Aplicații ale diferențiabilității. Extreme și extreme cu legături

Cursul 10

Matematică - anul I

Facultatea de Informatică, UAIC

e-mail: adrian.zalinescu@info.uaic.ro

web: https://profs.info.uaic.ro/~adrian.zalinescu

13 Decembrie 2021

Sumarul cursului

Probleme de optimizare în spații euclidiene

- O posibilă aplicație a diferențiabilității: stabilirea punctelor de extrem ale unor funcții implicate în probleme de optimizare.
- Mai precis, aceste probleme urmăres minimizarea (sau maximizarea) unei *funcționale de cost* (*funcție de profit*), în prezența sau în absența unor anumite restricții.

Exemplul 1. Metoda celor mai mici pătrate

Să presupunem că experimentăm asupra unei cantități fizice:

- valori de “input”: $a_1, a_2, \ldots, a_p$;
- obținem valorile de “output” $b_1, b_2, \ldots, b_p$;
- reprezentăm punctele (a_k, b_k) $(k = \overline{1, p})$ în planul euclidian;
- putem estima (vizual) natura funcției necunoscute φ ce satisface $\varphi(a_k) = b_k$.
- Mai exact, vom presupune că φ este de un anumit tip (polinomial, exponențial, trigonometric, etc.), ale cărui caracteristici (parametri reali) $c_1, c_2, \ldots, c_n$ trebuie identificate.

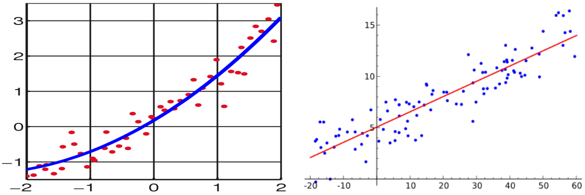


Figure: φ is quadratic, respectively linear

- Pentru aceasta, vom utiliza *the method of least squares*, *metoda celor mai mici pătrate*, prin considerarea problemei minimizării expresiei

$$\sum_{k=1}^{p} \left(\varphi(a_k; c_1, c_2, \ldots, c_n) - b_k\right)^2,$$

în raport cu $(c_1, c_2, \ldots, c_n) \in \mathbb{R}^n$.

- A rezolva această problemă (de extrem) implică găsirea $(c_1^0, c_2^0, \ldots, c_n^0) \in \mathbb{R}^n$ astfel încât

$$\min \left\{ \sum_{k=1}^{p} \left(\varphi(a_k; c_1, c_2, \ldots, c_n) - b_k\right)^2 \right\} = \sum_{k=1}^{p} \left(\varphi(a_k; c_1^0, c_2^0, \ldots, c_n^0) - b_k\right)^2.$$

- Putem conchide că mărimea fizică în studiu urmează legea $y = \varphi(x; c_1^0, c_2^0, \ldots, c_n^0)$.
- Dacă reprezentarea grafică a $\{(a_k, b_k) | k = \overline{1, p}\}$ sugrează că φ este liniară, putem lua $n = 2$ și $\varphi(x) := c_1 x + c_2$.
- În acest caz, metoda celor mai mici pătrate duce la minimizarea expresiei

$$\sum_{k=1}^{p} (c_1 a_k + c_2 - b_k)^2.$$

Exemplul 2. Realizarea unui profit maxim sau a unui cost minim

- În teoria economică, spațiul $\mathbb{R}^n$ este interpretat ca spațiul *bunurilor de consum*:
- fiecare *bun de consum* este identificat de un index $i \in \{1, 2, \ldots, n\}$;
- *vectorul bunurilor de consum* este $\mathrm{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, unde x_i este cantitatea în care bunul de consum i este produs.
- Într-un astfel de context, un *sistem de prețuri* este caracterizat de o funcție care asociază fiecărui vector de bunuri de consum o anumită valoare.
- Este natural să considerăm că o astfel de funcție este liniară, deci caracterizată de un vector $\mathrm{p} = (p_1, p_2, \ldots, p_n) \in \mathbb{R}^n$, unde p_i este prețul "unitar" al bunului de consum i; sistemul de prețuri este dat așadar de $\sum\limits_{i=1}^{n} p_i x_i$.

- Atunci când o companie producătoare produce anumite bunuri de consum, interesul acesteia este să realizeze producția astfel încât costul de producție să fie minimal și/sau profitul rezultat în urma producției să fie maximal.
- Economiștii folosesc o *funcție de cost* și/sau o *funcție de profit* alese convenabil, numite generic *funcții obiectiv*.
- În cazul în care spațiul bunurilor de consum este $\mathbb{R}^n$, avem de a face cu o *problemă de extrem necondiționat*.
- Dacă spațiul bunurilor de consum este $K \subsetneq \mathbb{R}^n$, vorbim de o *problemă de extrem condiționat*.
- În ambele cazuri, aceasta este o problemă de *programare liniară* dacă funcția obiectiv este liniară și mulțimea K este un subspațiu liniar al lui $\mathbb{R}^n$.
- Dacă funcția obiectiv este pătratică sau convexă, atunci problema se numește problemă de *optimizare pătratică*, respectiv *convexă*.

Exemplul 3. Entropia informațională

- *Entropia* a fost introdusă drept cantitate matematică de către Claude E. Shannon (1947).;
- este o funcție ce corespunde cantității de informație oferită de o anumită sursă, pe baza unui anumit limbaj, semnal electric sau fișier de date.
- Această funcție, notată H, este definită pe o mulțime de variabile aleatoare

$$X = \begin{pmatrix} 1 & 2 & \dots & n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}$$

și are expresia

$$H(X) = -\sum_{k=1}^{n} p_k \cdot \log_2 p_k,$$

unde p_k este probabilitatea ($p_k \in (0,1)$, $k = \overline{1,n}$ and $\sum_{k=1}^{n} p_k = 1$) ca un mesaj particular k să fie transmis de sursa amintită mai sus.

- O problemă este de a găsi distribuția optimală a unei variabile aleatoare X astfel încât valoarea $H(X)$ să fie maximală.
- Așadar, vom studia problema maximizării funcției

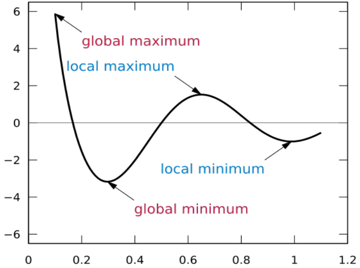
$$-\sum_{k=1}^{n} p_k \log_2 p_k$$

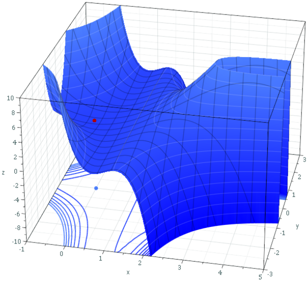
în raport cu $(p_1, p_2, \ldots, p_n)$, cu restricția $\sum_{k=1}^{n} p_k = 1$, $p_k \in (0, 1)$, $k = \overline{1, n}$.

Definition

Fie $\emptyset \neq A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}$ o funcție și $x_0 \in A$. Spunem că x_0 este:

- un *punct de minim* al lui f dacă $f(x) \geq f(x_0)$, $\forall x \in A$;
- un *punct de maxim* al lui f dacă $f(x) \leq f(x_0)$, $\forall x \in A$;
- un *punct de minim strict* al lui f dacă $f(x) > f(x_0)$, $\forall x \in A \setminus \{x_0\}$;
- un *punct de maxim strict* al lui f dacă $f(x) < f(x_0)$, $\forall x \in A \setminus \{x_0\}$;
- un *punct de minim local* al lui f dacă există $V \in \mathscr{V}(x_0)$ astfel încât $f(x) \geq f(x_0)$, $\forall x \in A \cap V$;
- un *punct de maxim local* al lui f dacă există $V \in \mathscr{V}(x_0)$ astfel încât $f(x) \leq f(x_0)$, $\forall x \in A \cap V$;
- un *punct de minim local strict* al lui f dacă există $V \in \mathscr{V}(x_0)$ astfel încât $f(x) > f(x_0)$, $\forall x \in A \cap (V \setminus \{x_0\})$;
- un *punct de maxim local strict* al lui f dacă există $V \in \mathscr{V}(x_0)$ astfel încât $f(x) < f(x_0)$, $\forall x \in A \cap (V \setminus \{x_0\})$;
- un *punct de extrem* (*extrem strict*, *extrem local*, *extrem local strict*) al lui f dacă x_0 este punct de minim (minim strict, minim local, minim local strict) sau punct de maxim (maxim strict, maxim local, respectiv maxim local strict) al lui f.





Observații. Un punct de extrem global al lui f este întotdeauna un punct de extrem local pentru f. Reciproca nu este adevărată.

1. Fie $D \subseteq \mathbb{R}^n$ o mulțime deschisă și $f : D \to \mathbb{R}$ o funcție. Problema determinării extremelor globale și/sau locale (și a valorilor extreme asociate) ale lui f se numește o *problemă de extrem necondiționat*.

2. Fie $D \subseteq \mathbb{R}^n$ o mulțime deschisă, $f : D \to \mathbb{R}$ o funcție și $A \subseteq \mathbb{R}^m$ o mulțime de restricții. Problema determinării extremelor globale și/sau locale (și a valorilor extreme asociate) ale lui $f\big|_{A \cap D}$ se numește o *problemă de extrem condiționat*.

Teorema lui Fermat

Teoremă

Fie $A \subseteq \mathbb{R}^n$ o mulțime nevidă, $f : A \to \mathbb{R}$ o funcție și x_0 un punct interior lui A. Dacă x_0 este un punct de extrem local al lui f și f admite derivate parțiale (de ordinul întâi) în x_0, atunci

$$\frac{\partial f}{\partial x_1}(\mathrm{x}_0) = \cdots = \frac{\partial f}{\partial x_n}(\mathrm{x}_0) = 0.$$

Observație. Concluzia teoremei poate fi scrisă $(\nabla f)(\mathrm{x}_0) = 0_{\mathbb{R}^n}$.
Dacă f este în plus Fréchet diferențiabilă în x_0, acest lucru este de asemenea echivalent cu

$$\mathrm{d}f(\mathrm{x}_0) = 0_{\mathrm{L}(\mathbb{R}^n;\mathbb{R})}.$$

Definiție

Fie $A \subseteq \mathbb{R}^n$, $\mathrm{x}_0 \in \mathring{A}$ și $f : A \to \mathbb{R}$ o funcție Fréchet diferențiabilă în x_0. Elementul x_0 se numește *punct critic* (sau *punct staționar*) al lui f dacă $\mathrm{d}f(\mathrm{x}_0) = 0_{\mathrm{L}(\mathbb{R}^n;\mathbb{R})}$ (*adică* $(\nabla f)(\mathrm{x}_0) = 0_{\mathbb{R}^n}$).

Remarcă. Teorema lui Fermat afirmă că dacă $x_0 \in \mathring{A}$ este un punct de extrem local al unei funcții f in x_0, atunci x_0 este un punct critic al lui f. Reciproca nu este adevărată:
Exemplu. Fie $f : \mathbb{R}^2 \to \mathbb{R}$,

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

- Atunci $(0, 0)$ este un punct critic al lui f, deoarece $\frac{\partial f}{\partial x}(0, 0) = 2x\big|_{x=0,y=0} = 0$ și $\frac{\partial f}{\partial y}(0, 0) = -2y\big|_{x=0,y=0} = 0$.
- Pe de altă parte, $f(x, 0) = x^2 > f(0, 0) = 0 > -y^2 = f(0, y)$ pentru orice $x \neq 0$ și $y \neq 0$. Deci $(0, 0)$ nu este nici punct de minim local, nici punct de maxim local pentru f.

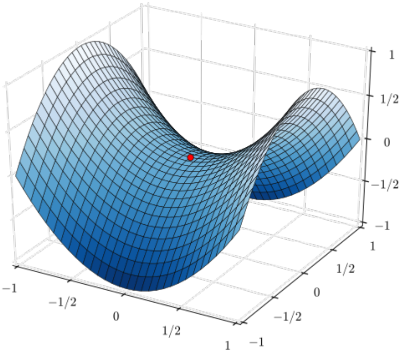


Figure: graph of $f(x, y) = x^2 - y^2$

Definiție

Fie $A \subseteq \mathbb{R}^n$, $x_0 \in \mathring{A}$ și $f : A \to \mathbb{R}$ o funcție Fréchet diferențiabilă în x_0. Dacă x_0 este un punct critic al lui f, dar nu este un punct de extrem local al lui f, spunem că x_0 este un *punct șa* al lui f.

Condiții suficiente de extrem

Cazul n=1

Teoremă

Fie $a, b \in \mathbb{R}$ cu $a < b$, $x_0 \in (a, b)$ și $f : (a, b) \to \mathbb{R}$ o funcție de n ori ($n \geq 2$) derivabilă în x_0. Să presupunem că $f'(x_0) = f''(x_0) = \ldots = f^{(n-1)}(x_0) = 0$ și $f^{(n)}(x_0) \neq 0$.

i) Dacă n este par, atunci x_0 este un punct de extrem local al lui f. Mai precis, dacă $f^{(n)}(x_0) > 0$, atunci x_0 este punct de minim local al lui f, iar dacă $f^{(n)}(x_0) < 0$ atunci x_0 este punct de maxim local al lui f.

ii) Dacă n este impar, atunci x_0 nu este un punct de extrem local al lui f.

Cazul n>1

Stabilim dacă un punct critic x_0 al lui f este punct de extrem local prin analiza derivatei Fréchet de ordinul 2 în x_0:

$$\left(\mathrm{d}^2 f(\mathrm{x}_0)\right)(\mathrm{v}) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathrm{x}_0) v_i v_j, \forall \mathrm{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n.$$

Teoremă

Fie $\emptyset \neq A \subseteq \mathbb{R}^n$ și $f : A \to \mathbb{R}$ o funcție de clasă C^2 într-o vecinătate a unui punct critic $\mathrm{x}_0 \in \mathring{A}$.

i) Dacă $\left(\mathrm{d}^2 f(\mathrm{x}_0)\right)(\mathrm{v}) > 0$, $\forall \mathrm{v} \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$, atunci x_0 este un punct de minim local al lui f.

ii) Dacă $\left(\mathrm{d}^2 f(\mathrm{x}_0)\right)(\mathrm{v}) < 0$, $\forall \mathrm{v} \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$, atunci x_0 este un punct de maxim local al lui f.

iii) Dacă $\mathrm{d}^2 f(\mathrm{x}_0)$ este o formă pătratică nedefinită (adică $\exists \mathrm{v}', \mathrm{v}'' \in \mathbb{R}^n$ astfel încât $\left(\mathrm{d}^2 f(\mathrm{x}_0)\right)(\mathrm{v}') < 0$ și $\left(\mathrm{d}^2 f(\mathrm{x}_0)\right)(\mathrm{v}'') > 0$), atunci x_0 este un punct șa pentru f.

Observații.
1. Dacă $\mathrm{d}^2 f(\mathrm{x}_0)$ este o formă pătratică pozitiv sau negativ-semidefinită (adică $\left(\mathrm{d}^2 f(\mathrm{x}_0)\right)(\mathrm{v}) \geq 0$, $\forall \mathrm{v} \in \mathbb{R}^n$ sau $\left(\mathrm{d}^2 f(\mathrm{x}_0)\right)(\mathrm{v}) \leq 0$, $\forall \mathrm{v} \in \mathbb{R}^n$, respectiv și $\exists \mathrm{v} \in \mathbb{R}^n$ astfel încât $\left(\mathrm{d}^2 f(\mathrm{x}_0)\right)(\mathrm{v}') = 0$), atunci nu putem determina natura punctului critic.

2. Din ipoteze, avem (teorema lui Schwarz): $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(\mathrm{x}_0) = \dfrac{\partial^2 f}{\partial x_j \partial x_i}(\mathrm{x}_0)$, $\forall i \neq j$, deci matricea asociată acestei forme pătratice este matricea simetrică

$$H_f(\mathrm{x}_0) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathrm{x}_0)\right)_{1 \leq i,j \leq n},$$

numită *Hessianul* lui f.

Din cursul 7, avem mai multe metode de a determina dacă $\mathrm{d}^2 f(x_0)$ este pozitiv definită, negativ definită, nedefinită sau niciuna dintre acestea.

Propoziție

Fie $\emptyset \neq A \subseteq \mathbb{R}^n$ și $f : A \to \mathbb{R}$ o funcție C^2 într-o vecinătate a unui punct critic $x_0 \in \mathring{A}$.

i) Dacă toate valorile proprii ale lui $H_f(x_0)$ sunt strict pozitive, atunci x_0 este un punct de minim local al lui f.

ii) Dacă toate valorile proprii ale lui $H_f(x_0)$ sunt strict negative, atunci x_0 este un punct de maxim local al lui f.

iii) Dacă $H_f(x_0)$ are cel puțin o valoare proprie strict negativă și una strict pozitivă, atunci x_0 este un punct șa al lui f.

Bineînțeles, nu putem spune nimic în cazul în care toate valorile proprii sunt doar pozitive sau doar negative.

Propoziție

Fie $\emptyset \neq A \subseteq \mathbb{R}^n$ și $f : A \to \mathbb{R}$ o funcție C^2 într-o vecinătate a unui punct critic $x_0 \in \mathring{A}$. Fie $\Delta_k = \det \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) \right)_{1 \leq i,j \leq k}$, $k = \overline{1, n}$, minorii principali ai lui $H_f(x_0)$.

i) Dacă $\Delta_k > 0$, $\forall k = \overline{1, n}$, atunci x_0 este un punct de minim local al lui f.

ii) Dacă $(-1)^{k+1}\Delta_k < 0$, $\forall k = \overline{1, n}$, atunci x_0 este un punct de maxim local al lui f.

iii) Dacă există $j, k = \overline{1, n}$ astfel încât $\Delta_j < 0$ și $(-1)^{k+1}\Delta_k > 0$, atunci x_0 este un punct șa al lui f.

Din nou, nu putem spune nimic în cazul în care $\Delta_k \geq 0$, $\forall k = \overline{1, n}$ sau $(-1)^{k+1}\Delta_k \leq 0$, $\forall k = \overline{1, n}$.

Cazul n=2

Remarcă. În cazul $n = 2$, rezultatul de mai sus afirmă că dacă $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$ este C^2 într-o vecinătate a unui punct critic $\mathrm{x}_0 \in \mathring{A}$ și notăm $p := \dfrac{\partial^2 f}{\partial x^2}(\mathrm{x}_0)$, $q := \dfrac{\partial^2 f}{\partial x \partial y}(\mathrm{x}_0)$, $r := \dfrac{\partial^2 f}{\partial y^2}(\mathrm{x}_0)$, atunci:

i) dacă $p > 0$ și $pr - q^2 > 0$, x_0 este un punct de minim local al lui f;

ii) dacă $p < 0$ și $pr - q^2 > 0$, x_0 este un punct de maxim local al lui f;

iii) dacă $pr - q^2 < 0$, x_0 este punct șa al lui f;

iv) dacă $pr - q^2 = 0$, nu putem stabili natura lui x_0 prin această metodă.

Metoda celor mai mici pătrate (continuare)

Revenind la metoda celor mai mici pătrate, cu φ liniară, observăm că funcția $f : \mathbb{R}^2 \to \mathbb{R}$, definită prin

$$f(c_1, c_2) = \sum_{k=1}^{\ell} (c_1 a_k + c_2 - b_k)^2$$

este de clasă C^2 pe $\mathbb{R}^2$ și are ca puncte critice elementele $\left(c_1^0, c_2^0\right) \in \mathbb{R}^2$ ce satisfac

$$\begin{cases} \dfrac{\partial f}{\partial c_1}\left(c_1^0, c_2^0\right) = 2\displaystyle\sum_{k=1}^{\ell} (c_1^0 a_k + c_2^0 - b_k) a_k = 0; \\ \dfrac{\partial f}{\partial c_2}\left(c_1^0, c_2^0\right) = 2\displaystyle\sum_{k=1}^{\ell} (c_1^0 a_k + c_2^0 - b_k) = 0. \end{cases}$$

Pentru a rezolva această ecuație, echivalentă cu sistemul liniar

$$\begin{cases} c_1^0 \displaystyle\sum_{k=1}^{\ell} a_k^2 + c_2^0 \sum_{k=1}^{\ell} a_k = \sum_{k=1}^{\ell} b_k a_k; \\ c_1^0 \displaystyle\sum_{k=1}^{\ell} a_k^2 + \ell c_2^0 = \sum_{k=1}^{\ell} b_k, \end{cases}$$

calculăm determinantul sistemului,

$$\begin{vmatrix} \sum\limits_{k=1}^{\ell} a_k^2 & \sum\limits_{k=1}^{\ell} a_k \\ \sum\limits_{k=1}^{\ell} a_k & \ell \end{vmatrix} = \ell \sum_{k=1}^{\ell} a_k^2 - \left(\sum_{k=1}^{\ell} a_k \right)^2 ,$$

care este $\neq 0$ dacă $a_1, a_2, \ldots, a_\ell$ nu sunt toți egali, din inegalitatea Cauchy-Buniakowski-Schwarz.
În acest caz, avem

$$c_1^0 = \frac{\ell \sum_{k=1}^{\ell} a_k b_k - \left(\sum_{k=1}^{\ell} a_k \right) \left(\sum_{k=1}^{\ell} b_k \right)}{\ell \sum_{k=1}^{\ell} a_k^2 - \left(\sum_{k=1}^{\ell} a_k \right)^2}$$

și

$$c_2^0 = \frac{\left(\sum_{k=1}^{\ell} b_k \right) \left(\sum_{k=1}^{\ell} a_k^2 \right) - \left(\sum_{k=1}^{\ell} a_k \right) \left(\sum_{k=1}^{\ell} a_k b_k \right)}{\ell \sum_{k=1}^{\ell} a_k^2 - \left(\sum_{k=1}^{\ell} a_k \right)^2}.$$

Atunci

- $p = \dfrac{\partial^2 f}{\partial c_1^2}\left(c_1^0, c_2^0\right) = 2\sum_{k=1}^{\ell} a_k^2$,
- $q = \dfrac{\partial^2 f}{\partial c_1 \partial c_2}\left(c_1^0, c_2^0\right) = 2\sum_{k=1}^{\ell} a_k$ and
- $r = \dfrac{\partial^2 f}{\partial c_2^2}\left(c_1^0, c_2^0\right) = 2\ell$.
- Din nou din inegalitatea Cauchy-Buniakowski-Schwarz, avem $p > 0$ și $pr - q^2 > 0$, deci $\left(c_1^0, c_2^0\right)$ este un punct de minim global pentru f (nu doar local, deoarece $\left(c_1^0, c_2^0\right)$ este unicul punct critic al lui f).

Extreme cu legături

În practică apar probleme în care trebuie să calculăm extreme ale unor funcții ale căror argumente sunt supuse unor anumite condiții (restricții). În acest caz, respectivele extreme sunt numite *extreme cu legături*.

Cadrul de lucru. Fie D o mulțime deschisă în $\mathbb{R}^n \times \mathbb{R}^m$ $(n, m \in \mathbb{N}^*)$ și $f : D \to \mathbb{R}$, $g : D \to \mathbb{R}^m$ funcții de clasă C^1 pe D.

- Căutăm puncte de extrem ale funcției f restricționată la condiția suplimentară $g(\mathrm{x}, \mathrm{y}) = 0$, unde $(\mathrm{x}, \mathrm{y}) \in D$.
- Notând $A := \{(\mathrm{x}, \mathrm{y}) \in D | g(\mathrm{x}, \mathrm{y}) = 0_{\mathbb{R}^n}\}$ și prin $g_1, g_2, \ldots, g_m$ componentele lui g, putem vedea că problema de mai sus cere aflarea punctelor de extrem ale restricției $f\big|_A$.
- *Terminologie:* spunem că un punct $(\mathrm{x}_0, \mathrm{y}_0) \in A$ este un *punct de extrem (minim, maxim) local* sau *punct șa* al lui f, *cu legăturile* $g_1(\mathrm{x}, \mathrm{y}) = 0$, $g_2(\mathrm{x}, \mathrm{y}) = 0$, …, $g_m(\mathrm{x}, \mathrm{y}) = 0$, dacă este punct de extrem (minim, maxim) local sau punct șa al lui $f\big|_A$.

Metoda multiplicatorilor lui Lagrange

Teoremă

Fie $\emptyset \neq D \subseteq \mathbb{R}^n \times \mathbb{R}^m$ o mulțime deschisă, $f : D \to \mathbb{R}$ și $g : D \to \mathbb{R}^m$ funcții de clasă C^1. Fie $(x_0, y_0) \in D$ un punct de extrem al lui f, cu legăturile $g_1(x, y) = 0$, $g_2(x, y) = 0$, …, $g_m(x, y) = 0$. Dacă $\frac{D(g_1, g_2, \ldots, g_m)}{D(y_1, y_2, \ldots, y_m)}(x_0, y_0) \neq 0$, atunci există $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{R}$ astfel încât (x_0, y_0) este un punct critic pentru funcția de clasă C^1, $L : D \to \mathbb{R}$, definită de

$$L(x, y) = f(x, y) + \lambda_1 g_1(x, y) + \lambda_2 g_2(x, y) + \cdots + \lambda_m g_m(x, y), \ (x, y) \in D,$$

adică

$$(\star) \begin{cases} \dfrac{\partial L}{\partial x_k}(x_0, y_0) = 0, \ \forall k = \overline{1, n}; \\ \dfrac{\partial L}{\partial y_j}(x_0, y_0) = 0, \ \forall j = \overline{1, m}; \\ g_j(x_0, y_0) = 0, \ \forall j = \overline{1, m}. \end{cases}$$

Observații.

1. Numerele $\lambda_1, \lambda_2, \ldots, \lambda_m$ se numesc *multiplicatorii lui Lagrange*.

2. Concluzia rezultatului de mai sus poate fi reformulată astfel: punctele de extrem local condiționat ale lui f sunt printre punctele critice ale funcției asociate $L = f + \sum_{k=1}^{m} \lambda_k g_k$.

3. Funcția $\mathscr{L} : \mathbb{R}^m \times D \to \mathbb{R}$ definită de

$$\mathscr{L}(\boldsymbol{\lambda}, \mathrm{x}, \mathrm{y}) = f(\mathrm{x}, \mathrm{y}) + \lambda_1 g_1(\mathrm{x}, \mathrm{y}) + \lambda_2 g_2(\mathrm{x}, \mathrm{y}) + \cdots + \lambda_m g_m(\mathrm{x}, \mathrm{y}),$$

pentru $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$, se numește *Lagrangianul* asociat lui f și restricției g.
Atunci $(\boldsymbol{\lambda}_0, \mathrm{x}_0, \mathrm{y}_0)$ satisface sistemul $(\star)$ (adică $(\mathrm{x}_0, \mathrm{y}_0)$ este un punct critic pentru L, ce satisface în plus restricția $g(\mathrm{x}_0, \mathrm{y}_0) = 0_{\mathbb{R}^m}$) dacă și numai dacă $(\boldsymbol{\lambda}_0, \mathrm{x}_0, \mathrm{y}_0)$ este un punct critic pentru $\mathscr{L}$, deoarece

$$\frac{\partial \mathscr{L}}{\partial \lambda_j}(\boldsymbol{\lambda}_0, \mathrm{x}_0, \mathrm{y}_0) = g_j(\mathrm{x}_0, \mathrm{y}_0),\ j = \overline{1, m}.$$

4. Să presupunem acum că $(\boldsymbol{\lambda}_0, \mathrm{x}_0, \mathrm{y}_0)$ satisface sistemul $(\star)$ și f, g sunt funcții de clasă C^2 (sau măcar pe o vecinătate a lui $(\boldsymbol{\lambda}_0, \mathrm{x}_0, \mathrm{y}_0)$). Deoarece

$$L(\mathrm{x}, \mathrm{y}) - L(\mathrm{x}_0, \mathrm{y}_0) = f(\mathrm{x}, \mathrm{y}) - f(\mathrm{x}_0, \mathrm{y}_0), \ \forall (\mathrm{x}, \mathrm{y}) \in A,$$

vedem că $(\mathrm{x}_0, \mathrm{y}_0)$ un punct de extrem local condiționat al lui f dacă și numai dacă $(\mathrm{x}_0, \mathrm{y}_0)$ este un punct de extrem local al funcției $L\big|_A$.
O modalitate practică de a determina condiții suficiente de extremalitate condițională locală este următoarea:
I. Fie $\mathrm{x} = (\tilde{x}_1, \ldots, \tilde{x}_n) \in \mathbb{R}^n$ și $\mathrm{y} = (\tilde{x}_{n+1}, \ldots, \tilde{x}_{n+m}) \in \mathbb{R}^m$; diferențiala Fréchet de ordinul 2 a lui L in $(\mathrm{x}_0, \mathrm{y}_0)$ este

$$(\star\star) \quad \sum_{i,j=1}^{n+m} \frac{\partial^2 L}{\partial \tilde{x}_i \partial \tilde{x}_j} (\mathrm{x}_0, \mathrm{y}_0) \mathrm{d}\tilde{x}_i \mathrm{d}\tilde{x}_j.$$

II. Dacă forma pătratică de mai sus este pozitiv definită sau negativ definită (aplicăm una din metodele de la extreme fără legături), putem concluziona că $(\mathrm{x}_0, \mathrm{y}_0)$ este este un punct de minim local condiționat, respectiv punct de maxim local condiționat al lui f.
Dacă nu, mergem mai departe.

III. Considerând sistemul obținut din diferențierea relațiilor $g_k(\mathrm{x}, \mathrm{y}) = 0$, $k = \overline{1, m}$, adică

$$\sum_{j=1}^{n+m} \frac{\partial g_k}{\partial \tilde{x}_j}(\mathrm{x}_0, \mathrm{y}_0)\mathrm{d}\tilde{x}_j = 0$$

putem scrie pe $\mathrm{d}\tilde{x}_{n+1}, \ldots, \mathrm{d}\tilde{x}_{n+m}$ în funcție de $\mathrm{d}\tilde{x}_1, \ldots, \mathrm{d}\tilde{x}_n$ și înlocuite în $(\star\star)$. Acest lucru transformă forma pătratică $(\star\star)$ într-o formă pătratică de tipul

$$\mathrm{d}^2 L(\mathrm{x}_0) = \sum_{i,j=1}^{n} a_{ij}\mathrm{d}\tilde{x}_i\mathrm{d}\tilde{x}_j.$$

- dacă $\mathrm{d}^2 L(\mathrm{x}_0)$ este pozitiv definită sau negativ definită, atunci $(\mathrm{x}_0, \mathrm{y}_0)$ este este un punct de minim local condiționat, respectiv punct de maxim local condiționat al lui f.
- dacă $\mathrm{d}^2 L(\mathrm{x}_0)$ este nedefinită, atunci $(\mathrm{x}_0, \mathrm{y}_0)$ este un punct șa condiționat pentru f;
- dacă $\mathrm{d}^2 L(\mathrm{x}_0)$ este doar pozitiv sau negativ semi-definită, nu putem decide asupra naturii lui $(\mathrm{x}_0, \mathrm{y}_0)$.

Entropia informațională (continuare)

Aplicăm metoda de mai sus pentru găsirea extremelor entropiei lui Shannon,

$$\tilde{H}(p_1, p_2, \dots, p_n) = -\sum_{k=1}^{n} p_k \log_2 p_k,\ \ p_k \in (0,1),\ \forall k = \overline{1,n},$$

condiționate de legătura $\sum_{k=1}^{n} p_k = 1$. Avem $m = 1$,
$g(p_1, p_2, \dots, p_n) := \sum_{k=1}^{n} p_k - 1$ și Lagrangianul este

$$\mathscr{L}(\lambda_1, p_1, p_2, \dots, p_n) = -\sum_{k=1}^{n} p_k \log_2 p_k + \lambda_1 \left(\sum_{k=1}^{n} p_k - 1 \right).$$

Atunci $(\lambda_1^0, p_1^0, p_2^0, \dots, p_n^0)$ este un punct critic pentru $\mathscr{L}$ dacă

$$\begin{cases} \dfrac{\partial \mathscr{L}}{\partial p_k}(\lambda_1^0, p_1^0, p_2^0, \dots, p_n^0) = -\left(\log_2 p_k^0 + \frac{1}{\ln 2} \right) + \lambda_1^0 = 0,\ \forall k = \overline{1,n} \\ \dfrac{\partial \mathscr{L}}{\partial \lambda_1}(\lambda_1^0, p_1^0, p_2^0, \dots, p_n^0) = \displaystyle\sum_{k=1}^{n} p_k^0 - 1 = 0. \end{cases}$$

Din primele n ecuații găsim $p_1^0 = p_2^0 = \ldots = p_n^0 = \psi(\lambda_1^0)$, unde $\psi(\lambda)$ este unica soluție a ecuației în $p \in (0,1)$: $\log_2 p + \frac{1}{\ln 2} = \lambda$, în timp ce din ultima găsim

$$p_1^0 = p_2^0 = \cdots = p_n^0 = \frac{1}{n},\ \lambda_1^0 = \log_2 \frac{1}{n} + \frac{1}{\ln 2}.$$

Avem deci

$$\mathrm{d}^2 L(p_1^0, p_2^0, \ldots, p_n^0) = \sum_{k=1}^{n} -\frac{1}{p_k^0 \ln 2} \left(\mathrm{d}p_k\right)^2,$$

care este o formă pătratică negativ definită.
Deducem că punctul $(p_1^0, p_2^0, \ldots, p_n^0) = \left(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\right)$ este punct de maxim (chiar global) pentru $\tilde{H}$, condiționat de $p_1 + p_2 + \cdots + p_n = 1$.
Astfel, entropia informațională maximă este obținută pentru variabila aleatoare dată de

$$X = \begin{pmatrix} 1 & 2 & \ldots & n \\ \frac{1}{n} & \frac{1}{n} & \ldots & \frac{1}{n} \end{pmatrix}.$$

- C. Canuto, A. Tabacco, *Mathematical Analysis II*, Springer, 2015.
- B. Edwards, R. Larson, *Calculus (9th Edition)*, Brooks/Cole-Cengage Learning, 2014.
- K. P. Evans, N. Jacob, *A Course in Analysis*, World Scientific, 2016.
- M. Gorunescu, *Lecții de analiză matematică pentru informaticieni*, Reprografia Universității Craiova, 2000.
- F. Iacob, *Matematică pentru anul II - ID*, seria 2004-2005.
- R. Luca-Tudorache, *Analiză matematică*, Editura Tehnopress, Iași, 2005.
- G. Păltineanu, *Analiză matematică. Calcul diferențial*, Editura AGIR, București, 2002.
- E. Popescu, *Analiză matematică. Calcul diferențial*, Editura Matrix Rom, București, 2006.
- M. Postolache, *Analiză matematică (teorie și aplicații)*, Ed. Fair Partners, București, 2011.
- V. Postolică, *Eficiență prin matematică aplicată. Analiză matematică*, Editura Matrix Rom, București, 2006.