Cursul 5

II) Sisteme de ecuații liniare supradeterminate. Regresii

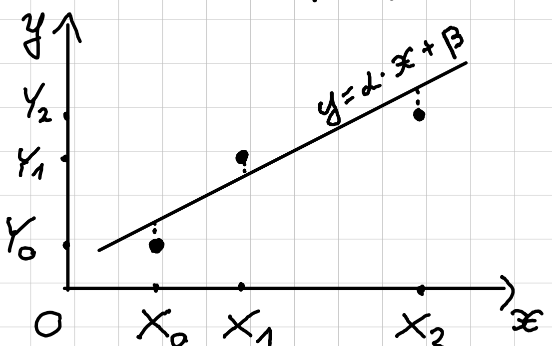
- Vrem să rezolvăm numeric sisteme de ecuații liniare supradeterminate, i.e.

$$A \cdot x = b, \quad (1)$$

unde $\begin{cases} A \in \mathcal{M}_{m,n}(\mathbb{R}), \ m > n, \ \text{rang}(A) = n \\ b \in \mathbb{R}^m. \end{cases}$

- Exemplu: Regresia liniară

Vrem să determinăm dreapta $y = \alpha \cdot x + \beta$ cea mai „apropiată" de punctele $\{(X_i, Y_i)\}_{i=\overline{0,2}}$



$$\begin{cases} \alpha \cdot X_0 + \beta = Y_0 \\ \alpha \cdot X_1 + \beta = Y_1 \\ \alpha \cdot X_2 + \beta = Y_2 \end{cases}$$

$$\Longleftrightarrow \underbrace{\begin{bmatrix} x_0 & 1 \\ x_1 & 1 \\ x_2 & 1 \end{bmatrix}}_{A} \cdot \underbrace{\begin{bmatrix} \alpha \\ \beta \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix}}_{b}$$

- Teorema Kronecker - Capelli ne spune că avem o soluție $\Leftrightarrow \operatorname{rang}(A) = \operatorname{rang}(A|b)$
- Cum $\operatorname{rang}(A) = 2 \Rightarrow$ sistemul are soluție $\Leftrightarrow \operatorname{rang}(A|b) = 2$

$$\Longleftrightarrow \begin{vmatrix} x_0 & 1 & y_0 \\ x_1 & 1 & y_1 \\ x_2 & 1 & y_2 \end{vmatrix} = 0 \Longleftrightarrow (x_i, y_i)_{i=\overline{0,2}} \text{ sunt coliniare.}$$

Q: Ce ne facem când punctele nu sunt coliniare?

Vom construi o dreaptă astfel încât aceasta să fie cât mai apropiată de punctele $(x_i, y_i)_{i=\overline{0,2}}$

- Mai exact, caut $\alpha, \beta \in \mathbb{R}$ a. î.

$$\sum_{i=0}^{2} (\alpha X_i + \beta - Y_i)^2 \text{ să fie minim}$$

- Observație

$$\| A x - b \|_2^2 = \left\| \begin{pmatrix} x_0 & 1 \\ x_1 & 1 \\ x_2 & 1 \end{pmatrix} \cdot \begin{pmatrix} \alpha \\ \beta \end{pmatrix} - \begin{pmatrix} Y_0 \\ Y_1 \\ Y_2 \end{pmatrix} \right\|_2^2$$

$$= \left\| \begin{pmatrix} \alpha X_0 + \beta - Y_0 \\ \alpha x_1 + \beta - Y_1 \\ \alpha X_2 + \beta - Y_2 \end{pmatrix} \right\|_2^2 = \sum_{i=0}^{2} (\alpha X_i + \beta - Y_i)^2$$

- Definiție

i) Numim problema celor mai mici pătrate (PCMMP) următoarea problemă

$$\hat{x} := \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \| A x - b \|_2^2 \quad (2)$$

$$\Leftrightarrow \| A \hat{x} - b \|_2^2 \leq \| A x - b \|_2^2, \ \forall x \in \mathbb{R}^n$$

ii) Dacă $\exists \hat{x} \in \mathbb{R}^n$ soluție pentru (2), aceasta se numește soluția PCMMP a sistemului supradeterminat de ecuații liniare (1).

• În continuare, vom răspunde la următoarele întrebări:

Q_1: Există o soluție a PCMMP?

Q_2: Dacă da, este unică?

Q_3: Algoritmi de determinare a soluției?

• Teoremă

Considerăm sistemul supradeterminat (1)
Dacă A este inversabilă la stânga, i.e. $\exists\, A^{-1} \in \mathcal{M}_{n,m}(\mathbb{R})$ a.î. $A^{-1}A = I_n$, atunci PCMMP (2) are o <u>unică</u> soluție dată de rezolvarea <u>sistemului de ecuații normale</u> asociat

$$A^T A \hat{x} = A^T b \quad (3)$$

Demonstrație:

Pasul I: Arăt că $A^T A \in \mathcal{M}_n(\mathbb{R})$ este simetrică și pozitiv definită (SPD)

Cum $(A^T A)^T = A^T A$, $A^T A$ este simetrică

Presupunem prin absurd că $A^T A$ nu este pozitiv definită $\Rightarrow \exists x \in \mathbb{R}^n \setminus \{0\}$ a.î.

$x^T (A^T A) x = 0 \Leftrightarrow (x^T A^T)(Ax) = 0 \Leftrightarrow$

$(Ax)^T (Ax) = 0 \Leftrightarrow \|Ax\|_2^2 = 0 \Leftrightarrow$

$Ax = 0 \in \mathbb{R}^m$

Înmulțim la stânga cu $A^{-1} \in \mathcal{M}_{n,m}(\mathbb{R})$:

$A^{-1} A x = 0 \in \mathbb{R}^n \Leftrightarrow x = 0$ ↯

$\Rightarrow A^T A$ este o matrice SPD

Pasul $\underline{II}$: Arăt că soluția $\hat{x} \in \mathbb{R}^n$ a sistemului $A^T A x = A^T b$ este și unica soluție PCMMP, i.e.

$$\|A\hat{x} - b\|_2^2 < \|Ax - b\|_2^2 \quad \forall x \neq \hat{x}$$

Fie $x \in \mathbb{R}^n$ a.î. $x \neq \hat{x}$.

$\|Ax - b\|_2^2 = \|(Ax - A\hat{x}) + (A\hat{x} - b)\|_2^2$

$= [A(x - \hat{x}) + (A\hat{x} - b)]^T \cdot [A(x - \hat{x}) + (A\hat{x} - b)]$

$$= \|A(x-\hat{x})\|_2^2 + 2[A(x-\hat{x})]^T(A\hat{x}-b) + \|A\hat{x}-b\|_2^2$$

Observăm că:

i) $\|A(x-\hat{x})\|_2^2 = (x-\hat{x})^T A^T A (x-\hat{x}) > 0$

deoarece $A^T A$ este pozitiv definită

ii) $[A(x-\hat{x})]^T (A\hat{x}-b) =$

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

$= (x-\hat{x})^T (\underbrace{A^T A\hat{x} - A^T b}_{0 \in \mathbb{R}^n}) = 0$

Din i) și ii) obținem că

$\|Ax-b\|_2^2 > \|A\hat{x}-b\|_2^2 \quad \forall x \neq \hat{x}$ $\square$

- Observație:

Cum $A^T A \in \mathcal{M}_n(\mathbb{R})$ este simetrică și pozitiv definită, putem aplica factorizarea Cholesky pentru a rezolva sistemul

$$A^T A x = A^T b$$

și a obține soluția PCMMP a $Ax = b$.

• Aplicație: Dreapta de regresie

Coeficienții dreptei de regresie, $\hat{x} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$, celei mai apropiată de punctele $(x_i, y_i)_{i=\overline{0,m-1}}$ se determină rezolvând sistemul

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

$$\Leftrightarrow \begin{bmatrix} x_0 & x_1 & \dots & x_{m-1} \\ 1 & 1 & \dots & 1 \end{bmatrix} \cdot \begin{bmatrix} x_0 & 1 \\ \vdots & \vdots \\ x_{m-1} & 1 \end{bmatrix} \cdot \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} x_0 & \dots & x_{m-1} \\ 1 & \dots & 1 \end{bmatrix} \cdot \begin{bmatrix} y_0 \\ \vdots \\ y_{m-1} \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} \sum_{i=0}^{m-1} x_i^2 & \sum_{i=0}^{m-1} x_i \\ \sum_{i=0}^{m-1} x_i & m \end{bmatrix} \cdot \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^{m-1} x_i y_i \\ \sum_{i=0}^{m-1} y_i \end{bmatrix}$$