⑦ Aproximação de funções: método dos mínimos quadrados

Quarteroni: 5.7 ⊛

Franco: cap. 8 → abordagem diferente.

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

- posto (A) = K = nº de linhas L.I. de A = nº colunas L.I. de A

  ↳ dimensão do espaço coluna = dimensão do espaço linha = dimensão do espaço coluna de A^T.

  ↳ posto (A) = posto (A^T).

  ↳ nº de linhas não-nulas da forma escalonada de A.

  ↳ $K \leq m$ e $K \leq n$.

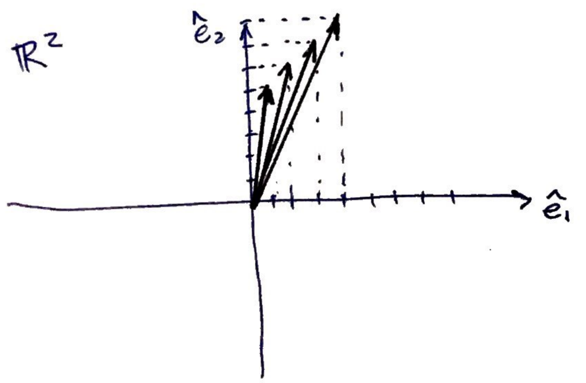
  ↳ se $m > n$, $k = n$ → posto completo

  se $m < n$, $K = m$ → posto completo

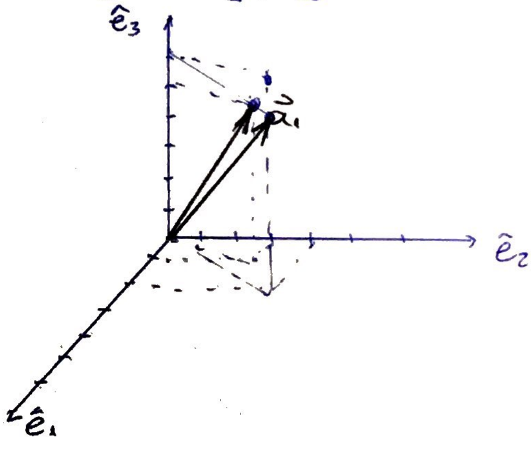
- se $m > n$ → sistema superdeterminado → mais eq. que inc.

  se $m < n$ → " subdeterminado → menos eq. que inc.

$$A = \begin{array}{c} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}_{2\times 4} \\ \begin{array}{cccc} \vec{a}_1 & \vec{a}_2 & \vec{a}_3 & \vec{a}_4 \end{array} \end{array}$$



$$A = \begin{array}{c} \begin{array}{cc} \vec{a}_1 & \vec{a}_2 \end{array} \\ \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}_{3\times 2} \end{array}$$



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

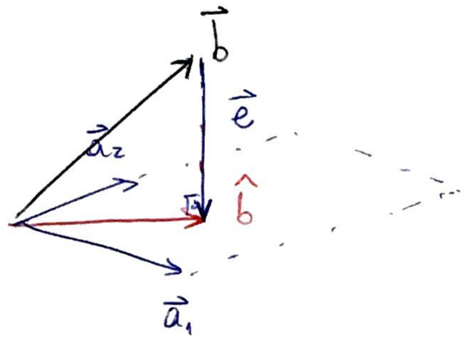
• se $\vec{b} \in \underbrace{col(A)}$ (espaço coluna de A): $\begin{cases} 1 \text{ solução} \\ \infty \text{ soluções} \end{cases}$

espaço formado pela comb. linear dos vetores coluna de A.

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

ex) $\mathbb{R}^3$

$A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 \end{bmatrix}_{3\times 2}$



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

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

$\boxed{A\hat{x} = \hat{b}}$ → melhor solução.

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

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

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

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

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

$\theta = 90°$, $\cos 90° = 0$.

$$A^T \vec{e} = \vec{0}$$

$$A^T(\vec{b} - \hat{b}) = \vec{0}$$

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

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

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

- sistema de equações 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.

Teorema:

$$\underbrace{A^T A}_{B_{n\times n} \text{ simétrica}} \hat{x} = \underbrace{A^T \vec{b}}_{\vec{d}}$$

sempre tem solução.

- se $k = n$: (A^TA) é não-singular

$\hat{x} = (A^TA)^{-1} A^T \vec{b}$ (única)

- se $k < n$: (A^TA) é singular: infinitas soluções $\hat{x}$.

3

* usando a decomposição QR: $A = QR$, A^TA não-singular.

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

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

$$R^T \underbrace{Q^T Q}_{I} R \hat{x} = R^T Q^T \vec{b}$$

I porque Q é ortogonal !

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

fácil de calcular

$$\boxed{\hat{x} = R^{-1} Q^T \vec{b}} \quad \text{ou} \quad \boxed{R \hat{x} = Q^T \vec{b}}$$

substituição inversa

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

* se A^TA singular: infinitas soluções: sempre que $m < n$, pois $k < n$.

- $A^T A \hat{x} = A^T \vec{b}$ → infinitas soluções $\hat{x}$ → critério adicional.

- escolhemos $\min \|\hat{x}\|_2$ ⇒ decomposição SVD.

$$\underset{m\times n}{A} = \underset{m\times m}{U}\ \underset{m\times n}{\Sigma}\ \underset{n\times n}{V^T} \quad \begin{cases} U, V \to \text{ortogonais} \\ \Sigma \to \text{diagonal} \end{cases} \quad \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \sigma_k & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}_{m\times n}$$

$k \to$ posto.

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

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

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

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

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

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

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

$$\hat{x}_{n\times 1} = V_{n\times n}\, \vec{y}_{n\times 1}$$

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

$$\left\{ \underset{n\times m}{A^+} = V\Sigma^+ U^T \to \text{pseudo-inversa de } A_{m\times n} \text{ (de Moore-Penrose)} \right\}$$

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

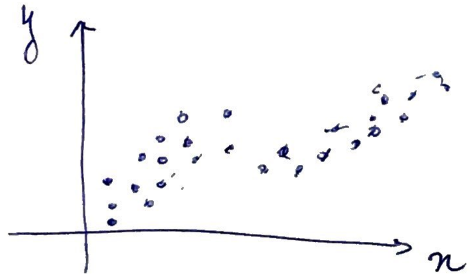
- válido p/ qualquer sistema $A\vec{x} = \vec{b}$
  - se solução única: $A^+ = A^{-1}$ e $\hat{x} = \vec{x} = A^{-1}\vec{b}$
  - se não há solução: $\hat{x} = A^+\vec{b} \rightarrow$ solução MMQ
    ($\min \| A\hat{x} - \vec{b} \|_2$)
  - se infinitas soluções: solução que $\min \| \hat{x} \|_2$
   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

---

* Aplicação 1) aproximação de dados por um polinômio de ordem p.

Quarteroni: 3.6
Franco: 8.2.2

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



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

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

- em geral: $m \gg p$. $\rightarrow$ sistema superdeterminado
  com $K = p \rightarrow$ 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)

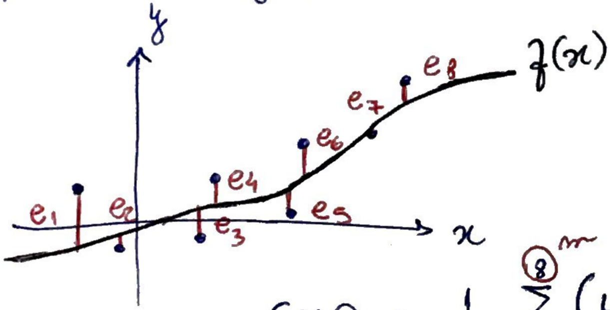
$$A = \begin{bmatrix} 1 & 0 & 0 & 2 & 0 \\ 0 & 1 & 1 & 1 & 2 \\ 1 & 0 & 2 & 1 & 1 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 1,3 \\ 3,3 \\ 9,7 \end{bmatrix} \rightarrow \text{tamanho do tumor}$$

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

——— // ———

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 $\hat{b}$)

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

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

- 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 $\hat{y}$.

obs: para aproximar uma função $f(x)$ qualquer por uma mais simples $\varphi(x)$:

opção 1: 1 seleciona m pontos (y_i, x_i), onde $y_i = f(x_i)$

2 calcula o MMQ nos pontos selecionados.

opção 2: livro Neide Franco, capítulo 8.