

CURSUL 10
APLICAȚII ALE DIFERENȚIABILITĂȚII. EXTREME ȘI EXTREME CU LEGĂTURI

Una dintre multiplele aplicații ale diferențiabilității funcțiilor este stabilirea punctelor de extrem pentru funcții ce apar în probleme de optimizare. Mai exact, aceste probleme urmăresc minimizarea (sau maximizarea) unor așa-zise *funcționale de cost (funcții de profit)*, în prezența sau absența unor anumite restricții.

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, \dots, b_p , corespunzând valorilor de “input” a_1, a_2, \dots, 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 referință în plan, putem estima (vizual) natura unei funcții necunoscute φ care satisface, cel puțin din punct de vedere teoretic, pentru orice $k = \overline{1, p}$, $\varphi(a_k) = b_k$. Mai exact, presupunem că funcția φ are un anumit tip (polinomial, exponențial, trigonometric, etc.), ai cărei parametri (caracteristici) c_1, c_2, \dots, c_n trebuie să fie identificați. În acest scop, utilizăm *metoda celor mai mici pătrate*, prin considerarea problemei minimizării expresiei

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

în raport cu $(c_1, c_2, \dots, c_n) \in \mathbb{R}^n$. Rezolvând această problemă (de extrem), adică găsim $(c_1^0, c_2^0, \dots, c_n^0) \in \mathbb{R}^n$ astfel încât

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

putem concluda că mărimea fizică în studiu urmează legea $y = \varphi(x; c_1^0, c_2^0, \dots, c_n^0)$.

Remarcăm că dacă reprezentarea grafică a mulțimii $\{(a_k, b_k) | k = \overline{1, p}\}$ ne sugerează că φ este liniară, atunci putem lua $n = 2$ și $\varphi(x) := c_1 x + c_2$. În acest caz, metoda celor mai mici pătrate implică minimizarea 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, spațiul \mathbb{R}^n este interpretat ca spațiul *bunurilor de consum*, unde fiecare *bun de consum* este identificat de un index $i \in \{1, 2, \dots, n\}$, în timp ce *vectorul bunurilor de consum* este un vector $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, 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 $\mathbf{p} = (p_1, p_2, \dots, p_n) \in \mathbb{R}^n$, unde p_i este prețul “unitar” al bunului de consum i ; sistemul de prețuri este dat așadar de $\sum_{i=1}^n p_i x_i$.

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. În acest scop, economiștii folosesc o *funcție de cost* și/sau o *funcție de profit* alese convenabil, și numite *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 acest spațiu este în schimb $K \subsetneq \mathbb{R}^n$ atunci avem de a face cu 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, de tipul

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

și are expresia

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

unde p_k este probabilitatea ($p_k \in (0, 1)$, $k = \overline{1, n}$ și $\sum_{k=1}^n p_k = 1$) 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. Cu alte cuvinte, vom studia problema maximizării funcției

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

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

2. EXTREME NECONDIȚIONATE

Reamintim câteva definiții legate de spații metrice, de care vom avea nevoie în continuare. Fie (X, d) un spațiu metric.

1. Bila deschisă centrată în $x_0 \in X$, de rază $r > 0$, este mulțimea

$$B(x_0; r) := \{x \in X \mid d(x, x_0) < r\}.$$

2. O vecinătate a unui punct $x_0 \in X$ este o submulțime $V \subseteq X$ pentru care există $\varepsilon > 0$ astfel încât $B(x_0; \varepsilon) \subseteq V$.
3. Un punct $x_0 \in X$ se numește *punct interior* al unei submulțimi $A \subseteq X$ dacă A este vecinătate a lui x_0 .
4. O submulțime $A \subseteq X$ se numește *deschisă* dacă orice element al său îi este punct interior.

DEFINIȚIE. Fie $A \subseteq \mathbb{R}^n$ o mulțime nevidă, $f : A \rightarrow \mathbb{R}$ o funcție și $x_0 \in A$. Spunem că x_0 este:

- a) un *punct de minim* al lui f dacă $f(x) \geq f(x_0)$, $\forall x \in A$;
- b) un *punct de maxim* al lui f dacă $f(x) \leq f(x_0)$, $\forall x \in A$;
- c) un *punct de minim strict* al lui f dacă $f(x) > f(x_0)$, $\forall x \in A \setminus \{x_0\}$;
- d) un *punct de maxim strict* al lui f dacă $f(x) < f(x_0)$, $\forall x \in A \setminus \{x_0\}$;
- e) un *punct de minim local* al lui f dacă există o vecinătate V a lui x_0 astfel încât $f(x) \geq f(x_0)$, $\forall x \in A \cap V$;
- f) un *punct de maxim local* al lui f dacă există o vecinătate V a lui x_0 astfel încât $f(x) \leq f(x_0)$, $\forall x \in A \cap V$;
- g) un *punct de minim local strict* al lui f dacă există o vecinătate V a lui x_0 astfel încât $f(x) > f(x_0)$, $\forall x \in A \cap (V \setminus \{x_0\})$;
- h) un *punct de maxim local strict* al lui f dacă există o vecinătate V a lui x_0 astfel încât $f(x) < f(x_0)$, $\forall x \in A \cap (V \setminus \{x_0\})$;
- i) 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 .

Observație. 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 $f : D \rightarrow \mathbb{R}$ 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, $f : D \rightarrow \mathbb{R}$ 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 $f|_{A \cap D}$ se numește o *problemă de extrem condiționat*.

Teorema 2.1 (Fermat). Fie $A \subseteq \mathbb{R}^n$ o mulțime oarecare, $f : A \rightarrow \mathbb{R}$ 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

$$\frac{\partial f}{\partial x_1}(x_0) = \dots = \frac{\partial f}{\partial x_n}(x_0) = 0.$$

Observație. 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 diferențiabilă în \mathbf{x}_0 , acest lucru este de asemenea echivalent cu $df(\mathbf{x}_0) = \mathbf{0}_{L(\mathbb{R}^n; \mathbb{R})}$.

DEFINIȚIE. Fie $A \subseteq \mathbb{R}^n$, \mathbf{x}_0 un punct interior lui A și $f : A \rightarrow \mathbb{R}$ o funcție, Fréchet diferențiabilă în \mathbf{x}_0 . Elementul \mathbf{x}_0 se numește *punct critic* (sau *punct staționar*) al lui f dacă $df(\mathbf{x}_0) = \mathbf{0}_{L(\mathbb{R}^n; \mathbb{R})}$ (adică $(\nabla f)(\mathbf{x}_0) = \mathbf{0}_{\mathbb{R}^n}$).

Observație. Teorema lui Fermat afirmă că dacă un punct interior mulțimii A , \mathbf{x}_0 , este un punct de extrem local al unei funcții $f : A \rightarrow \mathbb{R}$, Fréchet diferențiabile 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 $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = x^2 - y^2$, $x, y \in \mathbb{R}$. Atunci $(0, 0)$ este un punct critic al lui f , deoarece $\frac{\partial f}{\partial x}(0, 0) = 2x|_{x=0, y=0} = 0$ și $\frac{\partial f}{\partial y}(0, 0) = -2y|_{x=0, y=0} = 0$. Pe de altă parte, $f(x, 0) = x^2 > f(0, 0) = 0 > -y^2 = f(0, y)$ pentru orice $x \neq 0$ și 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 .

DEFINIȚIE. Fie $A \subseteq \mathbb{R}^n$, \mathbf{x}_0 un punct interior lui A și $f : A \rightarrow \mathbb{R}$ o funcție Fréchet diferențiabilă 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 șa* al lui f .

Pentru funcții diferențiabile de ordin cel puțin 2, există criterii (condiții 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 $a < b$, $x_0 \in (a, b)$ și $f : (a, b) \rightarrow \mathbb{R}$ o funcție de k ori ($k \geq 2$) derivabilă în x_0 . Să presupunem că $f'(x_0) = f''(x_0) = \dots = f^{(k-1)}(x_0) = 0$ și $f^{(k)}(x_0) \neq 0$.

- 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 șa al lui f .

Să trecem acum la cazul $n > 1$; vom stabili dacă un punct critic \mathbf{x}_0 al unei funcții f este punct de extrem local după comportamentul derivatei Fréchet de ordinul 2 în \mathbf{x}_0 . Să ne amintim că

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

Teorema 2.3. Fie $A \subseteq \mathbb{R}^n$ o mulțime nevidă și $f : A \rightarrow \mathbb{R}$ o funcție de clasă C^2 într-o vecinătate a unui punct critic $\mathbf{x}_0 \in A$.

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

Observație. Dacă $d^2 f(\mathbf{x}_0)$ este o formă pătratică pozitiv sau negativ-semidefinită (adică $(d^2 f(\mathbf{x}_0))(\mathbf{v}) \geq 0$, $\forall \mathbf{v} \in \mathbb{R}^n$ sau $(d^2 f(\mathbf{x}_0))(\mathbf{v}) \leq 0$, $\forall \mathbf{v} \in \mathbb{R}^n$, 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 $d^2 f(\mathbf{x}_0)$ este pozitiv definită, negativ definită, nedefinită sau niciuna dintre acestea.

Propoziția 2.4. Fie $A \subseteq \mathbb{R}^n$ o mulțime nevidă și $f : A \rightarrow \mathbb{R}$ o funcție 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 puțin o valoare proprie strict pozitivă și una strict negativă, atunci \mathbf{x}_0 este punct șa pentru f .

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

Propoziția 2.5. Fie $A \subseteq \mathbb{R}^n$ o mulțime nevidă și $f : A \rightarrow \mathbb{R}$ o funcție 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 $(-1)^{j+1} \Delta_j < 0$ și $(-1)^{k+1} \Delta_k > 0$, atunci \mathbf{x}_0 este un punct șa pentru f .

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

Observație. În cazul particular $n = 2$, rezultatul de mai sus ne spune că dacă $f : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ este C^2 pe o vecinătate a unui punct critic $\mathbf{x}_0 \in A$ și 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 $p > 0$ și $pr - q^2 > 0$;
ii) \mathbf{x}_0 este un punct de maxim local pentru f , dacă $p < 0$ și $pr - q^2 > 0$;
iii) \mathbf{x}_0 este punct șa 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 funcția $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, 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 și are ca puncte critice elementele $(c_1^0, c_2^0) \in \mathbb{R}^2$ ce satisfac

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

Pentru a rezolva această ecuație, echivalentă cu sistemul liniar

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

calculăm mai întâi determinantul sistemului,

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

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

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

și

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

Avem atunci $p = \frac{\partial^2 f}{\partial c_1^2}(c_1^0, c_2^0) = 2 \sum_{k=1}^l a_k^2$, $q = \frac{\partial^2 f}{\partial c_1 \partial c_2}(c_1^0, c_2^0) = 2 \sum_{k=1}^l a_k$ și $r = \frac{\partial^2 f}{\partial c_2^2}(c_1^0, c_2^0) = 2l$. Din nou datorită inegalității Cauchy-Buniakowski-Schwarz, avem $p > 0$ și $pr - q^2 > 0$, deci (c_1^0, c_2^0) este un punct de minim global pentru f (nu doar local, deoarece (c_1^0, c_2^0) este unicul punct critic al lui f).

3. EXTREME CU LEGĂTURI

După cum am remarcat deja, în afara situațiilor în care trebuie să rezolvăm probleme de extrem fără constrângeri, în practică apar de asemenea probleme în care trebuie să calculăm extreme ale unor funcții ale cărori argumente sunt supuse unor anumite condiții (restricții). În acest caz, respectivele extreme sunt numite *extreme cu legături*.

Să precizăm acum contextul în care ne vom plasa. Fie D o mulțime deschisă în $\mathbb{R}^n \times \mathbb{R}^m$ ($n, m \in \mathbb{N}^*$) și $f : D \rightarrow \mathbb{R}$, $g : D \rightarrow \mathbb{R}^m$ funcții de clasă C^1 pe D . Vom căuta puncte de extrem ale funcției f restricționată la condiția suplimentară $g(\mathbf{x}, \mathbf{y}) = 0$, unde $(\mathbf{x}, \mathbf{y}) \in D$.

De fapt, notând $A := \{(\mathbf{x}, \mathbf{y}) \in D \mid g(\mathbf{x}, \mathbf{y}) = \mathbf{0}_{\mathbb{R}^m}\}$ și prin g_1, g_2, \dots, 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 șa* al lui f , cu legăturile $g_1(\mathbf{x}, \mathbf{y}) = 0, g_2(\mathbf{x}, \mathbf{y}) = 0, \dots, g_m(\mathbf{x}, \mathbf{y}) = 0$, dacă este punct de extrem (minim, maxim) local, respectiv punct șa al lui $f|_A$.

Problema aceasta ar putea părea ca o problemă de extrem necondiționat, dar punctele de extrem $(\mathbf{x}_0, \mathbf{y}_0)$ ale lui $f|_A$ vor fi puncte interioare lui A dacă și numai dacă $g_1 = g_2 = \dots = g_m = 0$ pe o vecinătate a lui $(\mathbf{x}_0, \mathbf{y}_0)$, ceea ce ar implica faptul că aceste condiții nu sunt în realitate restrictive. De aceea în general nu putem aplica rezultatele din secțiunile 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 necondiționate pentru funcția $\mathbf{x} \mapsto f(\mathbf{x}, \varphi(\mathbf{x}))$, cu $\mathbf{x} \in B$. O astfel de procedură este, totuși, dificil de aplicat în practică, deoarece ecuația $g(\mathbf{x}, \mathbf{y}) = \mathbf{0}_{\mathbb{R}^m}$ nu este ușor de rezolvat.

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

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

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

adică

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

Observații.

1. Numerele $\lambda_1, \lambda_2, \dots, \lambda_m$ 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 $\mathcal{L} : \mathbb{R}^m \times D \rightarrow \mathbb{R}$ definită de

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

se numește *Lagrangianul* asociat lui f și restricției g (am adăugat o nouă variabilă $\boldsymbol{\lambda} \in \mathbb{R}^m$ funcției L pentru a obține funcția \mathcal{L}).

Tripletul $(\boldsymbol{\lambda}_0, \mathbf{x}_0, \mathbf{y}_0)$ satisface atunci sistemul $(*)$ (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 și numai dacă $(\boldsymbol{\lambda}_0, \mathbf{x}_0, \mathbf{y}_0)$ este un punct critic pentru \mathcal{L} , de vreme ce

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

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

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

vedem că $(\mathbf{x}_0, \mathbf{y}_0)$ este un punct de extrem local condiționat al lui f dacă și numai dacă $(\mathbf{x}_0, \mathbf{y}_0)$ este un punct de extrem local al funcției $L|_A$. Deoarece $\frac{D(g_1, g_2, \dots, g_m)}{D(y_1, y_2, \dots, y_m)}(\mathbf{x}_0, \mathbf{y}_0) \neq 0$, se poate rezolva ecuația $g(\mathbf{x}, \mathbf{y}) = 0$ și găsi o soluție $\mathbf{y} = \varphi(\mathbf{x})$ (cel puțin în jurul punctului $(\mathbf{x}_0, \mathbf{y}_0)$ – aceasta este teorema funcțiilor implicite, ce depășește cadrul acestui curs). De aceea, putem determina condiții suficiente de extrem condiționat studiind Hessianul funcției $\mathbf{x} \mapsto L(\mathbf{x}, \varphi(\mathbf{x}))$ (notată tot L) în \mathbf{x}_0 sau, echivalent, a formei pătratic asociate diferențialei Fréchet de ordinul 2 a lui L în \mathbf{x}_0 . O modalitate practică de a realiza acest lucru este următoarea:

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

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

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

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

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

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

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

Observație: Dacă forma pătratică din $(\star\star)$ este pozitiv sau negativ definită, este clar că și forma $(\star\star\star)$ va fi pozitiv, respectiv negativ definită și astfel nu trebuie să mai efectuăm acest ultim pas pentru a decide că $(\mathbf{x}_0, \mathbf{y}_0)$ este un punct de minim local condiționat, respectiv punct de maxim local condiționat al lui f . În schimb, lucrurile se pot schimba dacă $(\star\star)$ este nedefinită sau doar pozitiv sau negativ semi-definită, deci va trebui să studiem forma $(\star\star\star)$.

Pașii descriși 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 funcției

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

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

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

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

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

Din primele n ecuații găsim $p_1^0 = p_2^0 = \dots = p_n^0 = \psi(\lambda_1^0)$, 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

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

Datorită restricției $g(p_1, p_2, \dots, p_n) = 0$ găsim $dp_1 + dp_2 + \dots + dp_n = 0$; de asemenea avem

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

care este o formă pătratică negativ definită. Aceasta înseamnă că punctul $(p_1^0, p_2^0, \dots, p_n^0) = (\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$ este punct de maxim pentru \tilde{H} , condiționat de $p_1 + p_2 + \dots + p_n = 1$.

În consecință, entropia informațională maximă este obținută pentru variabila aleatoare dată de

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

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, *Lecții de analiză matematică pentru informaticieni*, Reprografia Universității Craiova, 2000.
- [5] F. Iacob, *Matematică pentru anul II - ID*, seria 2004-2005.
- [6] R. Luca-Tudorache, *Analiză matematică*, Editura Tehnopress, Iași, 2005.
- [7] G. Păltineanu, *Analiză matematică. Calcul diferențial*, Editura AGIR, București, 2002.
- [8] E. Popescu, *Analiză matematică. Calcul diferențial*, Editura Matrix Rom, București, 2006.
- [9] M. Postolache, *Analiză matematică (teorie și aplicații)*, Ed. Fair Partners, București, 2011.
- [10] V. Postolică, *Eficiență prin matematică aplicată. Analiză matematică*, Editura Matrix Rom, București, 2006.
- [11] A. Precupanu, *Bazele analizei matematice*, Editura Polirom, Iași, 1998.