CURSUL 10
APLICAŢII ALE DIFERENŢIABILITĂŢII. EXTREME ŞI EXTREME CU LEGĂTURI

Una dintre multiplele aplicaţii ale diferenţiabilităţii funcţiilor este stabilirea punctelor de extrem pentru funcţii ce apar în probleme de optimizare. Mai exact, aceste probleme urmăresc minimizarea (sau maximizarea) unor aşa-zise *funcţionale de cost* (*funcţii de profit*), în prezenţa sau absenţa unor anumite restricţii.

După ce vom oferi câteva exemple, vom expune aspectele teoretice de care avem nevoie pentru a aborda astfel de probleme.

1. Probleme de optimizare în $\mathbb{R}^n$

Exemplul 1. ***Metoda celor mai mici pătrate***

Să presupunem că am determinat experimental valorile unei anumite mărimi fizice, notate cu $b_1, b_2, \ldots, b_p$, corespunzând valorilor de "input" $a_1, a_2, \ldots, a_p$ ($p \in \mathbb{N}^*$). Mai presupunem că determinarea precisă a acestor valori este supusă erorilor de măsurătoare. Dacă reprezentăm punctele (a_k, b_k) ($k = \overline{1,p}$) într-un sistem ortogonal de referinţă în plan, putem estima (vizual) natura unei funcţii necunoscute φ care satisface, cel puţin din punct de vedere teoretic, pentru orice $k = \overline{1,p}$, $\varphi(a_k) = b_k$. Mai exact, presupunem că funcţia φ are un anumit tip (polinomial, exponenţial, trigonometric, etc.), ai cărei parametri (caracteristici) $c_1, c_2, \ldots, c_n$ trebuie să fie identificaţi. În acest scop, utilizăm *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 $(c_1, c_2, \ldots, c_n) \in \mathbb{R}^n$. Rezolvând această problemă (de extrem), adică găsind $(c_1^0, c_2^0, \ldots, c_n^0) \in \mathbb{R}^n$ 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 conclude că mărimea fizică în studiu urmează legea $y = \varphi(x; c_1^0, c_2^0, \ldots, c_n^0)$.

Remarcăm că dacă reprezentarea grafică a mulţimii $\{(a_k, b_k) | k = \overline{1,p}\}$ ne sugerează că φ este liniară, atunci putem lua $n = 2$ şi $\varphi(x) := c_1 x + c_2$. În acest caz, metoda celor mai mici pătrate implică mimizarea expresiei

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

în raport cu (c_1, c_2).

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

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

Atunci când o companie producătoare produce anumite bunuri de consum, interesul acesteia este să realizeze producţia astfel încât costul de producţie să fie minimal şi/sau profitul rezultat în urma producţiei să fie maximal. În acest scop, economiştii folosesc o *funcţie de cost* şi/sau o *funcţie de profit* alese convenabil, şi numite *funcţii obiectiv.*

În cazul în care spaţiul bunurilor de consum este $\mathbb{R}^n$, avem de a face cu o *problemă de extrem necondiţionat.* Dacă acest spaţiu este în schimb $K \subsetneq \mathbb{R}^n$ atunci avem de a face cu o *problemă de extrem condiţionat.* În ambele cazuri, aceasta este o problemă de *programare liniară* dacă funcţia obiectiv este liniară şi mulţimea K este un subspaţiu liniar al lui $\mathbb{R}^n$. Dacă funcţia obiectiv este pătratică sau convexă, atunci problema se numeşte problemă de *optimizare pătratică*, respectiv *convexă.*

Exemplul 3. ***Entropia informaţională***

Entropia a fost introdusă drept cantitate matematică de către Claude E. Shannon (1947). Este o funcţie ce corespunde cantităţii de informaţie oferită de o anumită sursă, pe baza unui anumit limbaj, semnal electric sau fişier de date. Această funcţie, notată H, este definită pe o mulţime de variabile aleatoare, de tipul

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

şi are expresia

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

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

O problemă este de a găsi distribuţia optimală a unei variabile aleatoare X astfel încât valoarea $H(X)$ să fie maximală. Cu alte cuvinte, vom studia problema maximizării funcţiei

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

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

2. Extreme necondiţionate

Reamintim câteva definiţii legate de spaţii metrice, de care vom avea nevoie în continuare. Fie (X, d) un spaţiu metric.

1. *Bila deschisă centrată* în $x_0 \in X$, de *rază* $r > 0$, este mulţimea
$$B(x_0; r) := \{x \in X \mid \mathrm{d}(x, x_0) < r\}.$$
2. O *vecinătate* a unui punct $x_0 \in X$ este o submulţime $V \subseteq X$ pentru care există $\varepsilon > 0$ astfel încât $B(x_0; \varepsilon) \subseteq V$.
3. Un punct $x_0 \in X$ se numeşte *punct interior* al unei submulţimi $A \subseteq X$ dacă A este vecinătate a lui x.
4. O submulţime $A \subseteq X$ se numeşte *deschisă* dacă orice element al său îi este punct interior.

Definiţie. Fie $A \subseteq \mathbb{R}^n$ o mulţime nevidă, $f : A \to \mathbb{R}$ o funcţie şi $\mathbf{x}_0 \in A$. Spunem că $\mathbf{x}_0$ este:

a) un *punct de minim* al lui f dacă $f(\mathbf{x}) \geq f(\mathbf{x}_0)$, $\forall \mathbf{x} \in A$;
b) un *punct de maxim* al lui f dacă $f(\mathbf{x}) \leq f(\mathbf{x}_0)$, $\forall \mathbf{x} \in A$;
c) un *punct de minim strict* al lui f dacă $f(\mathbf{x}) > f(\mathbf{x}_0)$, $\forall \mathbf{x} \in A \setminus \{\mathbf{x}_0\}$;
d) un *punct de maxim strict* al lui f dacă $f(\mathbf{x}) < f(\mathbf{x}_0)$, $\forall \mathbf{x} \in A \setminus \{\mathbf{x}_0\}$;
e) un *punct de minim local* al lui f dacă există o vecinătate V a lui $\mathbf{x}_0$ astfel încât $f(\mathbf{x}) \geq f(\mathbf{x}_0)$, $\forall \mathbf{x} \in A \cap V$;
f) un *punct de maxim local* al lui f dacă există o vecinătate V a lui $\mathbf{x}_0$ astfel încât $f(\mathbf{x}) \leq f(\mathbf{x}_0)$, $\forall \mathbf{x} \in A \cap V$;
g) un *punct de minim local strict* al lui f dacă există o vecinătate V a lui $\mathbf{x}_0$ astfel încât $f(\mathbf{x}) > f(\mathbf{x}_0)$, $\forall \mathbf{x} \in A \cap (V \setminus \{\mathbf{x}_0\})$;
h) un *punct de maxim local strict* al lui f dacă există o vecinătate V a lui $\mathbf{x}_0$ astfel încât $f(\mathbf{x}) < f(\mathbf{x}_0)$, $\forall \mathbf{x} \in A \cap (V \setminus \{\mathbf{x}_0\})$;
i) un *punct de extrem* (*extrem strict*, *extrem local*, *extrem local strict*) al lui f dacă $\mathbf{x}_0$ 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.

Observaţie. 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 mulţime deschisă şi $f : D \to \mathbb{R}$ o funcţie. Problema determinării extremelor globale şi/sau locale (şi a valorilor extreme asociate) ale lui f se numeşte o *problemă de extrem necondiţionat.*
2. Fie $D \subseteq \mathbb{R}^n$ o mulţime deschisă, $f : D \to \mathbb{R}$ o funcţie şi $A \subseteq \mathbb{R}^m$ o mulţime de restricţii. Problema determinării extremelor globale şi/sau locale (şi a valorilor extreme asociate) ale lui $f|_{A \cap D}$ se numeşte o *problemă de extrem condiţionat.*

Teorema 2.1 (Fermat). *Fie $A \subseteq \mathbb{R}^n$ o mulţime oarecare, $f : A \to \mathbb{R}$ o funcţie şi $\mathbf{x}_0$ un punct interior lui A. Dacă $\mathbf{x}_0$ este un punct de extrem local al lui f şi f admite derivate parţiale (de ordinul întâi) în $\mathbf{x}_0$, atunci*

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

Observaţie. Dacă f este derivabilă Gâteaux în $\mathbf{x}_0$, concluzia teoremei poate fi scrisă $(\nabla f)(\mathbf{x}_0) = \mathbf{0}_{\mathbb{R}^n}$. Dacă f este în plus Fréchet diferenţiabilă în $\mathbf{x}_0$, acest lucru este de asemenea echivalent cu $\mathrm{d}f(\mathbf{x}_0) = \mathbf{0}_{\mathrm{L}(\mathbb{R}^n;\mathbb{R})}$.

DEFINIŢIE. Fie $A \subseteq \mathbb{R}^n$, $\mathbf{x}_0$ un punct interior lui A şi $f : A \to \mathbb{R}$ o funcţie, Fréchet diferenţiabilă în $\mathbf{x}_0$. Elementul $\mathbf{x}_0$ se numeşte *punct critic* (sau *punct staţionar*) al lui f dacă $\mathrm{d}f(\mathbf{x}_0) = \mathbf{0}_{\mathrm{L}(\mathbb{R}^n;\mathbb{R})}$ (*adică* $(\nabla f)(\mathbf{x}_0) = \mathbf{0}_{\mathbb{R}^n}$).

Observaţie. Teorema lui Fermat afirmă că dacă un punct interior mulţimii A, $\mathbf{x}_0$, este un punct de extrem local al unei funcţii $f : A \to \mathbb{R}$, Fréchet diferenţiabile în $\mathbf{x}_0$, atunci $\mathbf{x}_0$ este un punct critic al lui f. Reciproca nu este adevărată, după cum putem vedea din următorul 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 $\frac{\partial f}{\partial x}(0,0) = 2x\big|_{x=0,y=0} = 0$ şi $\frac{\partial f}{\partial y}(0,0) = -2y\big|_{x=0,y=0} = 0$. Pe de altă parte, $f(x,0) = x^2 > f(0,0) = 0 > -y^2 = f(0,y)$ pentru orice $x \neq 0$ şi orice $y \neq 0$. Acest lucru implică faptul că $(0,0)$ nu este nici punct de minim local, nici punct de maxim local pentru f.

DEFINIŢIE. Fie $A \subseteq \mathbb{R}^n$, $\mathbf{x}_0$ un punct interior lui A şi $f : A \to \mathbb{R}$ o funcţie Fréchet diferenţiabilă în $\mathbf{x}_0$. Dacă $\mathbf{x}_0$ este un punct critic al lui f, dar nu este punct de extrem local pentru f, spunem că $\mathbf{x}_0$ este un *punct şa* al lui f.

Pentru funcţii diferenţiabile de ordin cel puţin 2, există criterii (condiţii suficiente) pentru identificarea punctelor de extrem local printre punctele critice. Să începem cu cazul $n = 1$:

Teorema 2.2. *Fie $a, b \in \mathbb{R}$ cu $a < b$, $x_0 \in (a,b)$ şi $f : (a,b) \to \mathbb{R}$ o funcţie de k ori ($k \geq 2$) derivabilă în x_0. Să presupunem că $f'(x_0) = f''(x_0) = \ldots = f^{(k-1)}(x_0) = 0$ şi $f^{(k)}(x_0) \neq 0$.*

- ***i)*** *Atunci când k este număr par, x_0 este punct de extrem local al lui f. Mai precis, dacă $f^{(k)}(x_0) > 0$ atunci x_0 este punct de minim local al lui f, în timp ce dacă $f^{(k)}(x_0) < 0$ atunci x_0 este punct de maxim local al lui f.*
- ***ii)*** *Atunci când k este număr impar, x_0 este punct şa al lui f.*

Să trecem acum la cazul $n > 1$; vom caracteriza punctele de extrem ale unei funcţii f în $\mathbf{x}_0$ după comportamentul derivatei Fréchet de ordinul 2 în $\mathbf{x}_0$. Să ne amintim că

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

Teorema 2.3. *Fie $A \subseteq \mathbb{R}^n$ o mulţime nevidă şi $f : A \to \mathbb{R}$ o funcţie de clasă C^2 într-o vecinătate a unui punct critic $\mathbf{x}_0 \in A$.*

- ***i)*** *Dacă $\left(\mathrm{d}^2 f(\mathbf{x}_0)\right)(\mathbf{v}) > 0$, $\forall \mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$, atunci $\mathbf{x}_0$ este un punct de minim local al lui f.*
- ***ii)*** *Dacă $\left(\mathrm{d}^2 f(\mathbf{x}_0)\right)(\mathbf{v}) < 0$, $\forall \mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$, atunci $\mathbf{x}_0$ este un punct de maxim local al lui f.*
- ***iii)*** *Dacă $\mathrm{d}^2 f(\mathbf{x}_0)$ este o formă pătratică nedifinită (adică $\exists \mathbf{v}', \mathbf{v}'' \in \mathbb{R}^n$ astfel încât $\left(\mathrm{d}^2 f(\mathbf{x}_0)\right)(\mathbf{v}') < 0$ şi $\left(\mathrm{d}^2 f(\mathbf{x}_0)\right)(\mathbf{v}'') > 0$), atunci $\mathbf{x}_0$ este un punct şa pentru f.*

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

Vom numi matricea asociată acestei forme pătratice (relativ la baza canonică) *Hessianul* lui f. Datorită ipotezelor din rezultatul de mai sus, avem $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}_0)$, $\forall i \neq j$, din teorema lui Schwarz. Rezultă că Hessianul lui f are forma

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

care este o matrice simetrică.

Datorită cursului 7, dispunem de mai multe metode de a determina dacă $H_f(\mathbf{x}_0)$ este pozitiv definită, negativ definită, nedefinită sau niciuna dintre acestea.

Propoziţia 2.4. *Fie $A \subseteq \mathbb{R}^n$ o mulţime nevidă şi $f : A \to \mathbb{R}$ o funcţie de clasă C^2 pe o vecinătate a unui punct critic $\mathbf{x}_0 \in A$.*

- ***i)*** *Dacă toate valorile proprii ale lui $H_f(\mathbf{x}_0)$ sunt strict pozitive, atunci $\mathbf{x}_0$ este un punct de minim local al lui f.*
- ***ii)*** *Dacă toate valorile proprii ale lui $H_f(\mathbf{x}_0)$ sunt strict negative, atunci $\mathbf{x}_0$ este un punct de maxim local al lui f.*
- ***iii)*** *Dacă $H_f(\mathbf{x}_0)$ are cel puţin o valoare proprie strict pozitivă şi una strict negativă, atunci $\mathbf{x}_0$ este punct şa pentru f.*

Desigur, nu putem spune nimic în cazul în care toate valorile proprii sunt toate fie pozitive sau negative (dar nu strict).

Propoziţia 2.5. *Fie $A \subseteq \mathbb{R}^n$ o mulţime nevidă şi $f : A \to \mathbb{R}$ o funcţie de clasă C^2 pe o vecinătate a unui punct critic $\mathbf{x}_0 \in A$. Fie $\Delta_k = \det\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)\right)_{1 \leq i,j \leq k}$, $k = \overline{1,n}$, minorii principali ai Hessianului $H_f(\mathbf{x}_0)$.*

i) *Dacă* $\Delta_k > 0$, $\forall k = \overline{1,n}$, *atunci* $\mathbf{x}_0$ *este un punct de minim local pentru* f.
ii) *Dacă* $(-1)^{k+1}\Delta_k < 0$, $\forall k = \overline{1,n}$, *atunci* $\mathbf{x}_0$ *este un punct de maxim local pentru* f.
iii) *Dacă există* $j, k = \overline{1,n}$ *astfel încât* $(-1)^{j+1}\Delta_j < 0$ *şi* $(-1)^{k+1}\Delta_k > 0$, *atunci* $\mathbf{x}_0$ *este un punct şa pentru* f.

Din nou, nu putem spune nimic dacă inegalităţile $\Delta_k \geq 0$, $\forall k = \overline{1,n}$ sau $(-1)^{k+1}\Delta_k \leq 0$, $\forall k = \overline{1,n}$ nu sunt toate stricte.

Observaţie. În cazul particular $n = 2$, rezultatul de mai sus ne spune că dacă $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$ este C^2 pe o vecinătate a unui punct critic $\mathbf{x}_0 \in A$ şi notăm $p := \frac{\partial^2 f}{\partial x^2}(\mathbf{x}_0)$, $q := \frac{\partial^2 f}{\partial x \partial y}(\mathbf{x}_0)$, $r := \frac{\partial^2 f}{\partial y^2}(\mathbf{x}_0)$, atunci:

i) $\mathbf{x}_0$ este un punct de minim local pentru f, dacă $p > 0$ şi $pr - q^2 > 0$;
ii) $\mathbf{x}_0$ este un punct de maxim local pentru f, dacă $p < 0$ şi $pr - q^2 > 0$;
iii) $\mathbf{x}_0$ este punct şa pentru f, dacă $pr - q^2 < 0$;
iv) dacă $pr - q^2 = 0$, nu putem stabili natura lui $\mathbf{x}_0$ prin această metodă.

Revenind acum la metoda celor mai mici pătrate (Exemplul 1), cu φ liniară, observăm că funcţia $f : \mathbb{R}^2 \to \mathbb{R}$, definită prin

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

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

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

Pentru a rezolva această ecuaţie, echivalentă cu sistemul liniar

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

calculăm mai întâi determinantul sistemului,

$$\begin{vmatrix} \displaystyle\sum_{k=1}^{l} a_k^2 & \displaystyle\sum_{k=1}^{l} a_k \\ \displaystyle\sum_{k=1}^{l} a_k & l \end{vmatrix} = l \sum_{k=1}^{l} a_k^2 - \left(\sum_{k=1}^{l} a_k\right)^2.$$

Acesta este diferit de 0 dacă nu toţi $a_1, a_2, \ldots, a_l$ sunt egali, din inegalitatea Cauchy-Buniakowski-Schwarz. În acest caz, avem

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

şi

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

Avem atunci $p = \frac{\partial^2 f}{\partial c_1^2}\left(c_1^0, c_2^0\right) = 2\sum_{k=1}^{l} a_k^2$, $q = \frac{\partial^2 f}{\partial c_1 \partial c_2}\left(c_1^0, c_2^0\right) = 2\sum_{k=1}^{l} a_k$ şi $r = \frac{\partial^2 f}{\partial c_2^2}\left(c_1^0, c_2^0\right) = 2l$. Din nou datorită inegalităţii Cauchy-Buniakowski-Schwarz, avem $p > 0$ şi $pr - q^2 > 0$, 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).

3. Extreme cu legături

După cum am remarcat deja, în afara situaţiilor în care trebuie să rezolvăm probleme de extrem fără constrângeri, în practică apar de asemenea probleme în care trebuie să calculăm extreme ale unor funcţii ale căror argumente sunt supuse unor anumite condiţii (restricţii). În acest caz, respectivele extreme sunt numite *extreme cu legături.*

Să precizăm acum contextul în care ne vom plasa. Fie D o mulţime deschisă în $\mathbb{R}^n \times \mathbb{R}^m$ ($n, m \in \mathbb{N}^*$) şi $f : D \to \mathbb{R}$, $g : D \to \mathbb{R}^m$ funcţii de clasă C^1 pe D. Vom căuta puncte de extrem ale funcţiei f restricţionată la condiţia suplimentară $g(\mathbf{x}, \mathbf{y}) = 0$, unde $(\mathbf{x}, \mathbf{y}) \in D$.

De fapt, notând $A := \{(\mathbf{x}, \mathbf{y}) \in D | g(\mathbf{x}, \mathbf{y}) = \mathbf{0}_{\mathbb{R}^n}\}$ şi prin $g_1, g_2, \ldots, g_m$ componentele lui g, observăm că problema de mai sus cere să găsim punctele de extrem ale lui $f|_A$. Pentru a fixa terminologia, spunem că punctul $(\mathbf{x}_0, \mathbf{y}_0) \in A$ este un punct de *extrem (minim, maxim) local* sau *punct şa* al lui f, *cu legăturile* $g_1(\mathbf{x}, \mathbf{y}) = 0$, $g_2(\mathbf{x}, \mathbf{y}) = 0, \ldots, g_m(\mathbf{x}, \mathbf{y}) = 0$, dacă este punct de extrem (minim, maxim) local, respectiv punct şa al lui $f|_A$.

Problema aceasta ar putea părea ca o problemă de extrem necondiţionat, dar punctele de extrem $(\mathbf{x}_0, \mathbf{y}_0)$ ale lui $f|_A$ vor fi puncte interioare lui A dacă şi numai dacă $g_1 = g_2 = \cdots = g_m = 0$ pe o vecinătate a lui $(\mathbf{x}_0, \mathbf{y}_0)$, ceea ce ar implica faptul că aceste condiţii nu sunt în realitate restrictive. De aceea în general nu putem aplica rezultatele din secţiunile precedente.

Pentru a aborda o astfel de problemă, am putea mai întâi să încercăm $g_k(\mathbf{x}, \mathbf{y}) = 0$, $\forall\, k = \overline{1, m}$, în raport cu $\mathbf{y}$. A găsi $\mathbf{y} = \varphi(\mathbf{x})$, $\mathbf{x} \in B$, va duce la studiul problemei de extrem necondiţionate pentru funcţia $\mathbf{x} \mapsto f(\mathbf{x}, \varphi(\mathbf{x}))$, cu $\mathbf{x} \in B$. O astfel de procedură este, totuşi, dificil de aplicat în practică, deoarece ecuaţia $g(\mathbf{x}, \mathbf{y}) = \mathbf{0}_{\mathbb{R}^n}$ nu este uşor de rezolvat.

De aceea vom utiliza metoda multiplicatorilor lui *Lagrange*, descrisă în continuare.

Teorema 3.1 (existenţa multiplicatorilor lui Lagrange). *Fie $D \subseteq \mathbb{R}^n \times \mathbb{R}^m$ o mulţime deschisă nevidă, $f : D \to \mathbb{R}$ şi $g : D \to \mathbb{R}^m$ funcţii de clasă C^1 pe D. Fie $g_1, g_2, \ldots, g_m$ componentele lui g şi $(\mathbf{x}_0, \mathbf{y}_0) \in D$ un punct de extrem local al lui f, cu legăturile $g_1(\mathbf{x}, \mathbf{y}) = 0$, $g_2(\mathbf{x}, \mathbf{y}) = 0, \ldots, g_m(\mathbf{x}, \mathbf{y}) = 0$. Dacă $\frac{D(g_1, g_2, \ldots, g_m)}{D(y_1, y_2, \ldots, y_m)}(\mathbf{x}_0, \mathbf{y}_0) \neq 0$, atunci există $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{R}$ astfel încât $(\mathbf{x}_0, \mathbf{y}_0)$ este un punct critic pentru funcţia de clasă C^1, $L : D \to \mathbb{R}$ definită prin*

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

adică

$$(\star)\ \begin{cases} \dfrac{\partial L}{\partial x_k}(\mathbf{x}_0, \mathbf{y}_0) = 0,\ \forall k = \overline{1, n}; \\ \dfrac{\partial L}{\partial y_j}(\mathbf{x}_0, \mathbf{y}_0) = 0,\ \forall j = \overline{1, m}; \\ g_j(\mathbf{x}_0, \mathbf{y}_0) = 0,\ \forall j = \overline{1, m}. \end{cases}$$

Observaţii.

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 condiţionat ale lui f sunt printre punctele critice ale funcţiei asociate $L = f + \sum_{k=1}^{m} \lambda_k g_k$.

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

$$\mathscr{L}(\boldsymbol{\lambda}, \mathbf{x}, \mathbf{y}) = f(\mathbf{x}, \mathbf{y}) + \lambda_1 g_1(\mathbf{x}, \mathbf{y}) + \lambda_2 g_2(\mathbf{x}, \mathbf{y}) + \cdots + \lambda_m g_m(\mathbf{x}, \mathbf{y}),\ \boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m,\ (\mathbf{x}, \mathbf{y}) \in D,$$

se numeşte *Lagrangianul* asociat lui f şi restricţiei g (am adăugat o nouă variabilă $\boldsymbol{\lambda} \in \mathbb{R}^m$ funcţiei L pentru a obţine funcţia $\mathscr{L}$).

Tripletul $(\boldsymbol{\lambda}_0, \mathbf{x}_0, \mathbf{y}_0)$ satisface atunci sistemul $(\star)$ (adică $(\mathbf{x}_0, \mathbf{y}_0)$ este un punct critic pentru L, ce satisface $g(\mathbf{x}_0, \mathbf{y}_0) = \mathbf{0}_{\mathbb{R}^m}$) dacă şi numai dacă $(\boldsymbol{\lambda}_0, \mathbf{x}_0, \mathbf{y}_0)$ este un punct critic pentru $\mathscr{L}$, de vreme ce

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

4. Să presupunem acum că $(\boldsymbol{\lambda}_0, \mathbf{x}_0, \mathbf{y}_0)$ satisface sistemul $(\star)$ şi f, g sunt funcţii de clasă C^2 pe D (cel puţin pe o vecinătate a lui $(\boldsymbol{\lambda}_0, \mathbf{x}_0, \mathbf{y}_0)$). Deoarece

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

vedem că $(\mathbf{x}_0, \mathbf{y}_0)$ este un punct de extrem local condiţionat al lui f dacă şi numai dacă $(\mathbf{x}_0, \mathbf{y}_0)$ este un punct de extrem local al funcţiei $L|_A$. Deoarece $\frac{D(g_1, g_2, \ldots, g_m)}{D(y_1, y_2, \ldots, y_m)}(\mathbf{x}_0, \mathbf{y}_0) \neq 0$, se poate rezolva ecuaţia $g(\mathbf{x}, \mathbf{y}) = 0$ şi găsi o soluţie $\mathbf{y} = \varphi(\mathbf{x})$ (cel puţin în jurul punctului $(\mathbf{x}_0, \mathbf{y}_0)$ – aceasta este teorema funcţiilor implicite, ce depăşeşte cadrul acestui curs). De aceea, putem determina condiţii suficiente de extrem condiţionat studiind Hessianianul funcţiei $\mathbf{x} \mapsto L(\mathbf{x}, \varphi(\mathbf{x}))$ (notată tot L) în $\mathbf{x}_0$ sau, echivalent, a formei pătratice asociate diferenţialei Fréchet de ordinul 2 a lui L în $\mathbf{x}_0$. O modalitate practică de a realiza acest lucru este următoarea:

Să notăm $\mathbf{x} = (\tilde{x}_1, \ldots, \tilde{x}_n) \in \mathbb{R}^n$ şi $\mathbf{y} = (\tilde{x}_{n+1}, \ldots, \tilde{x}_{n+m}) \in \mathbb{R}^m$; diferenţiala Fréchet de ordinul 2 a lui L în $(\mathbf{x}_0, \mathbf{y}_0)$ este

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

Considerând sistemul obţinut din diferenţierea relaţiilor $g_k(\mathbf{x}, \mathbf{y}) = 0, \ k = \overline{1, m}$, adică

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

putem scrie pe $\mathrm{d}\tilde{x}_{n+1}, \ldots, \mathrm{d}\tilde{x}_{n+m}$ ca funcţie de $\mathrm{d}\tilde{x}_1, \ldots, \mathrm{d}\tilde{x}_n$ şi înlocuite în $(\star\star)$. Acest lucru permite scrierea diferenţialei de ordin 2 a lui L sub forma

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

Este acum suficient să aplicăm teorema 2.3 pentru a decide dacă $\mathbf{x}_0$ este un punct de extrem local sau un punct şa pentru L: dacă $\mathrm{d}^2 L(\mathbf{x}_0)$ este o formă pozitiv sau negativ definită, atunci $(\mathbf{x}_0, \mathbf{y}_0)$ este un punct de minim local condiţionat, respectiv punct de maxim local condiţionat al lui f; dacă $\mathrm{d}^2 L(\mathbf{x}_0)$ este nedefinită, atunci $(\mathbf{x}_0, \mathbf{y}_0)$ este un punct şa condiţionat pentru f; dacă $\mathrm{d}^2 L(\mathbf{x}_0)$ este positiv sau negativ semi-definită, nu putem decide asupra naturii lui $(\mathbf{x}_0, \mathbf{y}_0)$.

Paşii descrişi mai sus constituie metoda multiplicatorilor lui Lagrange privitoare la problema extremelor cu legături.

Să aplicăm această metodă pentru entropia lui Shannon, adică problema găsirii extremelor funcţiei

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

condiţionate de legătura $\sum_{k=1}^{n} p_k = 1$. Deoarece avem doar o restricţie, avem $m = 1$, $g(p_1, p_2, \ldots, p_n) := \sum_{k=1}^{n} p_k - 1$ şi Lagrangianul este

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

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

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

Din primele n ecuaţii găsim $p_1^0 = p_2^0 = \ldots = p_n^0 = \psi(\lambda_1^0)$, unde $\psi(\lambda)$ este unica soluţie a ecuaţiei în $p \in (0,1)$:

$$\log_2 p + \frac{1}{\ln 2} = \lambda,$$

în timp ce din ultima găsim

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

Datorită restricţiei $g(p_1, p_2, \ldots, p_n) = 0$ găsim $\mathrm{d}p_1 + \mathrm{d}p_2 + \cdots + \mathrm{d}p_n = 0$; de asemenea avem

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

care este o formă pătratică negativ definită. Aceasta înseamnă 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 pentru $\tilde{H}$, condiţionat de $p_1 + p_2 + \cdots + p_n = 1$.

În consecinţă, entropia informaţională maximă este obţinută pentru variabila aleatoare dată de

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

Bibliografie selectivă

[1] C. Canuto, A. Tabacco, *Mathematical Analysis II,* Springer, 2015.
[2] B. Edwards, R. Larson, *Calculus (9th Edition)*, Brooks/Cole-Cengage Learning, 2014.
[3] K. P. Evans, N. Jacob, *A Course in Analysis*, World Scientific, 2016.
[4] M. Gorunescu, *Lecţii de analiză matematică pentru informaticieni*, Reprografia Universităţii Craiova, 2000.
[5] F. Iacob, *Matematică pentru anul II - ID*, seria 2004-2005.
[6] R. Luca-Tudorache, *Analiză matematică*, Editura Tehnopress, Iaşi, 2005.
[7] G. Păltineanu, *Analiză matematică. Calcul diferenţial*, Editura AGIR, Bucureşti, 2002.
[8] E. Popescu, *Analiză matematică. Calcul diferenţial*, Editura Matrix Rom, Bucureşti, 2006.
[9] M. Postolache, *Analiză matematică (teorie şi aplicaţii)*, Ed. Fair Partners, Bucureşti, 2011.
[10] V. Postolică, *Eficienţă prin matematică aplicată. Analiză matematică*, Editura Matrix Rom, Bucureşti, 2006.
[11] A. Precupanu, *Bazele analizei matematice*, Editura Polirom, Iaşi, 1998.