Curs #5 (Notițe)

Factorizarea Cholesky.

Def. Fie $A=(a_{ij})_{i,j=\overline{1,n}} \in M_n(\mathbb{R})$. Numim descompunerea Cholesky a matricei A, scrierea matricei A ca produs de două matrice L și L^T, unde $L=(l_{ij})_{i,j=\overline{1,n}}$ este o matrice inferior triunghiulară.

Astfel A se scrie: $\boxed{A = LL^T}$

Dacă matricea A admite factorizarea Cholesky, atunci sistemul $\boxed{Ax=b}$ se rezolvă imediat.

$$Ax = b \iff \underbrace{LL^T x}_{y} = b \iff \begin{cases} Ly = b \\ L^T x = y \end{cases}$$

Primul sistem $Ly=b$ este inferior triunghiular și se poate rezolva conform metodei substituției ascendente.

Cel de-al doilea sistem $L^T x = y$ este un sistem superior triunghiular și poate fi rezolvat conform metodei substituției descendente.

Se pune următoarea întrebare: Cum trebuie să fie matricea A ca să admită factorizarea Cholesky?

Th. Dacă $A \in M_n(\mathbb{R})$ este simetrică și pozitiv definită, atunci matricea A admite factorizarea Cholesky.

Def. a) $A \in \mathcal{M}_n(\mathbb{R})$ s.n. semipozitiv definită (seminegativ definită) $\Leftrightarrow$ $\langle Av, v\rangle \geq 0, \forall v \in \mathbb{R}^n$ $(\langle Av, v\rangle \leq 0, \forall v \in \mathbb{R}^n)$

b) $A \in \mathcal{M}_n(\mathbb{R})$ s.n. pozitiv definită (negativ definită) $\Leftrightarrow \langle Av, v\rangle \geq 0, \forall v \in \mathbb{R}^n$ și $\langle Av, v\rangle = 0 \Leftrightarrow v = 0_n$

$(\langle Av, v\rangle \leq 0, \forall v \in \mathbb{R}^n$ și $\langle Av, v\rangle = 0 \Leftrightarrow v = 0_n)$

Cum verificăm dacă o matrice simetrică este pozitiv definită?

Criteriul lui Sylvester: Fie $A \in \mathcal{M}_n(\mathbb{R})$ simetrică.

Atunci A este pozitiv definită $\Leftrightarrow \det A_k > 0, k = \overline{1,n}$

unde $A_k = (a_{ij})_{i,j=\overline{1,k}}$

Unde se întâlnesc matricele simetrice și pozitiv definite?

Fie forma pătratică $f(x,y) = \langle A \cdot \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \rangle$, $A \in \mathcal{M}_2(\mathbb{R})$ simetrică. Funcția f se scrie desfășurat astfel:

$$f(x,y) = \langle \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \rangle = \langle \begin{pmatrix} a_{11}x + a_{12}y \\ a_{12}x + a_{22}y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \rangle$$

$$= (a_{11}x + a_{12}y) \cdot x + (a_{12}x + a_{22}y) \cdot y =$$

$$= a_{11}x^2 + 2a_{12}x \cdot y + a_{22}y^2$$

Exemplu 1 a) $f(x,y) = \langle A \cdot \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \rangle$, $A = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}$

A este simetrică, mai mult $\Delta_1 = 4 > 0$, $\Delta_2 = \begin{vmatrix} 4 & 1 \\ 1 & 2 \end{vmatrix} = 7 > 0$,

deci A este pozitiv definită

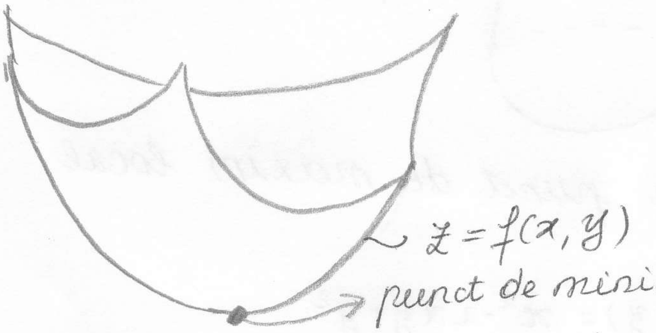
Altfel: $f(x,y) = \left\langle \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 4x+y \\ x+2y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle$

$$= 4x^2 + 2xy + 2y^2 = (2x)^2 + 2\cdot 2x \cdot \frac{y}{2} + \left(\frac{y}{2}\right)^2 - \underbrace{\left(\frac{y}{2}\right)^2}_{4} + 2y^2$$

$$= \left(2x + \frac{y}{2}\right)^2 + \frac{7}{4}y^2 \geq 0, \quad \forall\ (x,y) \in \mathbb{R}^2$$

$$f(x,y) = 0 \iff \begin{cases} 2x + \frac{y}{2} = 0 \\ y = 0 \end{cases} \iff \begin{cases} x = 0 \\ y = 0 \end{cases} \Rightarrow (x,y) = (0,0)$$

Reprezentarea grafică:



Se observă că suprafața $z = f(x,y)$ admite punct de minim local.

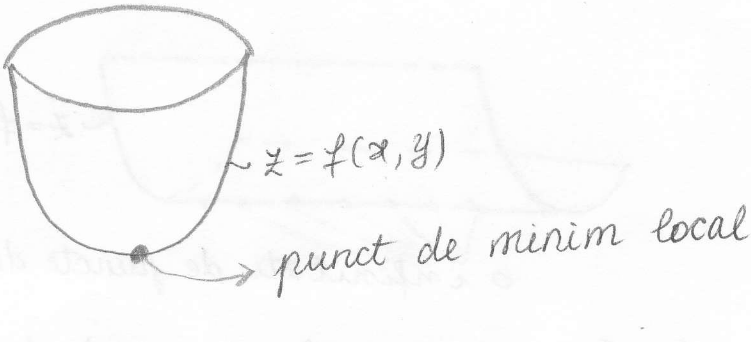
b) $f(x,y) = x^2 + y^2$

$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $\quad \Delta_1 = 1 > 0$, $\Delta_2 = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 > 0$ $\Rightarrow A(f)$ – este pozitiv definită

Conform definiției: $f(x,y) = x^2 + y^2 \geq 0, \ \forall\ (x,y) \in \mathbb{R}^2$

$f(x,y) = 0 \iff x^2 + y^2 = 0 \iff x = y = 0$

Reprezentarea grafică:

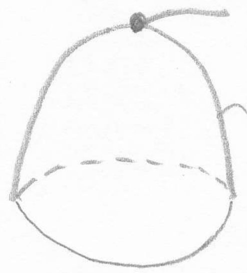


Exemplu 2. $f(x,y) = -(x^2+y^2)$

$$A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\left.\begin{array}{l} f(x,y) \leq 0, \ \forall (x,y) \in \mathbb{R}^2 \\ f(x,y) = 0 \iff x = y = 0 \end{array}\right| \Rightarrow f \text{ este negativ definită}$$

Reprezentarea grafică:



f - admite punct de maxim local

Exemplu 3. $f(x,y) = x^2 - 2xy + y^2$

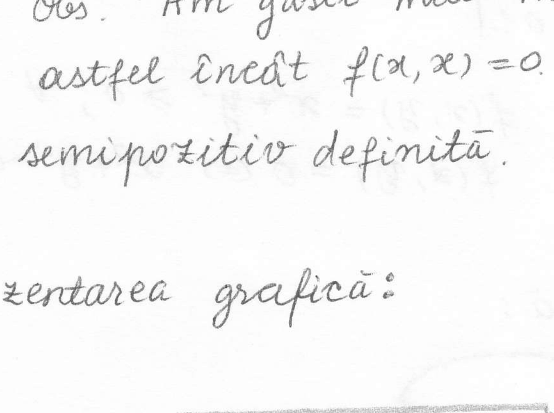
$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

$f(x,y) = (x-y)^2 \geq 0, \ \forall (x,y) \in \mathbb{R}^2$

$f(x,y) = 0 \iff x - y \iff x = y \Rightarrow (x,x), \ \forall x \in \mathbb{R}$

Obs. Am găsit mai multe puncte de forma (x,x) astfel încât $f(x,x) = 0$, deci f este cel mult semipozitiv definită.

Reprezentarea grafică:

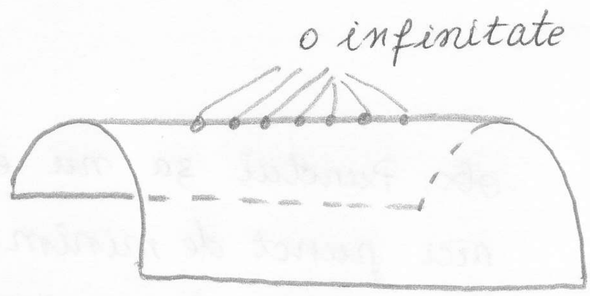


f - admite o infinitate de puncte de minim (nu avem unicitate)

Exemplu 4 $f(x,y) = -(x^2-2xy+y^2) \Rightarrow$

$$\left.\begin{array}{l} f(x,y) = -(x-y)^2 \leq 0, \ \forall (x,y)\in\mathbb{R}^2 \\ f(x,y)=0 \Leftrightarrow x=y \end{array}\right| \Rightarrow f \text{ este seminegativ definită}$$

Reprezentarea grafică:



f - admite o infinitate de puncte de maxim

(nu avem unicitate)

Exemplu 5 a) $f(x,y) = 2x^2 + 8xy + 2y^2$

$$A = \begin{pmatrix} 2 & 4 \\ 4 & 2 \end{pmatrix}$$

$$f(x,y) = (\sqrt{2}\,x)^2 + 2\sqrt{2}\,x \cdot 2\sqrt{2}\,y + (2\sqrt{2}\,y)^2 - (2\sqrt{2}\,y)^2 + 2y^2$$

$$= (\sqrt{2}\,x + 2\sqrt{2}\,y)^2 - 6y^2$$

Dacă alegem $(x,y)=(0,1)$ se obține $f(0,1) = 8-6=2>0$,

însă dacă alegem $(x,y)=(-1,1) \Rightarrow f(-1,1) = -4 < 0$

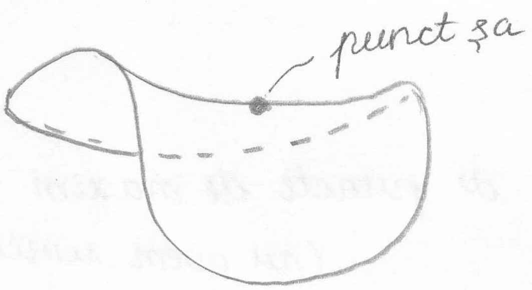
În concluzie, f nu este nici pozitiv definită, nici negativ definită. În acest caz o vom numi nedefinită.

b) $f(x,y) = x^2 - y^2$

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\left.\begin{array}{l} f(1,0) = 1 \\ f(0,1) = -1 \end{array}\right| \Rightarrow f \text{ este nedefinită}$$

Reprezentarea grafică:



obs. Punctul şa nu este nici punct de minim, nici punct de maxim.

În concluzie: Forma pătratică f a cărei matrice este pozitiv definită (negativ definită) admite un punct de maxim (respectiv minim).

Se prezintă în continuare două metode de determinare a matrice L

Metoda I: $A = LL^T$

$$\underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1i} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} & \cdots & a_{ki} & \cdots & a_{kn} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ik} & \cdots & a_{ii} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nk} & \cdots & a_{ni} & \cdots & a_{nn} \end{pmatrix}}_{A} = \underbrace{\begin{pmatrix} l_{11} & & & & & & \\ \vdots & \ddots & & & & 0 & \\ l_{k1} & \cdots & l_{kk} & & & & \\ \vdots & & \vdots & \ddots & & & \\ l_{i1} & \cdots & l_{ik} & \cdots & l_{ii} & & \\ \vdots & & \vdots & & \vdots & \ddots & \\ l_{n1} & \cdots & l_{nk} & \cdots & l_{ni} & \cdots & l_{nn} \end{pmatrix}}_{L} \cdot$$

linia k din L

linia i din L

$$\begin{array}{ccc} & \text{coloana } k \text{ din } L^T & \text{coloana } i \text{ din } L^T \\ & \downarrow & \downarrow \end{array}$$

$$\begin{pmatrix} l_{11} & \cdots & l_{k1} & \cdots & l_{i1} & \cdots & l_{n1} \\ & \ddots & \vdots & & \vdots & & \vdots \\ & & l_{kk} & \cdots & l_{ik} & \cdots & l_{nk} \\ & & & \ddots & \vdots & & \vdots \\ & & & & l_{ii} & \cdots & l_{ni} \\ & & & & & \ddots & \vdots \\ & & & & & & l_{nn} \end{pmatrix}, \quad k < i$$

Se determină primul element $\boxed{l_{11}}$:

$$a_{11} = (l_{11}, 0, \ldots, 0) \cdot \begin{pmatrix} l_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = l_{11}^2, \quad \text{fie } l_{11} > 0$$

$$\Rightarrow l_{11} = \sqrt{a_{11}}$$

Prop. Dacă $A \in M_n(\mathbb{R})$ este simetrică, pozitiv definită și $\boxed{l_{kk} > 0, \ \forall\, k = \overline{1,n}}$, atunci factorizarea Cholesky este unică.

Se determină în continuare și restul elementelor de pe coloana întâi, i.e. $\boxed{l_{i1}, \ i > 1}$

$$a_{i1} = (l_{i1} \ldots l_{ik} \ldots l_{ii} \ldots 0) \cdot \begin{pmatrix} l_{11} \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix} = l_{i1} \cdot l_{11}$$

$$\Rightarrow l_{i1} = a_{i1} / l_{11} \Rightarrow \boxed{l_{i1} = a_{i1} / \sqrt{a_{11}}, \quad i > 1}$$

Presupunem că printr-o anumită metodologie s-au calculat primele $k-1$ coloane din matricea L. Aceasta presupune cunoscute toate elementele de forma l_{i1}, $i = \overline{1,n}$; l_{i2}, $i = \overline{2,n}$; $\ldots$ $l_{i,k-1}$, $i = \overline{k-1,n}$

Evaluând a_{kk} se obține primul element $\boxed{l_{kk}}$ de pe coloana k

$$a_{kk} = (l_{k1} \dots l_{kk} \dots 0) \cdot \begin{pmatrix} l_{k1} \\ \vdots \\ l_{kk} \\ \vdots \\ 0 \end{pmatrix} = l_{k1}^2 + \dots + l_{kk}^2 =$$

$$= \underbrace{\left(l_{k1}^2 + \dots + l_{k,k-1}^2\right)}_{\text{se cunoaște}} + l_{kk}^2 \Rightarrow$$

$$\boxed{l_{kk} = \sqrt{a_{kk} - \sum_{j=1}^{k-1} l_{kj}}}$$

Continuăm să calculăm restul elementelor de pe coloana k, i.e. $\boxed{l_{ik},\ i = \overline{k+1, n}}$, evaluând a_{ik}

$$a_{ik} = (l_{i1} \dots l_{ik} \dots l_{ii} \dots 0) \begin{pmatrix} l_{k1} \\ \vdots \\ l_{kk} \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix} = l_{i1} l_{k1} + \dots + l_{ik} l_{kk}$$

$$a_{ik} = \underbrace{\left(l_{i1} l_{k1} + \dots + l_{i,k-1} l_{k,k-1}\right)}_{\text{se cunoaște}} + l_{ik} l_{kk}$$

$$\boxed{l_{ik} = \left(a_{ik} - \sum_{j=1}^{k-1} l_{ij} l_{kj}\right) / l_{kk} \quad , i > k}$$

Metoda II. Vom aplica aceeași metodologie, dar lucrând cu blocuri de matrice. Vom partiționa matricile A și L după cum urmează:

$A = L L^T$

$$\begin{array}{c} \\ A_{21} \left\{ \begin{array}{c} \\ \\ \\ \end{array} \right. \end{array} \left(\begin{array}{c|ccc} a_{11} & \multicolumn{3}{c}{\overbrace{a_{12} \cdots a_{1n}}^{A_{21}^T}} \\ \hline a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \\ a_{1n} & a_{n2} & \cdots & a_{nn} \end{array}\right) = \left(\begin{array}{c|ccc} l_{11} & & 0 & \\ \hline l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & 0 \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{array}\right) \cdot \left(\begin{array}{c|ccc} l_{11} & \multicolumn{3}{c}{\overbrace{l_{21} \cdots l_{n1}}^{L_{21}^T}} \\ \hline & l_{22} & \cdots & l_{n2} \\ 0 & & \ddots & \vdots \\ & 0 & & l_{nn} \end{array}\right)$$

$$\underbrace{\phantom{a_{22} \cdots a_{nn}}}_{A_{22}} \qquad \underbrace{}_{L_{21}} \quad \underbrace{\phantom{l_{22} \cdots l_{nn}}}_{L_{22}} \qquad \underbrace{\phantom{l_{22} \cdots l_{nn}}}_{L_{22}^T}$$

$$\left(\begin{array}{c|c} a_{11} & A_{21}^T \\ \hline A_{21} & A_{22} \end{array}\right) = \left(\begin{array}{c|c} l_{11} & 0 \\ \hline L_{21} & L_{22} \end{array}\right) \cdot \left(\begin{array}{c|c} l_{11} & L_{21}^T \\ \hline 0 & L_{22}^T \end{array}\right)$$

$$a_{11} = l_{11}^2 \Rightarrow \boxed{l_{11} = \sqrt{a_{11}}}$$

$$A_{21} = L_{21} \cdot l_{11} + L_{22} \cdot 0 \quad \Rightarrow \boxed{L_{21} = \frac{1}{l_{11}} \cdot A_{21}}$$

Astfel se calculează prima coloană din matricea L.

În final: $A_{22} = L_{21} \cdot L_{21}^T + L_{22} \cdot L_{22}^T \Rightarrow$

$$\underbrace{A_{22} - L_{21} \cdot L_{21}^T}_{\substack{\overset{not}{=} \\ S \in \mathcal{M}_{n-1}(\mathbb{R})}} = \underset{\substack{\cap \\ \mathcal{M}_{n-1}(\mathbb{R})}}{L_{22} \cdot L_{22}^T}$$

S se mai numește și complementul lui Schur.

Se poate demonstra că S este simetrică și pozitiv definită.

Întrucât L_{22} este o matrice inferior triunghiulară, am obținut că S admite factorizare Cholesky.

Se repetă procedeul pentru S pentru a determina prima coloană în L_{22}, care de fapt reprezintă a doua coloană în L. Se va obține factorizarea unei matrice de ordinul $n-2$, ș.a.m.d.

Exemplu

Fie matricea $A=\begin{pmatrix}4&2&2\\2&10&4\\2&4&6\end{pmatrix}$

a) Să se demonstreze că A este poz. definită

b) Să se determine factorizarea Cholesky

c) să se determine soluția sistemului $Ax=b$, $b=\begin{pmatrix}2\\-2\\-4\end{pmatrix}$

Rezolvare: a) A - sim, $\Delta_1=4>0$, $\Delta_2=\begin{vmatrix}4&2\\2&10\end{vmatrix}=40-4=36>0$, $\Delta_3=|A|>0$

cf. crit Sylvester $\Rightarrow$ f este pozitiv definită

b) $$\left(\begin{array}{c|cc}4&2&2\\\hline 2&10&4\\2&4&6\end{array}\right)=\left(\begin{array}{c|c}l_{11}&0\\\hline L_{21}&L_{22}\end{array}\right)\cdot\left(\begin{array}{c|c}l_{11}&L_{21}^T\\\hline 0&L_{22}^T\end{array}\right)$$

$$4=l_{11}^2 \underset{l_{11}>0}{\Rightarrow} l_{11}=2$$

$$A_{21}=L_{21}\cdot l_{11}+L_{22}\cdot 0 \Rightarrow L_{21}=A_{21}:l_{11}\Rightarrow L_{21}=\begin{pmatrix}2\\2\end{pmatrix}\cdot\frac{1}{2}=\begin{pmatrix}1\\1\end{pmatrix}$$

$$\begin{pmatrix}10&4\\4&6\end{pmatrix}=L_{21}L_{21}^T+L_{22}L_{22}^T\Rightarrow\begin{pmatrix}10&4\\4&6\end{pmatrix}-\begin{pmatrix}1\\1\end{pmatrix}\cdot(1\ \ 1)=L_{22}L_{22}^T$$

$$\Rightarrow\begin{pmatrix}10&4\\4&6\end{pmatrix}-\begin{pmatrix}1&1\\1&1\end{pmatrix}=L_{22}L_{22}^T\Rightarrow\underbrace{\begin{pmatrix}9&3\\3&5\end{pmatrix}}_{S}=L_{22}L_{22}^T$$

S - simetrică și poz. definită ($\Delta_1=9>0$, $\Delta_2=|S|=45-9>0$)

Se aplică aceeași metodologie:

$$\left(\begin{array}{c|c} 9 & 3 \\ \hline 3 & 5 \end{array}\right) = \left(\begin{array}{c|c} \ell_{22} & 0 \\ \hline \ell_{32} & \ell_{33} \end{array}\right) \cdot \left(\begin{array}{c|c} \ell_{22} & \ell_{32} \\ \hline 0 & \ell_{33} \end{array}\right)$$

$$9 = \ell_{22}^2 \Rightarrow \ell_{22} = 3$$

$$3 = \ell_{32} \cdot \ell_{22} \Rightarrow \ell_{32} = 3 : 3 = 1$$

$$5 = \ell_{32}^2 + \ell_{33}^2 \Rightarrow \ell_{33} = \sqrt{5 - \ell_{32}^2} = \sqrt{4} = 2$$

$$L = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ 1 & 1 & 2 \end{pmatrix}, \quad L^T = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

c) $L\underbrace{L^T x}_{y} = b \iff \begin{cases} Ly = b \\ L^T x = y \end{cases}$

$$\begin{cases} 2y_1 = 2 \\ y_1 + 3y_2 = -2 \\ y_1 + y_2 + 2y_3 = -4 \end{cases} \Rightarrow \begin{cases} y_1 = 1 \\ y_2 = (-2 - 1) : 3 = -3 : 3 = -1 \\ y_3 = (-4 - 1 + 1) : 2 = -2 \end{cases}$$

$$y = \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}$$

$$\begin{cases} 2x_1 + x_2 + x_3 = 1 \\ 3x_2 + x_3 = -1 \\ 2x_3 = -2 \end{cases} \Rightarrow \begin{cases} x_1 = (1 - 0 + 1) : 2 = 1 \\ x_2 = (-1 + 1) : 3 = 0 \\ x_3 = -1 \end{cases}$$

Soluție: $x = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$