CURS#6

10. Sisteme de ecuaţii liniare supradeterminate (problema celor mai mici pătrate):
    - (i) definiţia soluţiei în sensul celor mai mici pătrate;
    - (ii) sistemul de ecuaţii normale asociat;
    - (iii) problema de minimizare pătratică;
    - (iv) rezolvare folosind factorizarea Cholesky; algoritm;
    - (v) aplicaţie: problema regresiei liniare;
    - (vi) aplicaţie: problema regresiei - cazul general;
    - (vii) metoda sistemului augmentat.

PROBLEME

1) Fie $\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2 \dots \mathbf{a}_n] \in \mathscr{M}_{m,n}(\mathbb{R})$, $m \geq n$, unde $\mathbf{a}_k = \left(a_{ik}\right)_{i=\overline{1,m}} \in \mathbb{R}^m$, $k = \overline{1,n}$.

   Atunci sunt echivalente următoarele afirmaţii:
   - (i) $\mathbf{A}$ este inversabilă la stânga;
   - (ii) $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \in \mathbb{R}^m$ sunt liniar independenţi;
   - (iii) $\mathbf{A}^\mathsf{T}\mathbf{A} \in \mathscr{M}_n(\mathbb{R})$ este inversabilă;
   - (iv) $\mathrm{rang}(\mathbf{A}) = n$;
   - (v) $\mathbf{A}^\mathsf{T}\mathbf{A} \in \mathscr{M}_n(\mathbb{R})$ este SPD.

2) Fie setul de date $D \coloneqq \left\{\left(x_i, y_i\right) \in \mathbb{R}^2 \;\middle|\; i = \overline{1,m}\right\} \subset \mathbb{R}^2$, unde $m \geq 2$.
   - (i) Arătaţi că *dreapta de regresie* asociată setului de date D este dată de ecuaţia
   $$y = \overline{y} + c_1\left(x - \overline{x}\right) , \tag{1}$$
   unde
   $$\overline{x} \coloneqq \frac{1}{m}\sum_{i=1}^{m} x_i \, , \tag{2a}$$
   $$\overline{y} \coloneqq \frac{1}{m}\sum_{i=1}^{m} y_i \, , \tag{2b}$$
   $$c_1 \coloneqq \left(\sum_{i=1}^{m}\left(x_i - \overline{x}\right) y_i\right) \Bigg/ \left(\sum_{i=1}^{m}\left(x_i - \overline{x}\right) x_i\right) . \tag{2c}$$
   - (ii) Determinaţi componentele vectorului eroare reziduală $\mathbf{r} = \left(r_i\right)_{i=\overline{1,m}} \in \mathbb{R}^m$, precum şi norma sa euclidiană $\|\mathbf{r}\|_2$.

SISTEME DE ECUAȚII LINIARE SUPRADETERMINATE (SUPRAABUNDENTE)

Considerăm sisteme de ecuații liniare de forma:

$$A\underline{x} = \underline{b} \qquad (1)$$

unde

$A \in \mathcal{M}_{m,n}(\mathbb{R})$, $m > n$

$\underline{b} \in \mathbb{R}^m$.

Echivalent, sistemul (1) poate scrie:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases} \qquad (2)$$

Sistemele de tipul (1) sau (2) sn sisteme de ecuații liniare <u>supradeterminate (supraabundente)</u>.

MOTIVAȚIE:

Considerăm o problemă <u>modelată</u> prin:

$$\left.\begin{aligned} f(\cdot,\cdot) &: \mathbb{R}^n \times \mathbb{R} \longrightarrow \mathbb{R} \\ y(\cdot) &: \mathbb{R} \longrightarrow \mathbb{R} \end{aligned}\right\} \ \underline{\text{cunoscute}}$$

$$\boxed{f(\underline{x}, t) = y(t)} \quad (*)$$

Vrem să determinăm <u>parametrul necunoscut</u> $\underline{x} \in \mathbb{R}^n$ în urma a $m > n$ măsurători pt $t_i \in \mathbb{R}$, $i = \overline{1,m}$.

Mai mult, presupunem că f depinde liniar de $\underline{x} = (x_1\ x_2 \dots x_n)^T \in \mathbb{R}^n$, i.e.

$$\boxed{f(\underline{x}, t) = \sum_{j=1}^{n} \phi_j(t)\, x_j} \quad (**)$$

$(**)$ în $(*)$ pt $t_i \in \mathbb{R}$, $i = \overline{1,m}$

$$\sum_{j=1}^{n} \underbrace{\phi_j(t_i)}_{=:a_{ij}}\, x_j = \underbrace{y(t_i)}_{=:b_i}, \quad i = \overline{1,m}$$

<u>OBS</u>: <u>Problema majoră</u>: $m > n$!

Cum definim o "soluție" a sistemului supradeterminat (1) sau (2)?

1. PROBLEMA CELOR MAI MICI PĂTRATE (PCMMP)

<u>DEFINIȚIE</u>: (soluția în sensul celor mai mici pătrate):

Fie $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $m > n$ și $\underline{b} \in \mathbb{R}^m$ asociate sistemului supraabundent (1) sau (2).

(i) Sn <u>problema celor mai mici pătrate (PCMMP)</u> următoarea problemă de minimizare:

$$\boxed{\begin{array}{l} \text{Determină } \underline{\hat{x}} \in \mathbb{R}^n: \\ \|A\underline{\hat{x}} - \underline{b}\|_2 \leq \|A\underline{x} - \underline{b}\|_2, \ \forall \underline{x} \in \mathbb{R}^n \end{array}} \tag{3a}$$

sau

$$\boxed{\underline{\hat{x}} := \arg\min_{\underline{x} \in \mathbb{R}^n} \|A\underline{x} - \underline{b}\|_2} \tag{3b}$$

(ii) Dacă există $\underline{x} \in \mathbb{R}^n$ soluție a PCMMP, atunci $\underline{x} \in \mathbb{R}^n$ în <u>soluția problemei celor mai mici pătrate</u> a sistemului supradeterminat de ecuații liniare (1) sau (2).

<u>OBS</u>: Se pun următoarele <u>probleme</u>:

(i) <u>Există</u> o soluție a PCMMP?

(ii) Dacă da, este <u>unică</u>?

(iii) <u>Algoritm</u> de determinare a soluției?

<u>TEOREMA #1</u> (existența soluției PCMMP):

Fie $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $m > n$ și $\underline{b} \in \mathbb{R}^m$ corespunzători PCMMP (3a)/(3b). Echivalente:

(i) $\hat{\underline{x}} \in \mathbb{R}^n$ <u>soluție a PCMMP</u> (3a);

(ii) $\hat{\underline{x}} \in \mathbb{R}^n$ soluție a <u>sistemului de ecuații normale</u>

$$\boxed{(A^T A)\,\hat{\underline{x}} = A^T \underline{b}} \qquad (4)$$

(iii) vectorul eroare reziduală

$$\boxed{\underline{\hat{r}} := \underline{b} - A\underline{\hat{x}} \in \mathbb{R}^m} \qquad (5)$$

satisface condiția de ortogonalitate

$$\boxed{A^T \underline{\hat{r}} = \underline{0}_n} \qquad (6)$$

Dem:

(ii) $\Leftrightarrow$ (iii):

Din definiția lui $\underline{r}$ dată de (5).

(i) $\Rightarrow$ (iii):

Fie $\underline{\hat{x}} \in \mathbb{R}^n$ soluție a POMMP (3a)/(3b):

$$\| A\underline{\hat{x}} - \underline{b} \|_2 \le \| A\underline{x} - \underline{b} \|_2, \ \forall \underline{x} \in \mathbb{R}^n$$

Presupunem că $A^T \underline{\hat{r}} = \underline{z} \in \mathbb{R}^n \setminus \{\underline{0}_n\}$.

Fie $\underline{x} := \underline{\hat{x}} + \varepsilon \underline{z} \in \mathbb{R}^n$, $\varepsilon \in \mathbb{R}$ $\Rightarrow$

$$\underline{r} := \underline{b} - A\underline{x} = \underline{b} - A(\underline{\hat{x}} + \varepsilon \underline{z})$$

$$= (\underline{b} - A\underline{\hat{x}}) - \varepsilon A\underline{z} = \underline{\hat{r}} - \varepsilon A\underline{z} \Rightarrow$$

$$\|\underline{\hat{r}}\|_2^2 = \underline{r}^T\underline{r} = (\underline{\hat{r}} - \varepsilon A\underline{z})^T (\underline{\hat{r}} - \varepsilon A\underline{z})$$

$$= \underline{\hat{r}}^T\underline{\hat{r}} - (\varepsilon A\underline{z})^T \underline{\hat{r}} - \underline{r}^T(\varepsilon A\underline{z}) + (\varepsilon A\underline{z})^T(\varepsilon A\underline{z})$$

$$= \|\hat{\underline{r}}\|_2^2 - 2\varepsilon \underline{z}^T \underbrace{A^T \hat{\underline{r}}}_{=\underline{z}} + \varepsilon^2 \|A\underline{z}\|_2^2$$

$$= \|\hat{\underline{r}}\|_2^2 - 2\varepsilon \|\underline{z}\|_2^2 + \varepsilon^2 \|A\underline{z}\|_2^2 \quad (*)$$

<u>Cazul I</u>: $\|A\underline{z}\|_2 = 0$

Relația (*) devine:

$$\|\underline{r}\|_2^2 = \|\hat{\underline{r}}\|_2^2 - 2\varepsilon \|\underline{z}\|_2^2 < \|\hat{\underline{r}}\|_2^2, \ \forall \varepsilon > 0$$

contradicție cu $\hat{\underline{x}} \in \mathbb{R}^n$ soluție a PCMMP (3a).

<u>Cazul II</u>: $\|A\underline{z}\|_2 > 0$

Alegem $\varepsilon \in \left(0, \frac{2\|\underline{z}\|_2^2}{\|A\underline{z}\|_2^2}\right)$ și obținem din (*):

$$\|\underline{r}\|_2^2 = \|\hat{\underline{r}}\|_2^2 + \varepsilon \underbrace{\left(\varepsilon \|A\underline{z}\|_2^2 - 2\|\underline{z}\|_2^2\right)}_{<0} < \|\hat{\underline{r}}\|_2^2$$

ceea ce contrazice, din nou, faptul că $\hat{\underline{x}} \in \mathbb{R}^n$ este o soluție a PCMMP (3a).

(iii) $\Longrightarrow$ (i):

Fie $\underline{\hat{x}} \in \mathbb{R}^n$ a.î. $A^T A \underline{\hat{x}} = A^T \underline{b}$, deci

$$A^T(\underline{b} - A\underline{\hat{x}}) = A^T \underline{\hat{r}} = \underline{0}_n$$

$\forall\, \underline{x} \in \mathbb{R}^n$,

$$\underline{r} := \underline{b} - A\underline{x} = (\underline{b} - A\underline{\hat{x}}) + (A\underline{\hat{x}} - A\underline{x})$$

$$= \underline{\hat{r}} + A(\underline{\hat{x}} - \underline{x}) \Longrightarrow$$

$\forall\, \underline{x} \in \mathbb{R}^n$,

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

$$= \underline{\hat{r}}^T\underline{\hat{r}} + (\underline{\hat{x}} - \underline{x})^T \underbrace{A^T\underline{\hat{r}}}_{=\underline{0}_n} + \underbrace{\underline{\hat{r}}^T A(\underline{\hat{x}} - \underline{x})}_{=[(\underline{\hat{x}} - \underline{x})^T \underbrace{A^T\underline{\hat{r}}}_{=\underline{0}_n}]^T}$$

$$[A(\underline{\hat{x}} - \underline{x})]^T [A(\underline{\hat{x}} - \underline{x})]$$

$$= \|\underline{\hat{r}}\|_2^2 + \|A(\underline{\hat{x}} - \underline{x})\|_2^2 \geq \|\underline{\hat{r}}\|_2^2$$

$\square$

<u>OBS</u>: Fie $A = [\underline{a}_1\ \underline{a}_2 \ldots \underline{a}_n] \in \mathcal{M}_{m,n}(\mathbb{R})$, $\underline{a}_j := (a_{ij})_{i=\overline{1,m}} \in \mathbb{R}^m$, $j = \overline{1,n}$. $\Longrightarrow$

$$\text{range}(A) := \{A\underline{x} \mid \underline{x} \in \mathbb{R}^n\}$$
$$= \text{span}\{\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_n\} \subset \mathbb{R}^m;$$

$\ker(A^T) := \{ \underline{y} \in \mathbb{R}^m \mid A^T \underline{y} = \underline{0}_n \} \subset \mathbb{R}^m$
complementul ortogonal în $\mathbb{R}^m$ al lui
$\mathrm{range}(A)$;

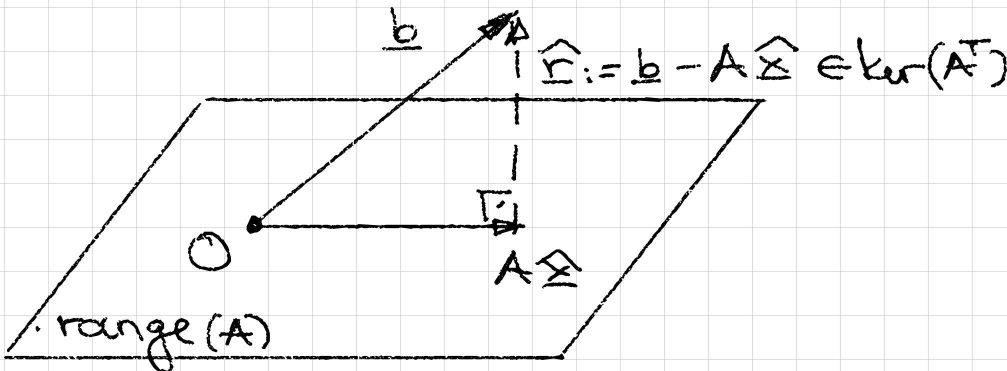
$\mathbb{R}^m = \mathrm{range}(A) \oplus \ker(A^T)$.

<u>**OBS**</u>: Cf Teoremei#1, dacă $\underline{\hat{x}} \in \mathbb{R}^n$ este o <u>soluție a PCMMP</u> (3a), atunci <u>vectorul eroare reziduală</u> $\underline{\hat{r}} := \underline{b} - A\underline{\hat{x}} \in \ker(A^T)$.

Prin urmare, orice soluție a PCMMP $\underline{\hat{x}} \in \mathbb{R}^n$ descompune $\underline{b} \in \mathbb{R}^m$, în mod unic, în două componente ortogonale:

$$\boxed{\begin{aligned} &\underline{b} = A\underline{\hat{x}} + \underline{\hat{r}}, \\ &\qquad A\underline{\hat{x}} \in \mathrm{range}(A) \\ &\qquad \underline{\hat{r}} \in \ker(A^T) \\ &\qquad (A\underline{\hat{x}})^T \underline{\hat{r}} = 0 \end{aligned}} \tag{7}$$

OBS: Interpretare geometrică



$$\boxed{\begin{aligned} A\hat{x} &= pr_{\text{range}(A)}\, b \\ &= pr_{\text{span}\{a_1, a_2, \dots, a_n\}}\, b \end{aligned}} \qquad (8)$$

OBS: Sistemul de ecuații normale (4) este consistent, i.e.

$$\boxed{\text{range}(A^T A) = \text{range}(A^T)} \qquad (9)$$

Dem:

„⊂": Fie $y \in \text{range}(A^T A) \Rightarrow$

$\exists\, x \in \mathbb{R}^n : y = A^T A x = A^T(Ax) \in \text{range}(A^T)$

"$\supseteq$": Fie $\underline{y} \in \text{range}(A^T) \Rightarrow$

$\exists\, \underline{z} \in \mathbb{R}^m: \quad A^T \underline{z} = \underline{y}$

Dar $\mathbb{R}^m = \text{range}(A) \oplus \ker(A^T) \Rightarrow$

$$\underline{z} = A\underline{x} + \underline{\tilde{z}}, \quad \underline{x} \in \mathbb{R}^n \quad \Rightarrow$$

$$\underline{\tilde{z}} \in \ker(A^T)$$

$$\Rightarrow \underline{y} = A^T \underline{z} = A^T (A\underline{x} + \underline{\tilde{z}})$$

$$= A^T A \underline{x} + \underbrace{A^T \underline{\tilde{z}}}_{= \underline{0}_n} = A^T A \underline{x} \in \text{range}(A^T A) \qquad \square$$

TEOREMA #2 (unicitatea soluției PCMMP):

Fie $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $m > n$ și $\underline{b} \in \mathbb{R}^m$ corespunzători PCMMP (3a) sau (3b).

Dacă $A \in \mathcal{M}_{m,n}(\mathbb{R})$ este inversabilă la stânga ($\exists\, B \in \mathcal{M}_{n,m}(\mathbb{R})$: $BA = I_n$), atunci PCMMP (3a) sau (3b) are soluție unică.

$$\boxed{\begin{array}{l} \underline{\hat{x}} = (A^T A)^{-1} A^T \underline{b} \\ \hline \underline{\hat{r}} := \underline{b} - A (A^T A)^{-1} A^T \underline{b} \end{array}} \qquad \begin{array}{r} (10) \\ (11) \end{array}$$

Dem:

- $A \in \mathcal{M}_{m,n}(\mathbb{R})$ inversabilă la stânga $\Leftrightarrow$

$\Leftrightarrow$ $A^T A \in \mathcal{M}_n(\mathbb{R})$ inversabilă

$\Leftrightarrow$ Sistemul de ecuații normale are <u>soluție unică</u> dată de

$$\underline{\hat{x}} := (A^T A)^{-1} A^T \underline{b} \Rightarrow$$

$$\underline{\hat{r}} := \underline{b} - A\underline{\hat{x}} = \underline{b} - A(A^T A)^{-1} A^T \underline{b}$$

$\square$

TEOREMA #3 (problema de minimizare pătratică):

Fie $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $m > n$ și $\underline{b} \in \mathbb{R}^m$ corespunzători PCMMP (3a) sau (3b).

Fie $\mathcal{J} : \mathbb{R}^n \to [0, \infty)$, $\mathcal{J}(\underline{x}) = \frac{1}{2} \| A\underline{x} - \underline{b} \|_2^2$.

Echivalente:

(i) $\underline{\hat{x}} \in \mathbb{R}^n$ <u>soluție a PCMMP</u> (3a);

(ii) $\underline{\hat{x}} \in \mathbb{R}^n$ <u>soluție a problemei de minimizare pătratică</u>

$$\boxed{\underline{\hat{x}} := \arg\min_{\underline{x} \in \mathbb{R}^n} \mathcal{J}(\underline{x})} \qquad (12)$$

Dem:

Fie $\underline{x} \in \mathbb{R}^n$ arbitrar fixat.

Fie $\underline{e} \in \mathbb{R}^n$ arbitrar. Atunci:

$$J(\underline{x}+\underline{e}) = \frac{1}{2} \| A(\underline{x}+\underline{e}) - \underline{b}\|_2^2$$

$$= \frac{1}{2} [A(\underline{x}+\underline{e}) - \underline{b}]^T [A(\underline{x}+\underline{e}) - \underline{b}]$$

$$= \frac{1}{2} [(A\underline{x}-\underline{b}) + A\underline{e}]^T [(A\underline{x}-\underline{b}) + A\underline{e}]$$

$$= \frac{1}{2} \Big[(A\underline{x}-\underline{b})^T (A\underline{x}-\underline{b}) + (A\underline{e})^T (A\underline{x}-\underline{b})$$

$$+ (A\underline{x}-\underline{b})^T (A\underline{e}) + (A\underline{e})^T (A\underline{e})\Big]$$

$$= \frac{1}{2} \|A\underline{x}-\underline{b}\|_2^2 + \underline{e}^T [A^T(A\underline{x}-\underline{b})]$$

$$+ \frac{1}{2} \underline{e}^T (A^T A) \underline{e}$$

$$= J(\underline{x}) + \underline{e}^T (A^T A \underline{x} - A^T \underline{b}) + \frac{1}{2} \underline{e}^T (A^T A) \underline{e}$$

$$|\underline{e}^T (A^T A) \underline{e}| \le \|\underline{e}\|_2 \cdot \|(A^T A)\underline{e}\|_2 \le$$

$$\le \|\underline{e}\|_2 \, \|A^T A\| \, \|\underline{e}\|_2 = O(\|\underline{e}\|_2^2)$$

Obținem astfel:

$$\mathcal{J}(\underline{x}+\underline{e})-\mathcal{J}(\underline{x}) = \underline{e}^T(A^TA\underline{x}-A^T\underline{b})+O(\|\underline{e}\|_2^2)$$
$$\forall \underline{e}\in\mathbb{R}^n$$

i.e. derivata Fréchet a lui $\mathcal{J}$ în $\underline{x}\in\mathbb{R}^n$ este dată de:

$$\boxed{\nabla\mathcal{J}(\underline{x}) = A^T(A\underline{x}-\underline{b})}$$

„$\Rightarrow$": Fie $\hat{\underline{x}}\in\mathbb{R}^n$ a.î. $\hat{\underline{x}} = \arg\min_{\underline{x}\in\mathbb{R}^n} \mathcal{J}(\underline{x})$

$\Rightarrow \nabla\mathcal{J}(\hat{\underline{x}}) = \underline{0}_n \Rightarrow A^T(A\hat{\underline{x}}-\underline{b}) = \underline{0}_n$

i.e. $\hat{\underline{x}}$ soluția sistemului de ecuații normale (5). $\square$

„$\Leftarrow$": Fie $\hat{\underline{x}}\in\mathbb{R}^n$ a.î. $A^T(A\hat{\underline{x}}-\underline{b}) = \underline{0}_n \Rightarrow$

$$\mathcal{J}(\hat{\underline{x}}+\underline{e})-\mathcal{J}(\hat{\underline{x}}) = \underline{e}^T\underbrace{\left[A^T(A\hat{\underline{x}}-\underline{b})\right]}_{=\underline{0}_n}+$$
$$+\frac{1}{2}\|A\underline{e}\|_2^2 \geq 0,\ \forall \underline{e}\in\mathbb{R}^n \Rightarrow$$
$$\Rightarrow \mathcal{J}(\underline{x}) \geq \mathcal{J}(\hat{\underline{x}}),\ \forall \underline{x}\in\mathbb{R}^n \quad \square$$

OBSERVAȚII:

1) Cf Teoremei#1, o soluțiile PCMMP (3a) sau (3b) se determină rezolvând sistemul de ecuații normale asociat (4), i.e.

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

unde $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $m > n$ și $\underline{b} \in \mathbb{R}^m$.

2) Dacă $A \in \mathcal{M}_{m,n}(\mathbb{R})$ este inversabilă la stânga, atunci PCMMP (3a) sau (3b) are soluție unică, dată de

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

3) Mai mult, în acest caz, $A^T A \in \mathcal{M}_n(\mathbb{R})$ este SPD și putem aplica factorizarea Cholesky a matricei $A^T A \in \mathcal{M}_n(\mathbb{R})$.

4) Factorizarea Cholesky este instabilă pentru rezolvarea a PCMMP! (laborator)

ALGORITM (PCMMP & factorizarea Cholesky):

Date: $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $m > n$, $\operatorname{rank}(A) = n$

$\underline{b} \in \mathbb{R}^m$

1) Calculează $C := A^T A \in \mathcal{M}_n(\mathbb{R})$

$\underline{d} := A^T \underline{b} \in \mathbb{R}^n$

2) Calculează factorizarea Cholesky a matricei C:

$C = L L^T$

$L \in \mathcal{M}_n(\mathbb{R})$ inferior triunghiulară

3) Rezolvă sistemul inferior triunghiular

$L \underline{y} = \underline{d} \Rightarrow \underline{y} := L^{-1} \underline{d}$

4) Rezolvă sistemul superior triunghiular

$L^T \underline{x} = \underline{y} \Rightarrow \underline{x} := L^{-T} \underline{y}$

APLICAȚIE (problema regresiei liniare):

Date fiind punctele din plan

$$D = \{(x_i, y_i) \in \mathbb{R}^2 \mid i = \overline{1,10}\} \subset \mathbb{R}^2$$

să se determine dreapta d de ecuație

$$d: \quad \alpha x + \beta = y \qquad (\alpha, \beta \in \mathbb{R})$$

care este "cea mai apropiată" de punctele mulțimii D.

Trebuie determinate $\alpha, \beta \in \mathbb{R}$ a.î.

$$\alpha x_i + \beta = y_i, \ i = \overline{1,10} \iff$$

$$\underbrace{\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \\ x_{10} & 1 \end{bmatrix}}_{A \in \mathcal{M}_{10,2}(\mathbb{R})} \underbrace{\begin{bmatrix} \alpha \\ \beta \end{bmatrix}}_{= \underline{x} \in \mathbb{R}^2} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{10} \end{bmatrix}}_{= \underline{b} \in \mathbb{R}^{10}} \iff$$

$$\boxed{\begin{gathered} \text{Determină } \underline{x} := (\alpha, \beta)^T \in \mathbb{R}^2 \text{ a.î.} \\ A\underline{x} = \underline{b} \end{gathered}}$$

IDEEA (Gauss):

$$\hat{\underline{x}} := \arg\min_{\underline{x}=(\alpha,\beta)^T\in\mathbb{R}^2} \|A\underline{x}-\underline{b}\|_2^2$$

$$= \arg\min_{\underline{x}:=(\alpha,\beta)^T\in\mathbb{R}^2} \sum_{i=1}^{10} (\alpha x_i+\beta-y_i)^2$$

OBS: Problema regresiei liniare se poate generaliza astfel:

$$p(\underline{x}) := \sum_{i=0}^{n} c_i x^i \in \mathbb{P}_n \text{ polinom de grad} \leq n$$

$$D_m := \{(x_i, y_i)\in\mathbb{R}^2 \mid i=\overline{1,m}\} \subset \mathbb{R}^2$$

Să se determine $p \in \mathbb{P}_n$ (ie $c_0, c_1, \ldots, c_n \in \mathbb{R}$) a.î. p este "cel mai apropiat" de punctele mulțimii D_m.

Fie $\underline{x} := (c_0, c_1, \ldots, c_n)^T \in \mathbb{R}^{n+1}$.

Problema se reformulează astfel:

Determinați

$$\underline{\hat{x}} := (\hat{c}_0, \hat{c}_1, \dots, \hat{c}_n)^T \in \mathbb{R}^{n+1}$$

a.î.

$$\underline{\hat{x}} := \arg\min_{\underline{x} \in \mathbb{R}^{n+1}} \sum_{i=0}^{n} [p(x_i) - y_i]^2$$

$$= \arg\min_{\underline{x} \in \mathbb{R}^{n+1}} \|A\underline{x} - \underline{y}\|_2^2$$

unde

$$\underbrace{\begin{bmatrix} x_1^n & x_1^{n-1} & \cdots & x_1 & 1 \\ x_2^n & x_2^{n-1} & \cdots & x_2 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ x_m^n & x_m^{n-1} & \cdots & x_m & 1 \end{bmatrix}}_{= A \in \mathcal{M}_{m,n+1}(\mathbb{R})} \underbrace{\begin{bmatrix} c_n \\ c_{n-1} \\ \vdots \\ c_0 \end{bmatrix}}_{= \underline{x} \in \mathbb{R}^{n+1}} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}}_{= \underline{b} \in \mathbb{R}^m}$$

$m > n+1$

<u>necunoscute</u>

METODA SISTEMULUI AUGMENTAT

Considerăm, din nou, sistemul de ecuații liniare supradeterminate

$$A \underline{x} = \underline{b} \qquad (1)$$

unde

- $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $m > n$, $\operatorname{rank}(A) = n$;
- $\underline{b} \in \mathbb{R}^m$

<u>OBSERVAȚII</u>:

1) Cf Teoremei #1, există și este unică soluția în sensul celor mai mici pătrate, $\underline{\hat{x}} \in \mathbb{R}^n$, a sistemului (1).

2) Soluția în sensul celor mai mici pătrate, $\underline{\hat{x}} \in \mathbb{R}^n$, a sistemului (1) se poate determina prin rezolvarea sistemului de ecuații normale

$$(A^T A)\,\underline{x} = A^T \underline{b} \qquad (2)$$

3) <u>Metoda sistemului augmentat</u> este o variantă a metodei sistemului de ecuații normale care folosește <u>vectorul eroare reziduală</u> definit prin

$$\underline{r} := \underline{b} - A\underline{x} \in \mathbb{R}^m \qquad (3)$$

4) Folosind definiția vectorului eroare reziduală, $\underline{r}$, <u>sistemul de ecuații normale</u> (2) se rescrie sub forma

$$A^T(\underline{b} - A\underline{x}) = \underline{0}_n$$

$$A^T\underline{r} = \underline{0}_n \qquad (4)$$

5) Relațiile (3) și (4) determină <u>sistemul augmentat</u>

$$\begin{cases} \underline{r} + A\underline{x} = \underline{b} \\ A^T\underline{r} \qquad\;\; = \underline{0}_n \end{cases} \qquad (5)$$

pentru <u>necunoscutele</u>

$\underline{x} \in \mathbb{R}^n$ și $\underline{r} \in \mathbb{R}^m$

care poate fi scris sub forma:

$$\underbrace{\begin{bmatrix} I_m & A \\ A^T & O_n \end{bmatrix}}_{\in \mathcal{M}_{m+n}(\mathbb{R})} \underbrace{\begin{bmatrix} \underline{r} \\ \underline{x} \end{bmatrix}}_{\in \mathbb{R}^{m+n}} = \underbrace{\begin{bmatrix} \underline{b} \\ \underline{0}_n \end{bmatrix}}_{\in \mathbb{R}^{m+n}} \qquad (6)$$

6) AVANTAJE / DEZAVANTAJE:

(i) Sistemul augmentat (5) sau (6) este simetric, dar nu este PD.

(ii) Sistemul augmentat este mai mare decât sistemul de ecuații normale.

(iii) Trebuie stocate două copii ale matricei A, i.e. A și A^T.

(iv) Aplicând MEGFP, se obține sistemul de ecuații normale din nou.

(v) Se poate aplica fie MEGPP/MEGPPS, fie factorizarea LU/PLU, în mod convenabil. În mod evident, alegerea pivotului depinde de mărimea relativă a coeficienților matricelor corespunzătoare celor două linii ale partiționării matricei augmentate.

(vi) Deoarece componentele lui $\underline{r}$ și $\underline{x}$ au ordine de mărime diferite în general, introducem un <u>parametru de scalare</u>, $\alpha > 0$, a vectorului rezidual și obținem:

$$\begin{bmatrix} \alpha I_m & A \\ A^T & \mathbb{O}_n \end{bmatrix} \begin{bmatrix} \underline{r}/\alpha \\ \underline{x} \end{bmatrix} = \begin{bmatrix} \underline{b} \\ \underline{\mathbb{O}}_n \end{bmatrix} \qquad (7)$$

În practică, $\alpha := \max_{i,j} |a_{ij}| \times 10^{3}$.