2023 (7) Aula (19)

Aproximação de funções: método dos mínimos quadrados

Quarteroni: 5.7

Franco: cap 8 (abordagem diferente)

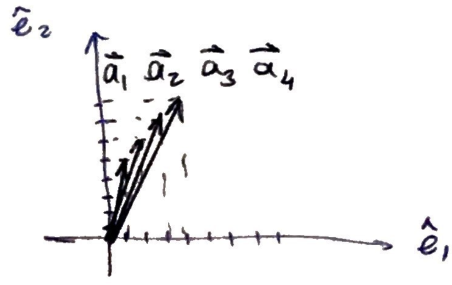
$$A\vec{x} = \vec{b} \quad , \quad A_{m\times n} \text{ de posto } K$$

- posto $(A) = K$ = nº linha L.I = nº colunas L.I de A
  - ↳ dim. espaço coluna = dim. espaço linha = dim espaço linha / coluna de A^T
  - ↳ posto (A) = posto (A^T)
  - ↳ nº linhas não-nulas da forma escalonada de A
  - ↳ $K \leq m$ e $K \leq n$
  - ↳ $\left.\begin{array}{l} \text{se } m > n \text{ e } K = n \\ \text{se } m < n \text{ e } K = m \end{array}\right\}$ posto completo.

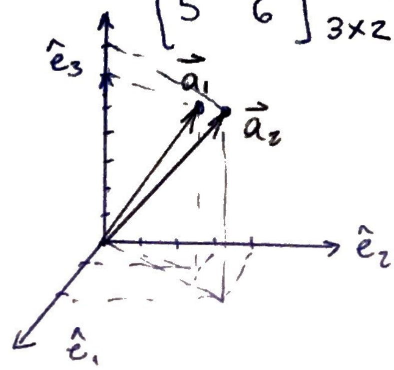
- se $m > n$ → mais eq. que incóg. → sist. superdeterminado
- se $m < n$ → menos " " " → " subdeterminado.

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}_{2\times 4}$$

$\mathbb{R}^2$



$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}_{3\times 2}$$



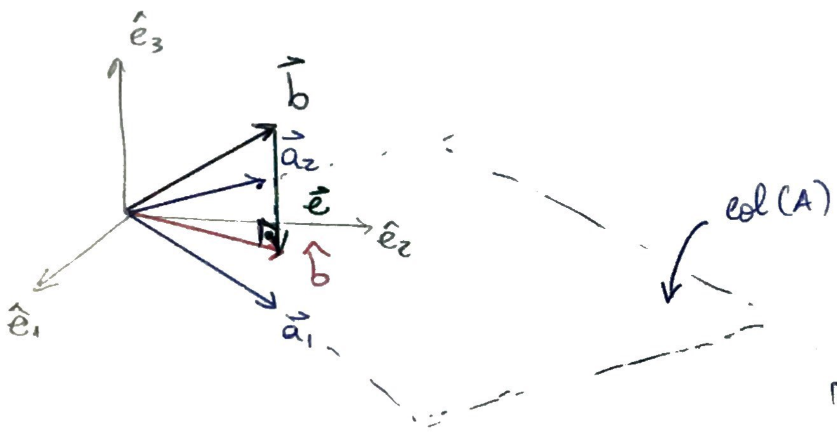
$$A\vec{x} = \vec{b}$$

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

- se $\vec{b} \in col(A)$ $\begin{cases} 1 \text{ solução} \\ \infty \text{ soluções} \end{cases}$

- se $\vec{b} \notin col(A)$: NÃO HÁ SOLUÇÃO → solução mais próxima: PROJEÇÃO

ex) $\mathbb{R}^3$



$\hat{b}$ → projeção de $\vec{b}$ em col(A)

$\vec{e} = \vec{b} - \hat{b}$ (erro)

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

↳ melhor solução aproximada

por definição: $\vec{e} \perp col(A) \Rightarrow A^T\vec{e} = \vec{0}$

$$A^T\vec{e} = \begin{bmatrix} - \vec{a}_1 - \\ - \vec{a}_2 - \\ \vdots \\ - \vec{a}_n - \end{bmatrix} \begin{bmatrix} | \\ \vec{e} \\ | \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

se $\vec{e} \perp \vec{a}_i$:

$\vec{a}_i^T \vec{e} = 0$

$\|\vec{a}_i\| \, \|\vec{e}\| \, \underbrace{\cos\theta} = 0$

$\theta = 90°$, $\cos 90° = 0$

$$A^T\vec{e} = \vec{0}$$
$$A^T(\vec{b} - \hat{b}) = \vec{0}$$
$$A^T(\vec{b} - A\hat{x}) = \vec{0}$$
$$A^T\vec{b} - A^TA\hat{x} = \vec{0}$$
$$\boxed{A^TA\hat{x} = A^T\vec{b}}$$

- sistema de eq. normais (estatística)
- solução MMQ para problemas superdeterminados ($m > n$)
- melhor solução de $A\vec{x} = \vec{b}$:
  - se 1 solução → exata
  - se não há solução → projeção

nome MMQ: $\min \| A\hat{x} - \vec{b} \|_2 = \min \| \vec{e} \|_2 \rightarrow$ projeção

Teorema: $\underbrace{A^T A}_{B_{n\times n}} \hat{x} = \underbrace{A^T \vec{b}}_{\vec{d}} \rightarrow B\hat{x} = \vec{d}$

$B_{n\times n}$ simétrica

↳ quadrada e sempre tem solução.

- se $K = n$: $(A^T A)$ é não-singular: $\hat{x} = (A^T A)^{-1} A^T \vec{b}$ é solução única.
- se $K < n$: $(A^T A)$ é singular: infinitas soluções $\hat{x}$

* algoritmo: decomposição QR: $A^T A$ não-singular:

$$A = QR$$

$$A^T A \hat{x} = A^T \vec{b}$$

$$(QR)^T (QR)\hat{x} = (QR)^T \vec{b}$$

$$\cancel{R^T}\cancel{Q^T Q}R\hat{x} = \cancel{R^T}Q^T\vec{b}$$

$$\boxed{R\hat{x} = Q^T\vec{b}}$$

fácil

subst. inversa.

3 etapas:

- $A = QR$
- Q^T
- $R\hat{x} = Q^T\vec{b}$

* se $A^T A$ singular: infinitas soluções: $\left\{\begin{array}{l} \text{sempre que } m < n \text{, pois} \\ K < n. \quad [\quad] \end{array}\right.$

→ critério adicional: $\min \| \hat{x} \|_2 \Rightarrow$ decomposição SVD

$$A_{m\times n} = U_{m\times m}\, \Sigma_{m\times n}\, V^T_{n\times n}$$

$\left\{\begin{array}{l} U, V \rightarrow \text{ortogonais} \\ \Sigma \rightarrow \text{diagonal} \end{array}\right.$

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \sigma_K & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}$$

$K \rightarrow$ posto de A.

MMQ:

$$\min \| A\hat{x} - \vec{b} \|_2 =$$

$$= \min \| U \Sigma \underbrace{V^T \hat{x}}_{\vec{y}} - \vec{b} \|_2 =$$

$$= \min \| U \Sigma \vec{y} - \vec{b} \|_2 =$$

$$= \min \| \Sigma \vec{y} - \underbrace{U^T \vec{b}}_{\vec{b}'} \|_2 =$$

$$= \min \| \Sigma \vec{y} - \vec{b}' \|_2 \quad \begin{cases} y_i = \dfrac{b_i'}{\sigma_i}, & i = 1, \dots, k \\ y_i = 0, & i = k+1, \dots, n \end{cases}$$

↳ escolha

as infinitas soluções surgem de $y_i = ?$, $i = k+1, \dots, n$
(todas tem o mesmo erro mínimo)

$$\hat{x} = V\vec{y} = V\Sigma^+ \vec{b}' = V\Sigma^+ U^T \vec{b}$$

onde

$$\Sigma^+ = \begin{bmatrix} 1/\sigma_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1/\sigma_2 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 1/\sigma_k & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}_{n \times m}$$

pseudo-inversa de $\Sigma_{m \times n}$

Pseudo inversa de Moore-Penrose: $\boxed{A^+ = V\Sigma^+ U^T}$

- dado $A\vec{x} = \vec{b}$, $\hat{x} = A^+\vec{b}$ é a solução de MMQ com $\|\hat{x}\|_2$ mínimo.

- válido para qualquer sistema $A\vec{x} = \vec{b}$

  * se solução única: $A^+ = A^{-1}$, $\hat{x} = \vec{x} = A^{-1}\vec{b}$
  * se não há solução: $\hat{x} = A^+\vec{b} \rightarrow$ solução MMQ ($\min \|A\hat{x} - \vec{b}\|_2$)
  * se infinitas soluções: $\min \|A\hat{x} - \vec{b}\|_2$ e $\min \|\hat{x}\|_2$.

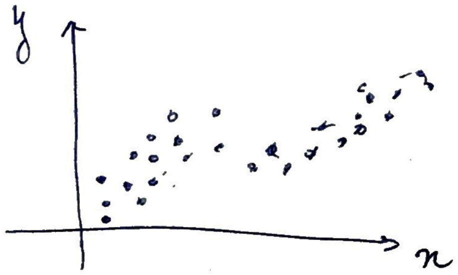
- generalização da matriz inversa.
- existe e é única $\forall$ matriz $A_{m \times n}$ real ou complexa
- se $\det(A) \neq 0$, $A^+ = A^{-1}$
- Propriedades:
  - $AA^+A = A$
  - $A^+AA^+ = A^+$
  - se col(A) são L.I.: $A^+A = I$
  - se linhas(A) são L.I: $AA^+ = I$
  - $(A^+)^+ = A$

- válido p/ qualquer sistema $A\vec{x} = \vec{b}$
  - se solução única: $A^+ = A^{-1}$ e $\hat{x} = \vec{x} = A^{-1}\vec{b}$
  - se não há solução: $\hat{x} = A^+\vec{b}$ → solução MMQ ($\min \| A\hat{x} - \vec{b} \|_2$)
  - se infinitas soluções: solução que $\min \| \hat{x} \|_2$ e $\min \| A\hat{x} - \vec{b} \|_2$
- matriz "mais próxima" da inversa (melhor aproximação).

---

* Aplicação 1) aproximação de dados por um polinômio de grau p. Quarteroni: 3.6
Franco: 8.2.2

conjunto de pontos: (x_i, y_i), $i = 1, \ldots, m$



$$y = \underbrace{a_0 + a_1 x + a_2 x^2 + \ldots + a_p x^p}_{\text{polinômio de grau } p}$$

$$\underbrace{\begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^p \\ 1 & x_2 & x_2^2 & \cdots & x_2^p \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^p \end{bmatrix}}_{X_{m\times(p+1)}} \underbrace{\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_p \end{bmatrix}}_{\vec{a}_{(p+1)\times 1}} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}}_{\vec{y}_{m\times 1}}$$

↳ matriz de Vandermonde

- em geral: $m \gg p$. → sistema superdeterminado
  com $K = p$ → MMQ tem solução única.

solução: $\vec{a} = (X^T X)^{-1} X^T \vec{y}$

ou

$$\vec{a} = X^+ \vec{y} \quad , \quad \begin{cases} X^+ = V \Sigma^+ U^T \\ X = U \Sigma V^T \end{cases}$$

* Aplicação 2) aproximação de dados por função trigonométrica

$$y = \frac{a_0}{2} + \sum_{k=1}^{p} \left(a_k \cos(kx) + b_k \operatorname{sen}(kx) \right)$$

ex) $p = 2$

$$y = \frac{a_0}{2} + a_1 \cos(x) + b_1 \operatorname{sen}(x) + a_2 \cos(2x) + b_2 \operatorname{sen}(2x)$$

$$\underbrace{\begin{bmatrix} 1/2 & \cos(x_1) & \operatorname{sen}(x_1) & \cos(2x_1) & \operatorname{sen}(2x_1) \\ 1/2 & \cos(x_2) & \operatorname{sen}(x_2) & \cos(2x_2) & \operatorname{sen}(2x_2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1/2 & \cos(x_m) & \operatorname{sen}(x_m) & \cos(2x_m) & \operatorname{sen}(2x_m) \end{bmatrix}}_{X_{m \times (2p+1)}} \underbrace{\begin{bmatrix} a_0 \\ a_1 \\ b_1 \\ a_2 \\ b_2 \end{bmatrix}}_{\vec{a}_{(2p+1) \times 1}} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}}_{\vec{y}_{m \times 1}}$$

* Aplicação 3) $y = \sum_{j=0}^{p} a_j \psi_j(x)$ (similar)

* Aplicação 4: problemas subdeterminados $(m < n)$

"matriz larga"

$m = 3$ indivíduos

$n = 5$ marcadores genéticos (# alelos de interesse)

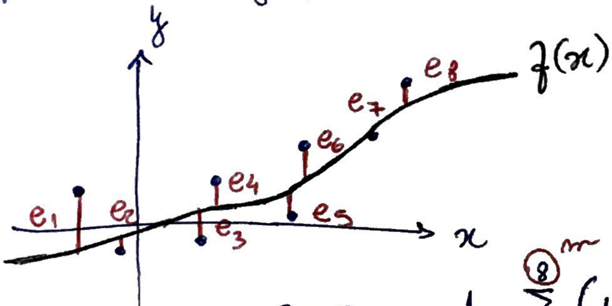
$$A = \begin{bmatrix} 1 & 0 & 0 & 2 & 0 \\ 0 & 1 & 1 & 1 & 2 \\ 1 & 0 & 2 & 1 & 1 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 1,3 \\ 3,3 \\ 9,7 \end{bmatrix} \rightarrow \text{tamanho do tumor}$$

$$\vec{x} = A^{+}\vec{b} = \begin{bmatrix} 1,7 \\ -1,1 \\ 3,9 \\ -0,2 \\ 0,3 \end{bmatrix}$$

---

obs: aproximação de dados por funções: $X\vec{a} = \vec{y}$

$m = 8$ pontos: (x_i, y_i), $i = 1, 8$



$$\vec{e} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ \vdots \\ e_8 \end{bmatrix} \rightarrow \text{vetor erro}$$

(mesmo vetor erro da projeção $\hat{b}$)

$$EMQ = \frac{1}{\underset{m}{8}} \sum_{i=1}^{\overset{m}{8}} (y_i - f(x_i))^2$$

$$\min \underbrace{(EMQ)}_{\text{erro médio quadrático}} \iff \min \|\vec{e}\|_2 = \underbrace{\min \sqrt{e_1^2 + e_2^2 + \dots + e_8^2}}_{\text{mínimos quadrados}}$$

- se pontos na curva: $\vec{e} = \vec{0}$: $\vec{y} \in col(X) \rightarrow$ projeção de $\vec{y}$ é o próprio $\vec{y}$, solução exata.
- se pontos fora da curva: $\vec{e} \neq \vec{0}$: $\vec{y} \notin col(X) \rightarrow$ método dos mínimos quadrados calcula projeção $\hat{y}$.