CURS #10

CONȚINUTUL CURSULUI #10:

V. Minimizarea funcțiilor de două variabile.

- V.2. Forme pătratice, tipuri de extreme locale.
- V.3. Metoda gradientului conjugat.

V. Minimizarea funcțiilor de două variabile.
V.2. Forme pătratice, tipuri de extreme locale.

Reamintim următoarele definiții:

Definiția (V.1.)

a) Matricea $A \in \mathcal{M}_n(\mathbb{R})$ se numește semipozitiv definită (seminegativ definită) $\Leftrightarrow \langle Av, v\rangle \geqslant 0, \forall v \in \mathbb{R}^n$ ($\langle Av, v\rangle \leqslant 0,\ \forall v \in \mathbb{R}^n$).
b) Matricea $A \in \mathcal{M}_n(\mathbb{R})$ s.n. pozitiv definită (negativ definită)

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

c) Dacă matricea A nu este nici de tipul a), nici de tipul b) zicem ca aceasta este nedefinită.

Teorema (V.2. Criteriul lui Sylvester:)

Fie $A \in \mathcal{M}_n(\mathbb{R})$ simetrică. Atunci A este pozitiv (negativ) definită

$$\Leftrightarrow \det A_k > 0, (\operatorname{sign}(\det A_k) = (-1)^k), k = \overline{1, n}$$

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

În continuare, vom da câteva exemple de forme pătratice și vom discuta tipul punctelor de extrem pe care acestea le admit. Fie forma patratica $f(x, y) = \left\langle A \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle, A \in \mathcal{M}_2(\mathbb{R})$ simetrică. Funcția f se scrie desfășurat astfel:

$$\begin{aligned} f(x, y) &= \left\langle \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle \\ &= \left\langle \begin{pmatrix} a_{11}x + a_{12}y \\ a_{12}x + a_{22}y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle \\ &= (a_{11}x + a_{12}y) \cdot x + (a_{12}x + a_{22}y) \cdot y \\ &= a_{11}x^2 + 2a_{12}x \cdot y + a_{22}y^2 \end{aligned}$$

Exemplu (1.)

a) $f(x, y) = \left\langle A \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle, A = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}$. Matricea 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. Conform reprezentării grafice, funcția f admite punct de minim local.

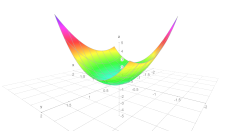


Figure: *Reprezentarea grafică a funcției $z = 4x^2 + 2xy + 2y^2$*

Exemplu (1.)

b) $f(x,y) = \left\langle A\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle, A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. *Matricea A este simetrică, mai mult* $\Delta_1 = 1 > 0, \Delta_2 = 1 > 0$, *deci A este pozitiv definită. Conform reprezentării grafice, funcția* $z = x^2 + y^2$ *admite punct de minim local. Pozitiv definirea se mai poate demonstra și conform definiției:*

$$f(x,y) = x^2 + y^2 \geqslant 0, \forall (x,y) \in \mathbb{R}^2$$
$$f(x,y) = 0 \Leftrightarrow x^2 + y^2 = 0 \Leftrightarrow x = y = 0$$

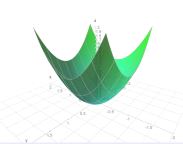


Figure: *Reprezentarea grafică a funcției* $z = x^2 + y^2$

Exemplu (2.)

b) $f(x,y) = \left\langle A\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle, A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$. *Matricea A este simetrică, mai mult* $\Delta_1 = -1 < 0, \Delta_2 = 1 > 0$, *deci A este negativ definită. Conform reprezentării grafice, funcția* $z = -(x^2 + y^2)$ *admite punct de maxim local. Negativ definirea se mai poate demonstra și conform definiției:*

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

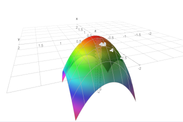


Figure: *Reprezentarea grafică a funcției* $z = -x^2 - y^2$

Exemplu (3.)

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

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$
$$f(x,y) = (x-y)^2 \geqslant 0, \forall (x,y) \in \mathbb{R}^2$$
$$f(x,y) = 0 \Leftrightarrow x - y \Leftrightarrow x = y \Rightarrow (x,x), \forall x \in \mathbb{R}$$

Am găsit mai multe puncte de forma (x,x) *astfel încât* $f(x,x) = 0$, *deci f este semipozitiv definită. Funcția f admite o infinitate de puncte de minim.*

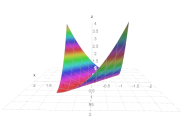


Figure: *Reprezentarea grafică a funcției* $z = x^2 - 2xy + y^2$

Exemplu (4.)

$$f(x,y) = -\left(x^2 - 2xy + y^2\right) \Rightarrow$$
$$f(x,y) = -(x-y)^2 \leq 0, \forall (x,y) \in \mathbb{R}^2$$
$$f(x,y) = 0 \Leftrightarrow x = y \quad \Rightarrow f \text{ este seminegativ definită}$$

Am găsit mai multe puncte de forma (x,x) *astfel încât* $f(x,x) = 0$, *deci f este semipozitiv definită. În acest caz funcția f admite o infinitate de puncte de maxim.*

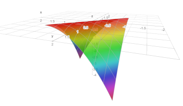


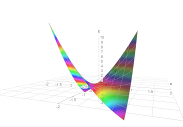
Figure: *Reprezentarea grafică a funcției* $z = x^2 - 2xy + y^2$

Exemplu (5.)

$$f(x,y) = 2x^2 + 8xy + 2y^2, A = \begin{pmatrix} 2 & 4 \\ 4 & 2 \end{pmatrix}$$

$$\begin{aligned} f(x,y) &= (\sqrt{2}x)^2 + 2\sqrt{2}x \cdot 2\sqrt{2}y + (2\sqrt{2}y)^2 - (2\sqrt{2}y)^2 + 2y^2 \\ &= (\sqrt{2}x + 2\sqrt{2}y)^2 - 6y^2 \end{aligned}$$

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/semipozitiv definită, nici negativ/seminegativ definită. In acest caz o vom numi nedefinită. Funcția nu admite nici punct de minim, nici punct de maxim, ci punct șa.



V. Minimizarea funcțiilor de două variabile.
V.3. Metoda gradientului conjugat.

Metoda gradientului conjugat presupune aflarea punctului de minim în doi pași.

Fie $(x_0, y_0) \in D$, unde D reprezintă domeniului de definiție al funcției $f : D \subset \mathbb{R}^2 \to \mathbb{R}$. Alegem la primul pas, analog metodei pasului descendent, $v^{(0)} = \nabla f|_{(x_0,y_0)}$. Vom construi, de data aceasta, $v^{(1)}$ astfel încât $\langle v^{(1)}, Av^{(0)}\rangle = 0$. Se spune că vectorul $v^{(1)}$ este $A-$ortogonal pe $v^{(0)}$.

La fiecare pas k, $k = \overline{1,2}$, vom nota $r^{(k)} = b - Ax^{(k)}$, semnificând restul la fiecare iterație. Amintim că dorim să minimizăm funcția

$$f(x,y) = \frac{1}{2}\left\langle A\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix}\right\rangle - \left\langle \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix}\right\rangle,$$

unde A este o matrice simetrică și pozitiv definită. Aceste cerințe asupra matricei A asigură existența unui punct de minim local și astfel să putem aplica metode de calcul al punctului de minim local.

Considerăm $v^{(1)}$ de forma $v^{(1)} = -r^{(1)} + \beta_0 v^{(0)}$, unde coeficientul β_0 se va calcula astfel încât $\langle v^{(1)}, Av^{(0)}\rangle = 0$, condiție echivalentă cu $\langle -r^{(1)} + \beta_0 v^{(0)}, Av^{(0)}\rangle = 0$, de unde rezultă

$$\beta_0 = \frac{\langle r^{(1)}, Av^{(0)}\rangle}{\langle v^{(0)}, Av^{(0)}\rangle}$$

Cu excepția modalității de calcul a noii direcții, algoritmul gradientului conjugat rămâne similar cu algoritmul pasului descendent.

ALGORITM (Metoda gradientulu conjugat)

Date de intrare: $A \in \mathcal{M}_2(\mathbb{R})$ - simetrică și pozitiv definită, $b \in \mathbb{R}^2$
Date de ieșire: $x^{(2)}$

1:
- $x^{(0)} = (x_0, y_0) \in D \subset \mathbb{R}^2$ - ales arbitrar
- $r^{(0)} = b - Ax^{(0)} = -\nabla f|_{(x_0,y_0)}$ - se calculează restul
- $v^{(0)} = -r^{(0)}$ - se alege $v^{(0)}$ după direcţia gradientului
- $\alpha_0 = -\dfrac{\langle v^{(0)}, r^{(0)}\rangle}{\langle v^{(0)}, Av^{(0)}\rangle}$

2:
- $x^{(1)} = x^{(0)} - \alpha_0 v^{(0)}$
- $r^{(1)} = b - Ax^{(1)}$
- $\beta_0 = \dfrac{\langle r^{(1)}, Av^{(0)}\rangle}{\langle v^{(0)}, Av^{(0)}\rangle}$
- $v^{(1)} = -r^{(1)} + \beta_0 v^{(0)}$
- $\alpha_1 = -\dfrac{\langle v^{(1)}, r^{(1)}\rangle}{\langle v^{(1)}, Av^{(1)}\rangle}$

3:
- $x^{(2)} = x^{(1)} - \alpha_1 v^{(1)}$
- STOP

Obs.: Se poate demonstra ușor că $x^{(2)}$ reprezintă soluția sistemului $Ax = b$ care este chiar punctul de minim al funcției f.