Cursul 6

- Strategie generală de rezolvat sisteme de ecuații liniare supradeterminate.

$$A \cdot x = b, \quad A \in \mathcal{M}_{m,n}(\mathbb{R}),\ m > n,\ \text{rang}(A) = n$$

$$\Big\downarrow \text{PCMMP}$$

$$A^T A \cdot x = A^T \cdot b, \quad A^T A \in \mathcal{M}_n(\mathbb{R}) \ \text{SPD}$$

$$\Big\downarrow \text{Gauss / Factorizări}$$

$$U \cdot x = b, \quad U \in \mathcal{M}_n(\mathbb{R}) \ \text{superior triunghiulară}$$

$$\Big\downarrow \text{Substituția descendentă}$$

$\hat{x} \in \mathbb{R}^n$, unica soluție PCMMP

$$\left(\|A\hat{x} - b\|_2 \leq \|Ax - b\|_2 \quad \forall x \neq \hat{x} \right)$$

- Aplicație: Dreapta de regresie $y = \alpha x + \beta$ cea mai apropiată de punctele $(x_i, y_i)_{i=\overline{0,m-1}}$ se determină rezolvând (Vezi cursul 5)

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

Temă bonus: Verificați că obțineți

$$\alpha = \frac{\sum_{i=0}^{m-1} \Delta X_i \cdot \Delta Y_i}{\sum_{i=0}^{m-1} \Delta X_i^2} \text{ și } \beta = \overline{Y} - \alpha \overline{X}$$

unde $\overline{X} = \frac{1}{m} \sum_{i=0}^{m-1} X_i$ media valorilor lui X_i

$\overline{Y} = \frac{1}{m} \sum_{i=0}^{m-1} Y_i$ media valorilor lui Y_i

$\Delta X_i = X_i - \overline{X}$ abaterea de la medie a lui X_i

$\Delta Y_i = Y_i - \overline{Y}$ abaterea de la medie a lui Y_i

• Generalizare: Polinomul de regresie

Coeficienții polinomului de regresie

$$P_n(x) = \sum_{i=0}^{n} c_i x^i \in P_n$$

cel mai apropiat de punctele $\{(x_i, y_i)\}_{i=\overline{0,m-1}}$

se determină găsind soluția PCMMP

a sistemului de ecuații supradeterminat

$$\underbrace{\begin{bmatrix} x_0^n & x_0^{n-1} & & x_0 & 1 \\ x_1^n & x_1^{n-1} & \cdots & x_1 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ x_{m-1}^n & x_{m-1}^{n-1} & & x_{m-1} & 1 \end{bmatrix}}_{A \in \mathcal{M}_{m,n+1}(\mathbb{R})} \cdot \underbrace{\begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix}}_{x \in \mathbb{R}^{n+1}} = \underbrace{\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{m-1} \end{bmatrix}}_{b \in \mathbb{R}^m}$$

i.e. rezolvând $A^T A x = A^T b$.

• Observație (Interpretarea geometrică)

Pentru $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $m > n$, $\text{rang}(A) = n$ și $b \in \mathbb{R}^m$ corespunzătoare PCMMP:

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

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

$$\Leftrightarrow \hat{x} = \underset{x \in \mathbb{R}^n}{\text{argmin}} \|Ax - b\|_2^2$$

$$= \underset{x \in \mathbb{R}^n}{\text{argmin}} \left[d(b, Ax) \right]^2$$

• Notăm cu $a_0, a_1 \ldots a_{n-1} \in \mathbb{R}^m$ coloanele lui A:

$$A = [a_0, a_1 \ldots a_{n-1}].$$

$$Ax = \begin{bmatrix} a_{00} & a_{01} & \ldots & a_{0n-1} \\ a_{10} & a_{11} & \ldots & a_{1n-1} \\ \vdots & \vdots & \ldots & \vdots \\ \underbrace{a_{m-1 0}}_{a_0} & \underbrace{a_{m-1 1}}_{a_1} & \ldots & \underbrace{a_{m-1 n-1}}_{a_{n-1}} \end{bmatrix} \cdot \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{bmatrix} =$$

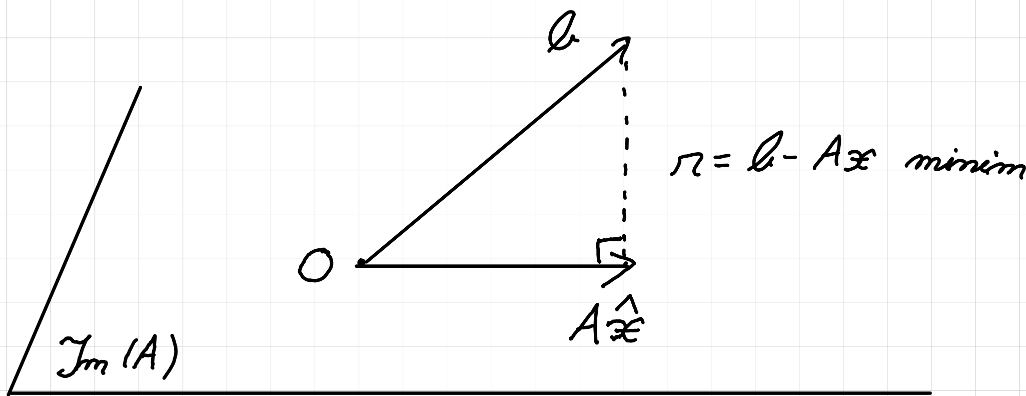
$$= a_0 x_0 + a_1 x_1 + \ldots + a_{n-1} x_{n-1} \in \text{Im}(A)$$

$$\Rightarrow \text{Im}(A) = \langle a_0, a_1 \ldots a_{n-1} \rangle \subsetneq \mathbb{R}^m \text{ sp. vect.}$$

- Definiție

Numim vector eroare reziduală

$$r := b - Ax \in \mathbb{R}^m$$



$\Rightarrow A\hat{x} = Pr_{Im(A)}\, b$

(Proiecția lui b pe $Im(A)$).

Distanța minimă de la b la $Im(A)$, vectorul eroare reziduală $r = b - A\hat{x}$, e perpendiculara din b pe $Im(A)$, i.e.

$$A^T r = 0$$

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

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

- Exemplu problematic de rezolvat cu sistemul de ecuații normale

$$A = \begin{bmatrix} 1 & 1 \\ \varepsilon & 0 \\ 0 & \varepsilon \end{bmatrix}, \text{ unde } \varepsilon \approx 10^{-8}$$

$$A^T A = \begin{bmatrix} 1+\varepsilon^2 & 1 \\ 1 & 1+\varepsilon^2 \end{bmatrix}, \quad \varepsilon^2 \approx 10^{-16} < \varepsilon_M$$

Precizia mașinii

$$fl(A^T A) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \text{ matrice neinversabilă!}$$

- Soluție: Metoda sistemului augmentat

Considerăm vectorul eroare reziduală $r = b - Ax$ necunoscut și obținem sistemul augmentat

$$\begin{cases} r + Ax = b \in \mathbb{R}^m \\ A^T r = 0 \in \mathbb{R}^n \end{cases} \Leftrightarrow \left[\begin{array}{c|c} I_m & A \\ \hline A^T & 0 \end{array}\right] \cdot \left[\begin{array}{c} r \\ \hline x \end{array}\right] = \left[\begin{array}{c} b \\ \hline 0 \end{array}\right]$$

a cărui rezolvare ne dă atât vectorul eroare reziduală cât și soluția PCMMP

- Deși matricea asociată sistemului augmentată nu e pozitiv definită și mai mare decât $A^T A$, avem libertatea de a scala sistemul augmentat în felul următor:

$$\left[\begin{array}{c|c} \alpha I_m & A \\ \hline A^T & 0 \end{array}\right] \cdot \left[\begin{array}{c} \frac{r}{\alpha} \\ \hline x \end{array}\right] = \left[\begin{array}{c} b \\ \hline 0 \end{array}\right]$$

pentru a evita obținerea de pivoți prea mici într-un algoritm de eliminare Gauss sau de pivotare.

O scalare bună, în general, e dată de

$$\alpha = \frac{\max |a_{ij}|}{1000}$$

- Revenind la exemplul de mai devreme, cu

$$A = \begin{bmatrix} 1 & 1 \\ \varepsilon & 0 \\ 0 & \varepsilon \end{bmatrix}, \quad \varepsilon \approx 10^{-8} \Rightarrow \alpha = \frac{1}{1000}$$

Sistemul augmentat:

$$\begin{bmatrix} 10^{-3} & 0 & 0 & 1 & 1 \\ 0 & 10^{-3} & 0 & \varepsilon & 0 \\ 0 & 0 & 10^{-3} & 0 & \varepsilon \\ 1 & \varepsilon & 0 & 0 & 0 \\ 1 & 0 & \varepsilon & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 10^3 r_0 \\ 10^3 r_1 \\ 10^3 r_2 \\ x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ 0 \\ 0 \end{bmatrix}$$

Seminar / Laborator: Matricea nu va deveni neinversabilă pe parcursul unei eliminări Gauss cu pivotare sau a unei factorizări cu pivotare!