto a uma reta

Dada uma **reta**  $\mathcal{R}$  e um **ponto**  $P \notin \mathcal{R}$ , existe um único **plano**  $\mathcal{P}$  perpendicular a  $\mathcal{R}$  e que contém  $P$ .



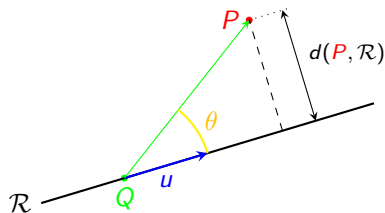
A **distância** do ponto  $P$  à **reta**  $\mathcal{R}$  é

$$d(P, \mathcal{R}) = d(P, Q),$$

em que  $Q$  é o ponto de interseção da **reta**  $\mathcal{R}$  com o **plano**  $\mathcal{P}$ .

# Distância de um ponto a uma reta (equação vetorial)

Dada uma **reta**  $\mathcal{R}$  que passa pelo **ponto**  $Q$  e que tem **vetor diretor**  $u$ ,



e um **ponto**  $P \notin \mathcal{R}$ , tem-se que

$$d(P, \mathcal{R}) = \|\overrightarrow{QP}\| |\sin(\theta)| = \frac{\|u \times \overrightarrow{QP}\|}{\|u\|},$$

sendo  $\theta$  o ângulo entre os vetores  $u$  e  $\overrightarrow{QP}$ .

# Problema dos mínimos quadrados

- ▶ Em muitas aplicações encontramos sistemas sem solução (impossíveis ou inconsistentes). Ao mesmo tempo gostaríamos de encontrar uma solução. **Como fazer?**
- ▶ Uma situação típica é a chamada **regressão linear**:  
*dada um conjunto de pontos do plano (que muitas vezes é o resultado de uma série de medidas com alguma incerteza), encontre a reta que “melhor se aproxima” desses pontos.*
- ▶ Este é um caso especial de uma classe maior de problemas de modelos lineares, onde pretendemos encontrar a **“melhor aproximação possível”**.
- ▶ Sistemas inconsistentes de equações lineares podem ser pensados como **como problemas de mínimos quadrados**, onde a (s) solução (ões) é (são) a (as) “melhor (es) possível (eis)” em certo sentido.

Considere o sistema de equações lineares  $Ax = b$ , onde  $A$  é  $m \times n$ ,  $x \in \mathbb{R}^n$  e  $b \in \mathbb{R}^m$ .

Sabemos que o sistema é possível

$$\Leftrightarrow b \in \mathcal{C}(A).$$

Se  $b$  não está em  $\mathcal{C}(A)$ , o que podemos fazer?

Em vez de tentar resolver o problema  $Ax = b$ , podemos tentar fazer com que o erro  $\|b - Ax\|$  seja o menor possível, ou seja, quanto menor for a distância entre  $b$  e  $Ax$ , dada por  $\|b - Ax\|$ , melhor é a aproximação.

Isto é chamado **problema de mínimos quadrados** (porque a expressão  $\|b - Ax\|^2$  é uma soma de quadrados).

### Definição

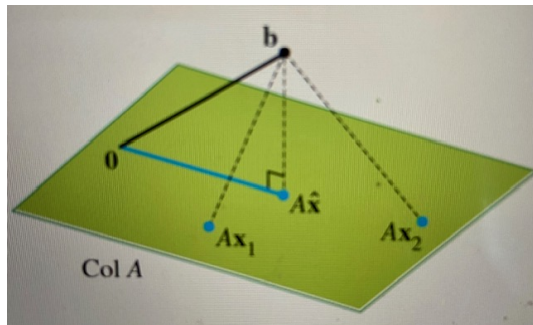
Seja  $A \in m \times n$  e  $b \in \mathbb{R}^m$ . Dizemos que  $\hat{x} \in \mathbb{R}^n$  é uma **solução dos mínimos quadrados** do sistema  $Ax = b$  se

$$\|b - A\hat{x}\| \leq \|b - Ax\|$$

para todo  $x \in \mathbb{R}^n$ .

**Nota:** Uma vez que  $\mathcal{C}(A) = \{Ax : x \in \mathbb{R}^n\}$ ,  $\hat{x}$  é tal que  $A\hat{x}$  dá a melhor aproximação de  $b$  entre todos os vetores em  $\mathcal{C}(A)$ .





**Figure:** O vetor  $b$  é o mais próximo a  $A\hat{x}$  do que a qualquer vetor  $Ax$  para qualquer  $x$ .

### Teorema da melhor aproximação

Seja  $W$  um subspaço de  $\mathbb{R}^n$  e  $y$  qualquer vetor em  $\mathbb{R}^n$ . Seja  $\hat{y}$  a projeção ortogonal de  $y$  em  $W$ . Então  $\hat{y}$  é o vetor em  $W$  mais próximo de  $y$ , no sentido que

$$\|y - \hat{y}\| < \|y - v\|,$$

para qualquer  $v$  em  $W$  distinto de  $\hat{y}$ .

- ▶ O vetor  $\hat{y}$  é chamado a melhor aproximação para  $y$  por elementos de  $W$ .
- ▶ A distância de  $y$  a  $v$ , dada por  $\|y - v\|$ , é encarada como o erro ao usar  $v$  em vez de  $y$ .
- ▶ O teorema anterior diz que esse erro é minimizado quando  $v = \hat{y}$ .

Do Teorema da melhor aproximação, obtemos que a melhor solução dos mínimos quadrados  $\hat{x}$  é determinada usando

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

onde  $\hat{b} := \text{Proj}_W(b)$  com  $W := \mathcal{C}(A)$ .

Se  $\hat{b} \in W = \mathcal{C}(A)$ , assim o sistema  $A\hat{x} = \hat{b}$  é sempre consistente. Logo, existem sempre soluções dos mínimos quadrados e estas são encontradas ao resolvermos o sistema  $A\hat{x} = \hat{b}$ .

Assim, podemos escrever a seguinte observação básica:

$\hat{x} \in \mathbb{R}^n$  é uma solução dos mínimos quadrados do sistema  $Ax = b$



$\hat{x}$  é uma solução do sistema (possível)  $A\hat{x} = \hat{b}$   
onde  $\hat{b} = \text{Proj}_W(b)$  e  $W = \mathcal{C}(A)$ .

### Nota:

- ▶ Assuma-se que o sistema  $Ax = b$  é consistente, ou seja,  $b \in W = \mathcal{C}(A)$ . Então  $\hat{b} = \text{Proj}_W(b) = b$ . Assim, soluções dos mínimos quadrados são as mesmas soluções usuais do sistema.
- ▶ Suponhamos que as colunas de  $A$  são linearmente independentes (por outras palavras, a forma escalonada de  $A$  tem pivots em todas as colunas). Então o sistema  $A\hat{x} = \hat{b}$  tem uma solução única. Isto significa que  $Ax = b$  tem uma única **solução dos mínimos quadrados**.
- ▶ Se  $A$  tem colunas linearmente dependentes, segue de um modo semelhante que  $Ax = b$  tem um número infinito de soluções dos mínimos quadrados.
- ▶  $\|b - A\hat{x}\| = \|b - \hat{b}\|$  é chamado o **erro dos mínimos quadrados**: dá a mínima distância entre  $b$  e vetores da forma  $Ax$ .

O método que apresentámos para calcular a solução dos mínimos quadrados é eficiente mas é necessário usar muitos cálculos.

Para calcular  $\hat{b} = \text{Proj}_W(b)$  precisamos primeiro de uma base ortogonal para  $W = \mathcal{C}(A)$  (ver método Gram-Schmidt).

Depois precisamos de calcular  $\hat{b}$ . E finalmente resolver o sistema  $A\hat{x} = \hat{b}$ .

Existe um método mais simples!

### Teorema

Considere o sistema  $Ax = b$  onde  $A$  é  $m \times n$ ,  $x \in \mathbb{R}^n$  and  $b \in \mathbb{R}^m$ . Então

$\hat{x} \in \mathbb{R}^n$  é uma solução dos mínimos quadrados do sistema  $Ax = b$



$\hat{x} \in \mathbb{R}^n$  é a solução (do sistema possível)

$$A^T Ax = A^T b.$$

Nota:

$A^T Ax = A^T b$  são muitas vezes chamadas equações normais para  $Ax = b$ .

**Exemplo 1** Encontre uma solução dos mínimos quadrados do sistema inconsistente  $Ax = b$  para

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \quad \text{e} \quad b = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}.$$

**Solução** Para usar as equações normais calculemos

$$A^T A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}.$$

$$A^T b = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}.$$

Então a equação  $A^T A x = A^T b$  fica:

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}.$$

Operações nas linhas podem ser usadas para resolver este sistema mas como  $A^T A$  é invertível e  $2 \times 2$ , é fácil calcular:

$$(A^T A)^{-1} = \frac{1}{84} \begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}$$

e, portanto, para resolver  $A^T A x = A^T b$ , vem

$$\hat{x} = (A^T A)^{-1} A^T b = \frac{1}{84} \begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix} \begin{bmatrix} 19 \\ 11 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 84 \\ 168 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Em muitas situações  $A^T A$  é invertível mas isso nem sempre acontece. Veja-se o próximo exemplo.

Encontre uma solução dos mínimos quadrados para  $Ax = b$  onde

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

$$A^T A = \begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix} \quad e \quad A^T b = \begin{bmatrix} 4 \\ -4 \\ 2 \\ 6 \end{bmatrix}.$$



A matriz ampliada para  $A^T Ax = A^T b$  é:

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

A solução geral é

$$x_1 = 3 - x_4$$

$$x_2 = -5 + x_4$$

$$x_3 = -2 + x_4$$

e  $x_4$  é livre. A solução geral dos mínimos quadrados de  $Ax = b$  é então:

$$\hat{x} = \begin{bmatrix} 3 \\ -5 \\ -2 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

# Critério para solução única dos mínimos quadrados

## Teorema

Seja  $A \in \mathbb{R}^{m \times n}$ . As afirmações seguintes são equivalentes:

1. A equação  $Ax = b$  tem uma única solução dos mínimos quadrados, para cada  $b \in \mathbb{R}^m$ .
2. As colunas de  $A$  são l.i.
3. a matriz  $A^T A$  é invertível.

Se o anterior for verdadeiro, a solução dos mínimos quadrados é dada por

$$\hat{x} = (A^T A)^{-1} A^T b.$$

# Regressão linear

**Exemplo 2:** Dado um conjunto de pontos  $(2, 1), (2, 2), (-2, 0), (-2, -1)$  em  $\mathbb{R}^2$  será que existe uma reta que contém estes pontos?

A reta será da forma  $y = \alpha x + \beta$ , onde  $\alpha, \beta \in \mathbb{R}$ . Portanto,  $\alpha$  e  $\beta$  têm de obedecer as seguintes equações

$$1 = 2\alpha + \beta$$

$$2 = 2\alpha + \beta$$

$$0 = -2\alpha + \beta$$

$$-1 = -2\alpha + \beta$$

o que é impossível!

Queremos determinar a reta "mais próxima" destes pontos. Note-se que facto de o sistema acima ser impossível quer dizer que

$$(1, 2, 0, -1) \notin \langle (2, 2, -2, -2), (1, 1, 1, 1) \rangle.$$

O vetor de  $U = \langle (2, 2, -2, -2), (1, 1, 1, 1) \rangle$  mais próximo de  $(1, 2, 0, -1)$  é  $\text{proj}_U(1, 2, 0, -1)$ .

Note-se que

$$\left(\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right), \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right)\right)$$

é uma base ortonormada de  $\langle (2, 2, -2, -2), (1, 1, 1, 1) \rangle$  (verifique!).

Logo,

$$\begin{aligned}\text{proj}_U((1, 2, 0, -1)) &= ((1, 2, 0, -1) \cdot \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right))\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) + \\ &\quad + ((1, 2, 0, -1) \cdot \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right))\left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right) \\ &= \left(\frac{3}{2}, \frac{3}{2}, -\frac{1}{2}, -\frac{1}{2}\right)\end{aligned}$$

Temos que  $\left(\frac{3}{2}, \frac{3}{2}, -\frac{1}{2}, -\frac{1}{2}\right)$  é o vetor de  $\langle (2, 2, -2, -2), (1, 1, 1, 1) \rangle$  mais próximo de  $(1, 2, 0, -1)$ .

Como  $(\frac{3}{2}, \frac{3}{2}, -\frac{1}{2}, -\frac{1}{2}) \in \langle (2, 2, -2, -2), (1, 1, 1, 1) \rangle$  temos que

$$\begin{aligned}\frac{3}{2} &= 2\alpha + \beta \\ \frac{3}{2} &= 2\alpha + \beta \\ -\frac{1}{2} &= -2\alpha + \beta \\ -\frac{1}{2} &= -2\alpha + \beta\end{aligned}$$

que tem solução  $\alpha = \frac{1}{2}$  e  $\beta = \frac{1}{2}$ .

A reta  $y = \frac{1}{2}x + \frac{1}{2}$  é a equação da reta mais próxima dos pontos  $(2, 1), (2, 2), (-2, 0), (-2, -1)$ .

## A Reta:

