

Notiuni introductive

Consideram problema de optimizare:

$$(\text{UNLP}) : \min_{x \in \mathbb{R}^n} f(x)$$

Presupuneri:

- ▶ $f \in \mathcal{C}^2$
- ▶ $\text{dom} f$ multime deschisa in \mathbb{R}^n

Problema de interes consta in gasirea punctelor de minim din $\text{dom} f$.

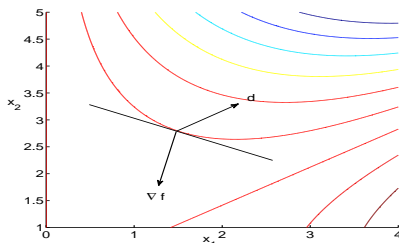
Reamintim ca un punct x^* se numeste punct de minim:

- ▶ global $\iff f(x^*) \leq f(x) \ \forall x \in \text{dom} f$
- ▶ local $\iff \exists \delta > 0$ a.i. $f(x^*) \leq f(x) \ \forall x \in \text{dom} f$ cu $\|x - x^*\| \leq \delta$

Directie de descrestere pentru (UNLP)

Definitie: Se numeste *directie de descrestere* al functiei f in punctul $x \in \text{dom}f$ un vector $d \in \mathbb{R}^n$ ce satisface inegalitatea:

$$\nabla f(x)^T d < 0$$



Interpretare: daca d directie de descrestere a lui f in x atunci functia descreste in vecinatatea lui x . Intr-adevar, exista $t > 0$ suficient de mic a.i. din continuitatea gradientului avem:

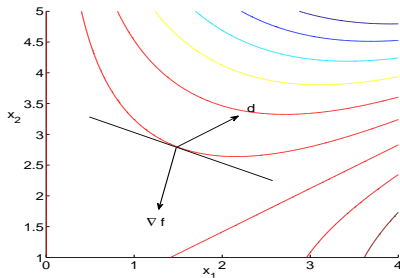
$$\nabla f(x + \tau d)^T d < 0 \quad \forall \tau \in [0, t]$$

Din aproximarea Taylor avem ca exista $\theta \in [0, t]$ a.i.:

$$f(x + td) = f(x) + \nabla f(x + \theta d)^T d < f(x)$$

Conditii necesare de ordinul I pentru (UNLP)

directie de descrestere d : $\nabla f(x)^T d < 0$



Teorema (Conditii necesare de ordinul I): Fie $f \in \mathcal{C}^1$ si $x^* \in \text{dom}f$ un punct de minim local al (UNLP). Atunci avem:

$$\nabla f(x^*) = 0$$

Conditii de ordinul I - Demonstratie

Demonstratie:

Consideram un punct de minim local x^* si presupunem prin absurd ca $\nabla f(x^*) \neq 0$. In acest caz vom arata ca $d = -\nabla f(x^*)$ este directie de descrestere in $x^* \Rightarrow$ contrazice ipoteza initiala.

Datorita continuitatii gradientului, putem alege un $t > 0$ suficient de mic a.i. $\forall \tau \in [0, t]$ avem:

$$\nabla f(x^* + \tau d)^T d = -\nabla f(x^* - \tau \nabla f(x^*))^T \nabla f(x^*) < 0$$

In plus, exista $\theta \in [0, t]$ ce satisface:

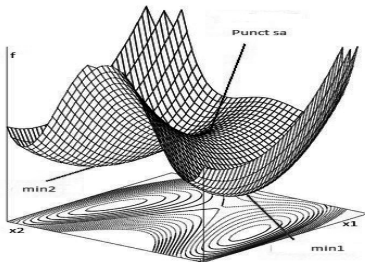
$$f(x^* - t \nabla f(x^*)) = f(x^*) - t \nabla f(x^* - \theta \nabla f(x^*))^T \nabla f(x^*) < f(x^*)$$

Ultima inegalitate contrazice ipoteza initiala, i.e. x^* este punct de minim local.

Conditii de ordinul I - Puncte stationare

Punctele $x^* \in \text{dom} f$ ce satisfac conditiile necesare de ordinul I $\nabla f(x^*) = 0$ se numesc puncte stationare si se pot imparti in:

- ▶ puncte de minim (globale si locale)
- ▶ puncte de maxim (globale si locale)
- ▶ puncte sa (de inflexiune)



Observatie: Pentru a gasi punctele stationare trebuie rezolvat sistemul de n ecuatii neliniare cu n necunoscute: $\nabla f(x) = 0$.

Conditii de ordinul I - Exemplu

Consideram problema (UNLP):

$$\min_{x \in \mathbb{R}^2} f(x) \left(= x_1^2 + \frac{1}{2}x_1x_2 + 5x_2 \sin(x_1) \right)$$

Pentru a gasi punctele stationare rezolvam sistemul $\nabla f(x) = 0$:

$$2x_1 + 0.5x_2 + 5x_2 \cos x_1 = 0 \quad \text{si} \quad 0.5x_1 + 5 \sin x_1 = 0$$

Punctele stationare apartin multimii $\{(0, 0), (-3.499, -1.672), (3.499, 1.672), (-5.679, 2.461), (5.679, -2.461)\}$.

Analiza punctelor: consideram punctul $(0, 0)$.

- ▶ daca alegem (x_1, x_2) in jurul lui $(0, 0)$, cu $x_1 > 0$ si $x_2 = -2x_1$ obtinem $f(x_1, x_2) < 0 = f(0, 0)$
- ▶ daca alegem (x_1, x_2) in jurul lui $(0, 0)$, cu $x_2 = 0$, avem $f(x_1, x_2) > 0 = f(0, 0)$

Concluzie: $(0, 0)$ punct de inflexiune.

Ce tip sunt restul punctelor stationare? Cum putem analiza natura punctelor in cazul problemelor de mari dimensiuni?

Conditii necesare de ordinul II pentru (UNLP)

Dupa cum am vazut anterior, natura punctelor stationare nu poate fi stabilita utilizand doar informatie de ordinul I \Rightarrow pentru a stabili ca un punct stationar este punct de extrem (minim sau maxim) este necesara informatie despre Hessiana.

Teorema (Conditii necesare de ordinul II): Fie $f \in \mathcal{C}^2$ si $x^* \in \text{dom}f$ un punct de minim local al problemei (UNLP). Atunci avem:

$$\nabla^2 f(x^*) \succcurlyeq 0.$$

Conditii necesare de ordinul II - Demonstratie

Demonstratie: Presupunem prin absurd ca pentru un x^* punct de minim local avem ca $\nabla^2 f(x^*) \succcurlyeq 0$ nu este satisfacuta si drept urmare $\exists d$ a.i. $d^T \nabla^2 f(x^*) d < 0$.

Datorita continuitatii lui $\nabla^2 f \Rightarrow \exists t > 0$ a.i. $\forall \tau \in [0, t]$ avem:

$$d^T \nabla^2 f(x^* + \tau d) d < 0.$$

Utilizand Teorema lui Taylor $\Rightarrow \exists \theta \in [0, t]$ a.i.:

$$f(x^* + td) = f(x^*) + \underbrace{t \nabla f(x^*)^T d}_{\substack{= 0 \\ \nabla f(x^*) = 0}} + \frac{1}{2} t^2 \underbrace{d^T \nabla^2 f(x^* + \theta d) d}_{\leq 0} < f(x^*).$$

Ultima inegalitate contrazice deci ipoteza initiala conform careia x^* este punct de minim local.

Conditii suficiente de ordinul II pentru (UNLP)

Teorema (Conditii suficiente de ordinul II): Fie $f \in \mathcal{C}^2$ si $x^* \in \text{dom} f$ un punct stationar, i.e. $\nabla f(x^*) = 0$. Daca Hessiana lui f este pozitiv definita in x^* , i.e. $\nabla^2 f(x^*) \succ 0$, atunci x^* este punct strict de minim local al problemei (UNLP).

Conditii suficiente de ordinul II - Demonstratie

Demonstratie: Notand cu $\lambda_{\min} > 0$ valoarea proprie minima a Hessianei $\nabla^2 f(x^*)$ avem:

$$d^T \nabla^2 f(x^*) d \geq \lambda_{\min} \|d\|^2 \quad \forall d \in \mathbb{R}^n.$$

Folosind in continuare aproximarea Taylor putem scrie:

$$\begin{aligned} f(x^* + d) - f(x^*) &= \nabla f(x^*)^T d + \frac{1}{2} d^T \nabla^2 f(x^*) d + \mathcal{R}(\|d\|^2) \\ &\geq \frac{\lambda_{\min}}{2} \|d\|^2 + \mathcal{R}(\|d\|^2) = \left(\frac{\lambda_{\min}}{2} + \frac{\mathcal{R}(\|d\|^2)}{\|d\|^2} \right) \|d\|^2 \geq 0 \end{aligned}$$

pentru orice $d \in \mathbb{R}^n$ suficient de mic (reamintim ca avem relatia $\lim_{\|d\| \rightarrow 0} \frac{\mathcal{R}(\|d\|^2)}{\|d\|^2} = 0$). Deci x^* este punct de minim local.

Conditii de ordinul II - Tipul punctelor de extrem

Ca o consecinta a conditiilor suficiente de ordinul II, putem stabili natura unui punct stationar x^* in functie de proprietatile Hessienei, astfel:

- ▶ punct de minim local $\iff \nabla^2 f(x^*) \succ 0$
- ▶ punct de maxim local $\iff \nabla^2 f(x^*) \prec 0$
- ▶ punct de inflexiune $\iff \nabla^2 f(x^*)$ este indefinita (atat valori proprii negative cat si pozitive)
- ▶ in cazul in care $\nabla^2 f(x^*)$ este singulara nu se poate preciza natura punctului

Conditii de ordinul II - Exemplu

Consideram problema (UNLP):

$$\min_{x \in \mathbb{R}^2} f(x) \quad \left(= \frac{1}{2}x_1^2 + x_1x_2 + 2x_2^2 - 4x_1 - 4x_2 - x_2^3 \right)$$

Pentru a gasi punctele stationare rezolvam sistemul de doua ecuatii neliniare $\nabla f(x) = 0$:

$$x_1 + x_2 - 4 = 0 \quad \text{si} \quad x_1 + 4x_2 - 4 - 3x_2^2 = 0.$$

Obtinem ca multimea punctelor stationare ale problemei este $\{(4, 0), (3, 1)\}$.

Pentru a observa natura acestora evaluam Hessiana:

$$\nabla^2 f(x) = \begin{bmatrix} 1 & 1 \\ 1 & 4 - 6x_2 \end{bmatrix}.$$

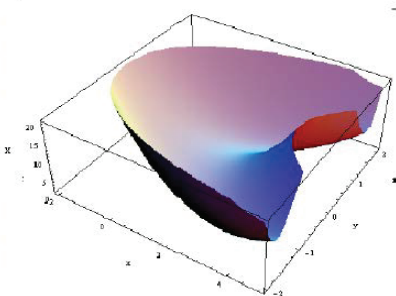
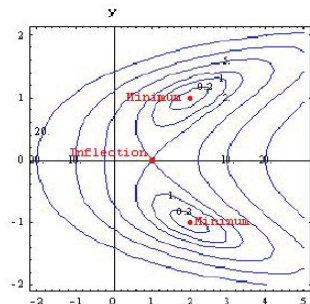
Deducem de aici ca $(4, 0)$ este punct de minim local iar $(3, 1)$ de inflexiune.

Conditii de ordinul II - Exemplul 2

Consideram problema (UNLP):

$$\min_{x \in \mathbb{R}^2} f(x) \quad (= (x_1 - 2)^2 + (x_1 - x_2^2)^2)$$

Din figura alaturata putem observa ca aceasta problema are doua puncte de minim si unul de inflexiune.



Cum putem gasi si stabili natura acestor puncte?

Conditii suficiente - Cazul convex

Sub presupunerea suplimentara ca f este o functie convexa vom arata ca in acest caz conditiile necesare de ordinul I sunt si suficiente:

Teorema (Conditii suficiente de ordinul I pentru cazul convex): Fie $f \in \mathcal{C}^1$ o **functie convexa**. Daca x^* este un punct stationar , i.e. $\nabla f(x^*) = 0$, atunci x^* este punct de minim global al problemei (UNLP).

Demonstratie: Folosind proprietatea de convexitate a lui f putem scrie:

$$f(x) \geq f(x^*) + \underbrace{\nabla f(x^*)^T}_{=0} (x - x^*) = f(x^*) \quad \forall x \in \mathbb{R}^n.$$

Concluzie: orice punct de minim local al unei probleme convexe este punct de minim global!

Probleme convexe sunt mult mai usor de rezolvat!

Cazul convex - Exemplu

Considerăm problema (UNLP) convexă:

$$\min_{x \in \mathbb{R}^2} f(x) \quad \left(= x_1^4 + 2x_1^2 - x_1x_2 + \frac{1}{2}x_2^2 \right).$$

În continuare obținem că $\nabla f(x)$ și $\nabla^2 f(x)$ au următoarele expresii:

$$\nabla f(x) = \begin{bmatrix} 4x_1^3 + 4x_1 - x_2 \\ -x_1 + x_2 \end{bmatrix} \quad \text{și} \quad \nabla^2 f(x) = \begin{bmatrix} 12x_1^2 + 4 & -1 \\ -1 & 1 \end{bmatrix}.$$

Observăm că $\nabla^2 f(x) \succ 0 \quad \forall x \in \mathbb{R}^2$ deci f este convexă.

Conform condițiilor suficiente de ordinul I, soluțiile sistemului $\nabla f(x) = 0$ sunt puncte de minim global ale problemei (UNLP) $\Rightarrow (0, 0)$ punct de minim global.