Cursul 11

Aplicaţii ale diferenţiabilităţii funcţiilor.

(Extreme fără restricţii. Inversabilitatea funcţiilor de mai multe variabile reale. Dependenţă/independenţă funcţională. Extreme condiţionate.)

Una dintre principalele aplicaţii practice a diferenţiabilităţii funcţiilor reale (de una sau mai multe variabile şi cu valori scalare) este aceea relativă la stabilirea elementelor de extrem pentru funcţiile implicate în anumite probleme de optimizare, adică în probleme care vizează minimizarea sau maximizarea unei aşa-numite funcţionale de cost, în absenţa sau în prezenţa unor condiţii (restricţii) precizate.

Exemple de probleme de optimizare în $\mathbb{R}^n$

În cele ce urmează, vom prezenta câteva exemple de probleme de optimizare.

1. Metoda celor mai mici pătrate pentru minimizarea abaterii unor estimări faţă de determinări

Presupunem că, în urma unor experimente asupra unei anumite mărimi fizice, s-au obţinut valorile $b_1, b_2, ..., b_p$, corespunzătoare valorilor (de "intrare") $a_1, a_2, \ldots, a_p$, unde $p \in \mathbb{N}^*$. Reprezentând punctele (a_k, b_k), $k \in \overline{1,p}$, într-un reper ortogonal din plan, facem o apreciere asupra naturii expresiei funcţiei φ care, necunoscută iniţial, ar avea, în a_k, valoarea b_k, $\forall\, k \in \overline{1,p}$. Potrivit acestei aprecieri, estimăm că φ ar avea o expresie de un anumit tip polinomial, exponenţial, trigonometric, etc. Se cere să determinăm parametrii $c_1, c_2, \ldots, c_n \in \mathbb{R}$, caracteristici funcţiei φ, definită mai sus.

În acest scop, folosind aşa-numita ***metodă a celor mai mici pătrate***, considerăm problema minimizării expresiei

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

în raport cu $(c_1, c_2, \ldots, c_n) \in \mathbb{R}^n$. Prin rezolvarea acestei probleme, de extrem fără restricţii, adică prin găsirea soluţiei $(c_1^0, c_2^0, \ldots, c_n^0) \in \mathbb{R}^n$ (când aceasta există şi este unică), pentru care

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

putem aprecia (în final) faptul că mărimea fizică asupra căreia s-au făcut măsurătorile ce au condus la determinările $b_1, b_2, \ldots, b_p$, se supune legii $y = \varphi(x; c_1^0, c_2^0, \ldots, c_n^0)$.

Dacă imaginea grafică a mulţimii $\{(a_k, b_k) \mid k \in \overline{1,p}\}$ ne sugerează faptul că φ ar avea o expresie liniară, atunci putem lua $n = 2$ şi $\varphi(x) = c_1 x + c_2$. Prin aceasta, metoda celor mai mici pătrate va consta în determinarea parametrilor c_1 şi c_2, aşa încât expresia

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

să fie minimă.

2. Realizarea unui profit maxim sau a unui cost minim într-o producţie economică

În economie, spaţiul $\mathbb{R}^n$ este interpretat ca fiind spaţiul complexelor de bunuri de consum, în care fiecare *bun* (produs) este caracterizat de un anumit indice $i \in \overline{1,n}$, iar un *complex de bunuri* este un vector $\mathrm{x} \in \mathbb{R}^n$, $\mathrm{x} = (x_1, x_2, \ldots, x_n)$, unde x_i reprezintă cantitatea în care se găseşte *bunul* i. *Elementul unitate* al bunului i

este $\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$. Dreapta reală este interpretată ca mulţimea valorilor. Într-un astfel de context, un *sistem de preţuri* este o funcţie ce asociază fiecărui complex de bunuri o anumită valoare. Se consideră astfel că un sistem de preţuri este un element din dualul lui $\mathbb{R}^n$, adică o aplicaţie liniară $\mathrm{p} = (p_1, p_2, \ldots, p_n)$, unde p_i este preţul unitar al *bunului i*. Astfel, pentru complexul de bunuri $\mathrm{x} \in \mathbb{R}^n$, valoarea sa în raport cu sistemul de preţuri p este dată de

$$\langle \mathrm{p}, \mathrm{x} \rangle_e = \sum_{i=1}^{n} p_i x_i.$$

Considerând mulţimea $\mathbb{R}^n_+$, a vectorilor $\mathrm{x} \in \mathbb{R}^n$ cu componente nenegative, comportamentul unui consumator este apreciat printr-o funcţie $u : \mathbb{R}^n_+ \to \mathbb{R}$, ce determină o aşa-numită *relaţie de preferinţă*, ”$\prec$”, pe mulţimea $\mathbb{R}^n_+$, definită prin: $\mathrm{y} \prec \mathrm{x} \Leftrightarrow u(\mathrm{x}) \le u(\mathrm{y})$.

Dacă o firmă produce un anumit *complex de bunuri*, se poate pune problema realizării respectivei producţii astfel încât cheltuielile de producţie sa fie minime sau/şi profitul de producţie să fie maxim. Prin studii economice adecvate, se utilizează, în acest sens, o funcţie de cost corespunzătoare contextului şi o funcţie de profit convenabil stabilită.

Când optimul funcţiei obiectiv (de cost sau /şi profit) se cere a fi găsit, în situaţia în care mulţimea complexelor de bunuri este $\mathbb{R}^n$, spunem că avem de-a face cu o problemă de extrem fără restricţii. În caz contrar, când mulţimea producţiilor nete ale unei firme, notată cu K, este o submulţime proprie a lui $\mathbb{R}^n_+$ ce conţine $\mathbf{0}_{\mathbb{R}^n}$, problema optimizării funcţiei u, pe K, este una de extrem condiţionat. Problema în cauză este de ”programare liniară” ori de câte ori atât funcţia de optimizat, cât şi relaţiile ce definesc mulţimea K sunt liniare. Dacă funcţia u este pătratică sau convexă, iar K este convexă, atunci problema de optimizare considerată se numeşte problemă de optimizare pătratică şi respectiv convexă, cu restricţii.

3. Problema entropiei informaţionale maxime

Introdusă, ca noţiune matematică, de Claude E. Shannon (1947), *entropia* reprezintă o funcţie ce corespunde cantităţii de informaţie livrată de o anumită sursă, prin intermediul unui anumit limbaj, semnal electric sau fişier (informatic) de date. Funcţia respectivă, notată cu H, este definită pe mulţimea variabilelor aleatoare

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

şi are expresia

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

unde $p_k \in (0, 1)$ este probabilitatea cu care se transmit k informaţii primite de la o sursă (sau mai multe surse ce acţionează concomitent), iar $\displaystyle\sum_{k=1}^{n} p_k = 1$.

O astfel de funcţie (entropie) se utilizează, de exemplu, în ”analiza” matematică a silabelor din limba română, caz în care, luând $n = 6496$ (numărul total al silabelor existente în limba română) şi p_k egală cu raportul dintre frecvenţa globală a silabei aflate pe poziţia k (în clasamentul obţinut prin ordonarea descrescătoare a silabelor după frecvenţa lor în limba română) şi numărul total de apariţii ale silabelor, s-a găsit valoarea lui H egală cu 8.621. Când se ia $n = 56$ (numărul total de tipuri consoană-vocală) care intră în componenţa silabelor şi p_k este probabilitatea de apariţie a tipului k (în ordinea frecvenţei de apariţie), se găseşte $H = 230$.

Relativ la funcţia H, se pune problema stabilirii clasei optime de distribuţie a variabilei aleatoare X, astfel încât valoarea $H(X)$ să fie maximă. Cu alte cuvinte, interesează care sunt probabilităţile $p_1, p_2, \ldots, p_n$ pentru care expresia lui $H(X)$, adică

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

are valoare maximă pe mulţimea $\displaystyle\left\{(p_1, p_2, \ldots, p_n) \mid p_i \in (0, 1), i \in \overline{1, n}, \sum_{k=1}^{n} p_k = 1\right\}$.

Extreme fără restricţii

Definiţia 11.1 *Fie $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ o funcţională reală pe $A \neq \varnothing$ şi $\mathrm{x}_0 \in A$.*

a) Punctul x_0 se numeşte ***punct de extrem local*** *al funcţiei f dacă există o vecinătate $V \in \mathcal{V}(\mathrm{x}_0)$ aşa încât $(f(\mathrm{x}) - f(\mathrm{x}_0))$ are semn constant sau este nulă pe $V \cap A$. Atunci $f(\mathrm{x}_0)$ se numeşte* ***valoare extremă locală a lui*** *f (pe $V \cap A$).*

b) Punctul x_0 se numeşte ***punct de maxim local*** *al funcţiei f când are loc relaţia:*

$$f(\mathrm{x}) - f(\mathrm{x}_0) \leq 0,\ \forall \mathrm{x} \in V \cap A.$$

Punctul x_0 se numeşte ***punct de minim local*** *al funcţiei f când are loc relaţia:*

$$f(\mathrm{x}) - f(\mathrm{x}_0) \geq 0,\ \forall \mathrm{x} \in V \cap A.$$

Dacă, în inegalităţile acestea, avem $f(\mathrm{x}) - f(\mathrm{x}_0) = 0$ numai când $\mathrm{x} = \mathrm{x}_0$, spunem că x_0 este un punct de ***maxim*** *(respectiv* ***minim)*** ***local, strict****.*

c) Dacă $f(\mathrm{x}) - f(\mathrm{x}_0) \leqslant 0, \forall \mathrm{x} \in A$, (respectiv $f(\mathrm{x}) - f(\mathrm{x}_0) \geqslant 0$, $\forall \mathrm{x} \in A$), atunci x_0 se numeşte ***punct de maxim absolut*** *sau* ***maxim global*** *(respectiv de* ***minim absolut*** *sau* ***minim global****) pentru f. În acest caz, $f(\mathrm{x}_0)$ este* ***valoarea maximă*** *(respectiv* ***minimă****) a lui f.*

Observaţie: Dacă x_0 este punct de extrem (maxim sau minim) absolut pentru o funcţie f, atunci el este întotdeauna punct de extrem local pentru f. Nu şi reciproc.

Definiţia 11.2

a) Problema determinării punctelor şi valorilor de extrem, locale sau globale, ale unei funcţii $f : \mathbb{R}^n \to \mathbb{R}$, în absenţa oricărei condiţii restrictive asupra argumentului lui f, se numeşte ***problemă de extrem liber*** *(****necondiţionat*** *sau* ***fără legături****).*

b) Problema determinării elementelor de extrem ale unei funcţii $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$, în condiţiile în care se cere punctului de extrem să aparţină unei anumite mulţimi nevide (de restricţii) $B \subseteq \mathbb{R}^n$, se numeşte ***problemă de extrem cu legături*** *(sau* ***problemă de extrem condiţionat****).*

Teorema 11.3 (Fermat) *Fie $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ şi fie $\mathrm{x}_0 \in \mathring{A}$. Dacă x_0 este un punct de extrem al lui f, iar f admite derivate parţiale de ordinul întâi în x_0, atunci derivatele respective se anulează în x_0.*

Demonstraţie: Dacă $\mathrm{x}_0 \in \mathring{A}$ şi există $\dfrac{\partial f}{\partial x_k}(\mathrm{x}_0)$, atunci avem: $\dfrac{\partial f}{\partial x_k}(\mathrm{x}_0) = \lim\limits_{t \to 0} \dfrac{f(\mathrm{x}_0 + t\mathbf{e}_k) - f(\mathrm{x}_0)}{t}$. Dacă x_0 este un punct de maxim al lui f, local sau global, avem: $f(\mathrm{x}_0 + t\mathbf{e}_k) - f(\mathrm{x}_0) \leq 0$, $\forall\, t \in V_0 \in \mathcal{V}(0)$, cu $V_0 = (a, b)$, $a < 0 < b$. Astfel, $\lim\limits_{\substack{t \to 0 \\ t<0}} \dfrac{f(\mathrm{x}_0 + t\mathbf{e}_k) - f(\mathrm{x}_0)}{t} \geq 0$ şi $\lim\limits_{\substack{t \to 0 \\ t>0}} \dfrac{f(\mathrm{x}_0 + t\mathbf{e}_k) - f(\mathrm{x}_0)}{t} \leq 0$. Deci $\dfrac{\partial f}{\partial x_k}(\mathrm{x}_0) = 0$.

Când x_0 este punct de minim al lui f, avem $\lim\limits_{\substack{t \to 0 \\ t<0}} \dfrac{f(\mathrm{x}_0 + t\mathbf{e}_k) - f(\mathrm{x}_0)}{t} \leq 0$ şi $\lim\limits_{\substack{t \to 0 \\ t>0}} \dfrac{f(\mathrm{x}_0 + t\mathbf{e}_k) - f(\mathrm{x}_0)}{t} \geq 0$.

Prin urmare, rezultă şi atunci că $\dfrac{\partial f}{\partial x_k}(\mathrm{x}_0) = 0$. ◀

Observaţie: Dacă $\mathrm{x}_0 \in \mathring{A}$ este un punct de extrem al unei funcţii $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$, cu gradient în x_0, atunci, potrivit Teoremei lui Fermat, avem că $(\nabla f)(\mathrm{x}_0)$ este vectorul nul $\mathbf{0}_{\mathbb{R}^n}$.

Definiţia 11.4 *Fie $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ şi $\mathrm{x}_0 \in \mathring{A}$, astfel încât f este diferenţiabilă Frèchet (de ordinul întâi) în x_0. Punctul x_0 se numeşte* ***punct critic*** *(sau* ***punct staţionar****) al funcţiei f dacă $(df)(\mathrm{x}_0) = \mathbf{0}_{\mathcal{L}(\mathbb{R}^n;\mathbb{R})}$, adică $(\nabla f)(\mathrm{x}_0) = \mathbf{0}_{\mathbb{R}^n}$.*

Observaţie: Fie $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$. Cum $((df)(\mathrm{x}_0))(\mathrm{v}) = \langle (\nabla f)(\mathrm{x}_0), \mathrm{v}\rangle$, $\forall\, \mathrm{v} \in \mathbb{R}^n$, Teorema 11.3 afirmă că orice punct de extrem ce aparţine interiorului mulţimii A şi în care f este diferenţiabilă Frèchet constituie un punct critic al lui f. Reciproca nu este adevărată.

Exemplu: Fie funcţia $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x_1, x_2) = x_1^3 + x_2^3 - 3x_1x_2$, cu punctul critic $(1, 1)$ (deoarece $\frac{\partial f}{\partial x_1}(1,1) = 3x_1^2 - 3x_2|_{x_1=1,x_2=1} = 0$ şi $\frac{\partial f}{\partial x_2}(1,1) = 3x_2^2 - 3x_1|_{x_1=1,x_2=1} = 0$). Se constată că diferenţa $f(x_1, x_2) - f(1, 1) = x_1^3 + x_2^3 - 3x_1x_2 + 1$ este, pentru $x_2 = 1$, egală cu $(x_1 - 1)(x_1 - 2)$ ceea ce înseamnă că $f(1 - a, 1) - f(1, 1) = (-a)(-1 - a) = a(1 + 1) \geq 0$, $\forall\, a \geq 0$ şi $f(1 + b, 1) - f(1, 1) = b(b - 1) \leq 0$, $\forall\, b \in [0, 1]$. Deci, în punctul $(1, 1)$, funcţia f din acest caz nu are nici minim şi nici maxim local, nesatisfăcând Definiţia 11.3.

Definiţia 11.5 *Un punct critic al unei funcţii diferenţiabile $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ care nu este punct de extrem al lui f se numeşte **punct şa** pentru funcţia f.*

Pentru funcţiile diferenţiabile de cel puţin ordinul al doilea într-un punct critic există criterii de identificare a punctelor de extrem sau a punctelor şa printre punctele critice ale respectivelor funcţii.

În acest sens, pentru cazul în care $n = 1$, are loc următorul rezultat:

Teorema 11.6 *Fie $A \subseteq \mathbb{R}$ un interval, $x_0 \in \mathring{A}$ şi $f : A \to \mathbb{R}$ o funcţie de n ori derivabilă în x_0, $n \geq 2$. Dacă $f'(x_0) = f''(x_0) = \ldots = f^{(n-1)}(x_0) = 0$ şi $f^{(n)}(x_0) \neq 0$, iar n este par, atunci x_0 este punct de extrem local pentru f şi anume: punct de minim local, când $f^{(n)}(x_0) > 0$ sau punct de maxim local, când $f^{(n)}(x_0) < 0$. Dacă n este impar, atunci x_0 nu este punct de extrem local al lui f.*

Demonstraţie: Fie funcţiile $R_f(\cdot; x_0) : A \to \mathbb{R}$ şi $g_0 : A \to \mathbb{R}$, $R_f(x; x_0) = f(x) - f(x_0) - \frac{x - x_0}{1!}f'(x_0) - \cdots - \frac{(x - x_0)^n}{n!}f^{(n)}(x_0)$ şi $g_0(x) = (x - x_0)^n$, $\forall\, x \in A$. Funcţiile $R_f(\cdot; x_0)$ şi g_0 sunt derivabile de n ori în $x_0 \in \mathring{A}$ şi $R_f^{(k)}(x_0; x_0) = g_0^{(k)}(x_0) = 0$, $\forall k \in \overline{1, n-1}$, iar $R_f^{(n)}(x_0; x_0) = 0$ şi $g_0^{(n)}(x_0) = n!$. Atunci: $\lim\limits_{x\to x_0} \frac{R_f(x; x_0)}{(x - x_0)^n} = \lim\limits_{x\to x_0} \frac{R'_f(x; x_0)}{n(x - x_0)^{n-1}} = \ldots = \lim\limits_{x\to x_0} \frac{R_f^{(n)}(x; x_0)}{g_0^{(n)}(x)} = \frac{0}{n!} = 0$. În consecinţă, avem

$$f(x) = f(x_0) + \frac{x - x_0}{1!}f'(x_0) + \cdots + \frac{(x - x_0)^n}{n!}f^{(n)}(x_0) + R_f(x; x_0), \forall\, x \in A,$$

Definind $\alpha : A \longrightarrow \mathbb{R}$ prin $\alpha(x) = \begin{cases} n!\dfrac{R_f(x; x_0)}{(x - x_0)^n} & , \text{ dacă } x \in A \setminus \{x_0\} \\ 0 & , \text{ dacă } x = x_0, \end{cases}$, putem scrie

$$f(x) = f(x_0) + \frac{x - x_0}{1!}f'(x_0) + \cdots + \frac{(x - x_0)^n}{n!}f^{(n)}(x_0) + \frac{(x - x_0)^n}{n!}\alpha(x), \forall\, x \in A,$$

unde α este continuă şi nulă în x_0. Adică, pentru f, are loc formula lui Taylor de ordin n, cu rest de tip Peano ($R_f(x; x_0) = \frac{(x - x_0)^n}{n!}\alpha(x)$) în vecinătatea lui x_0. Cum, prin ipoteză, $f'(x_0) = \ldots = f^{(n-1)}(x_0) = 0$ şi $f^{(n)}(x_0) \neq 0$, rezultă că avem:

$$f(x) = f(x_0) + \frac{(x - x_0)^n}{n!}\left[f^{(n)}(x_0) + \alpha(x)\right], \forall\, x \in A.$$

În acelaşi timp, deoarece $\lim\limits_{x\to x_0} \alpha(x) = 0$, avem:

$$\lim_{x\to x_0} \left[f^{(n)}(x_0) + \alpha(x)\right] = f^{(n)}(x_0) \neq 0.$$

Cum α este continuă în x_0, există $V \in \mathcal{V}(x_0)$ aşa încât semnul expresiei funcţiei $\alpha(x) + f^{(n)}(x_0)$ să fie constant şi egal cu semnul lui $f^{(n)}(x_0)$. Atunci $\operatorname{sign}(f(x) - f(x_0)) = \operatorname{sign}\left(\left[f^{(n)}(x_0)\right] \cdot \left[\frac{(x - x_0)^n}{n!}\right]\right)$, $\forall x \in V \cap A$ şi deci, dacă n este par, avem $(x - x_0)^n > 0$, $\forall\, x \in (V \cap A) \setminus \{x_0\}$, adică $\operatorname{sign}(f(x) - f(x_0)) = \operatorname{sign}\left(f^{(n)}(x_0)\right)$. Aşadar, când $f^{(n)}(x_0) > 0$, obţinem $f(x) > f(x_0)$, $\forall\, x \in (V \cap A) \setminus \{x_0\}$, adică x_0 este un punct de minim al lui f, iar când $f^{(n)}(x_0) < 0$, găsim că $f(x) > f(x_0)$, $\forall\, x \in (V \cap A) \setminus \{x_0\}$, adică x_0 este punct de maxim al lui f. Dacă n este impar, $(x - x_0)^n$ are semn variabil pe $V \cap A$ ceea ce ne spune că x_0 este un punct şa pentru f şi nu un punct de extrem. ◀

Teorema 11.7 *Fie $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ o funcţie ce are pe $\mathrm{x}_0 \in \mathring{A}$ ca punct critic. Dacă f are derivate parţiale de ordinul al doilea continue într-o vecinătate a lui x_0, atunci:*

i) când $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v}) \equiv \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathrm{x}_0) v_i v_j > 0, \forall \mathrm{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$, punctul x_0 este unul de minim pentru funcţia f, iar când $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v}) < 0$, $\forall\, \mathrm{v} \in \mathbb{R}^n \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$, punctul x_0 este unul de maxim pentru f;

ii) când $\left(d^2 f\right)(\mathrm{x}_0)$ este o formă pătratică nedefinită (adică există $\mathrm{v}' \in \mathbb{R}^n \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$ şi $\mathrm{v}'' \in \mathbb{R}^n \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$ astfel încât $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v}') < 0$ şi $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v}'') > 0$), atunci x_0 este punct şa pentru f;

iii) când $\left(d^2 f\right)(\mathrm{x}_0)$ este o formă pătratică pozitiv semidefinită, adică $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v}) \geq 0$, $\forall\, \mathrm{v} \in \mathbb{R}^n$ şi există $\mathrm{v}' \in \mathbb{R}^n \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$, astfel încât $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v}') = 0$, sau când $\left(d^2 f\right)(\mathrm{x}_0)$ este negativ semidefinită, adică $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v}) \leq 0$, $\forall\, \mathrm{v} \in \mathbb{R}^n$ şi există $\mathrm{v}'' \in \mathbb{R}^n \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$, astfel încât $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v}'') = 0$, nu putem stabili natura punctului staţionar x_0 cu ajutorul diferenţialei $\left(d^2 f\right)(\mathrm{x}_0)$.

Observaţie: Având în vedere că, în ipoteza $f \in \mathcal{C}^2(V)$, matricea formei pătratice $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)$, adică matricea

$$H_f(\mathrm{x}_0) \overset{not.}{=} \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathrm{x}_0)\right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}},$$

numită ***hessiana lui*** f ***în*** x_0 este simetrică, pe baza criteriului lui Schwarz (v. Curs 10). Prin urmare, toate valorile proprii ale acesteia sunt reale. Atunci, ţinând seama de Teorema 8.20 şi de Teorema de inerţie a lui Sylvester (Teorema 8.13), se poate afirma că forma pătratică $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v})$, adică $< H_f(\mathrm{x}_0)(\mathrm{v}), \mathrm{v} >_e$, este pozitiv-definită atunci când toate valorile proprii ale matricii $H_f(\mathrm{x}_0)$ sunt pozitive. Analog, forma pătratică $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)$ este negativ definită când toate valorile proprii ale matricii $H_f(\mathrm{x}_0)$ sunt negative. Dacă matricea $H_f(\mathrm{x}_0)$ are valori proprii atât din $\mathbb{R}_-^*$, cât şi din $\mathbb{R}_+^*$, atunci $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)$ este o formă pătratică nedefinită. Combinând aceste remarci cu Teorema 11.7, putem formula următorul rezultat.

Propoziţia 11.8 *Fie $A \neq \varnothing$, $A \subseteq \mathbb{R}^n$, $\mathrm{x}_0 \in \mathring{A}$, $V \in \mathcal{V}(\mathrm{x}_0)$ şi $f : A \to \mathbb{R}$ astfel încât $f \in \mathcal{C}^2(V)$. Dacă $\nabla f(\mathrm{x}_0) = \mathbf{0}_{\mathbb{R}^n}$, atunci avem că*

a) punctul critic x_0 este un punct de maxim local al lui f când $H_f(\mathrm{x}_0)$ are toate valorile proprii negative;

b) punctul critic x_0 este punct de minim local pentru f când $H_f(\mathrm{x}_0)$ are toate valorile proprii pozitive;

c) punctul critic x_0 este un punct şa pentru f când $H_f(\mathrm{x}_0)$ are cel puţin două valori proprii nenule şi de semne contrare;

d) nu putem decide natura lui x_0 când toate valorile proprii ale lui $H_f(\mathrm{x}_0)$ sunt nule.

Propoziţia 11.9 *Fie $A \neq \varnothing$, $A \subseteq \mathbb{R}^n$, $\mathrm{x}_0 \in \mathring{A}$, $V \in \mathcal{V}(\mathrm{x}_0)$ şi $f : A \to \mathbb{R}$ astfel încât $f \in \mathcal{C}^2(V)$. Fie $\Delta_0 = 1$, $\Delta_1 = \det[a_{11}]$, $\Delta_2 = \det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \ldots, \Delta_n = \det\left[(a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}\right]$ minorii principali ai matricii $(a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$, unde $a_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathrm{x}_0)$, $\forall\, i, j \in \overline{1,n}$. Atunci avem:*

i) Dacă $\Delta_j > 0$, $\forall j \in \overline{1,n}$, atunci $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)$ este pozitiv definită şi deci punctul critic x_0 este unul de minim pentru f.

ii) *Dacă* $(-1)^{j+1}\Delta_j < 0$, $\forall j \in \overline{1,n}$, *atunci* $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)$ *este negativ definită şi deci* x_0 *este punct de maxim al lui* f.

iii) *Dacă* $\Delta_j \geq 0$, $\forall j \in \overline{1,n}$ *sau* $(-1)^{j+1}\Delta_j \leq 0$, $\forall j \in \overline{1,n}$ *şi există cel puţin un rang* $i \in \overline{1,n}$ *pentru care* $\Delta_i = 0$, *în fiecare din cele două situaţii, atunci* $\left(d^2 f\right)(\mathrm{x}_0)$ *este semi-definită pozitiv, respectiv negativ, şi nu putem decide, cu ajutorul formei pătratice* $\left(d^2 f\right)(\mathrm{x}_0)$, *natura punctului* x_0.

iv) *Dacă şirul* $(\Delta_j)_{j\in\overline{1,n}}$ *nu este în nici unul dintre cazurile de la i), ii) sau iii), atunci forma pătratică* $\left(d^2 f\right)(\mathrm{x}_0)$ *este nedefinită şi deci* x_0 *nu este punct de extrem, ci punct şa al lui* f.

Observaţie: În cazul particular în care $n = 2$, Propoziţia 11.9 se poate reformula astfel:

Dacă $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$ este de clasă $\mathcal{C}^2$ pe o vecinătate a unui punct critic (pentru f) $\mathrm{x}_0 \in \mathring{A}$, atunci, notând $p = \dfrac{\partial^2 f}{\partial x_1^2}(\mathrm{x}_0)$, $q = \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(\mathrm{x}_0)$ şi $r = \dfrac{\partial^2 f}{\partial x_2^2}(\mathrm{x}_0)$, vom avea:

i) când $p > 0$ şi $pr - q^2 > 0$, punctul $\mathrm{x}_0 = \left(x_1^0, x_2^0\right)$ este de minim pentru f;

ii) când $p < 0$ şi $pr - q^2 > 0$, punctul x_0 este unul de maxim pentru f;

iii) când $pr - q^2 < 0$, x_0 nu este un punct de extrem al lui f;

iv) când $pr - q^2 = 0$, nu putem stabili natura lui x_0 prin intermediul diferenţialei a doua a lui f în x_0.

Spre exemplu, revenind la problema minimizării abaterii unor estimări cu ajutorul metodei celor mai mici pătrate, cu φ de expresie liniară, vedem 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, evident, de clasă $\mathcal{C}^2$ pe $\mathbb{R}^2$ şi are gradientul lui f nul în acel punct $\left(c_1^0, c_2^0\right)$, pentru care:

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

sau echivalent,

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

al cărui determinant este

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

Atunci $\Delta \neq 0$, când nu suntem în situaţia $a_1 = a_2 = \ldots = a_l$. Aşadar, pe baza inegalităţii lui Cauchy-Schwarz-Buniakowski $-$ obţinem:

$$c_1^0 = \frac{l\sum\limits_{k=1}^{l} a_k b_k - \left(\sum\limits_{k=1}^{l} a_k\right)\left(\sum\limits_{k=1}^{l} b_k\right)}{l\sum\limits_{k=1}^{l} a_k^2 - \left(\sum\limits_{k=1}^{l} a_k\right)^2} \quad \text{şi} \quad c_2^0 = \frac{\left(\sum\limits_{k=1}^{l} b_k\right)\left(\sum\limits_{k=1}^{l} a_k^2\right) - \left(\sum\limits_{k=1}^{l} a_k\right)\left(\sum\limits_{k=1}^{l} a_k b_k\right)}{l\sum\limits_{k=1}^{l} a_k^2 - \left(\sum\limits_{k=1}^{l} a_k\right)^2}.$$

Cum, în acest caz, avem $p = \dfrac{\partial^2 f}{\partial c_1^2}\left(c_1^0, c_2^0\right) = 2\sum\limits_{k=1}^{l} a_k^2$, $q = \dfrac{\partial^2 f}{\partial c_1 \partial c_2}\left(c_1^0, c_2^0\right) = 2\sum\limits_{k=1}^{l} a_k$ şi $r = \dfrac{\partial^2 f}{\partial c_2^2}\left(c_1^0, c_2^0\right) = 2l$, rezultă că, atâta timp cât $a_1, a_2, \ldots, a_n$ sunt nenule şi nu toate egale, suntem în situaţia în care $p > 0$ şi $pr - q^2 = 4\left[l\sum\limits_{k=1}^{l} a_k^2 - \left(\sum\limits_{k=1}^{l} a_k\right)^2\right] > 0$, ceea ce înseamnă că punctul $\left(c_1^0, c_2^0\right)$ este unul de minim pentru f.

În continuare, prezentăm câteva rezultate referitoare la noţiunile de funcţie implicită, inversabilitatea unei funcţii de mai multe variabile reale şi dependenţa funcţională a unui set de funcţii, utile pentru pentru abordarea teoretică a unei probleme de extrem cu restricţii.

Funcţii implicite

Definiţia 11.10 *Fie ecuaţia* $F(x_1, x_2, \ldots, x_n, y) = 0$, *unde* $F : A \times B \subseteq \mathbb{R}^{n+1} \to \mathbb{R}$, $A \subseteq \mathbb{R}^n$ *şi* $B \subseteq \mathbb{R}$.
O funcţie $f : A \to \mathbb{R}$ *este o* ***soluţie,*** *în raport cu* y, *a ecuaţiei* $F(x_1, x_2, \ldots, x_n, y) = 0$, *pe mulţimea* A, *dacă, pentru orice* $(x_1, x_2, \ldots, x_n) \in A$, *avem:*

$$F(x_1, x_2, \ldots, x_n, f(x_1, x_2, \ldots, x_n)) = 0.$$

O asemenea funcţie $y = f(x_1, x_2, \ldots, x_n)$, *definită prin intermediul ecuaţiei* $F(x_1, x_2, \ldots, x_n, y) = 0$ *se numeşte* ***funcţie implicită*** *sau* ***funcţie definită implicit***.

Observaţii:

a) Dacă notăm $\mathrm{x} \overset{not.}{=} (x_1, x_2, \ldots, x_n)$, atunci ecuaţia $F(x_1, x_2, \ldots, x_n, y) = 0$ se poate rescrie sub forma $F(\mathrm{x}, y) = 0$, iar soluţia $y = f(x_1, x_2, \ldots, x_n)$ se redă prin $y = f(\mathrm{x})$.

b) O ecuaţie $F(\mathrm{x}, y) = 0$ poate să aibă, pe A, mai multe soluţii sau nici una. De exemplu, ecuaţia $F(x, y) = 0$, în care $F : A \times \mathbb{R} \to \mathbb{R}$ este definită prin $F(\mathrm{x}, y) = y^2 - x$, unde $A \subseteq \mathbb{R}_-^*$, nu are nici o soluţie, pe când dacă $A \subseteq \mathbb{R}_+^*$ ea are două soluţii: $y = -\sqrt{x}$ şi $y = \sqrt{x}$.

Teorema 11.11 (Teorema funcţiilor implicite)
Fie $F : A \times B \subseteq \mathbb{R}^{n+1} \to \mathbb{R}$, $A \subseteq \mathbb{R}^n$, $B \subseteq \mathbb{R}$, $\mathrm{a} = (a_1, a_2, \ldots, a_n) \in \mathring{A}$ *şi* $b \in \mathring{B}$.
Dacă

j) $F(\mathrm{a}, b) = 0$,

jj) există $U \times V \subseteq A \times B$, *vecinătate a punctului* (a, b), *astfel încât* $F \in \mathcal{C}^1(U \times V)$,

jjj) $\dfrac{\partial F}{\partial y}(\mathrm{a}, b) \neq 0$,

atunci:

i) există $U_0 \subseteq U$, $U_0 \in \mathcal{V}(a)$, $V_0 \subseteq V$, $V_0 \in \mathcal{V}(b)$, *şi o funcţie unică* $f : U_0 \to V_0$, *astfel încât* $f(\mathrm{a}) = b$ *şi* $F\left(\mathrm{x}, f(\mathrm{x})\right) = 0$, $\forall\, \mathrm{x} \in U_0$.

ii) $f \in \mathcal{C}^1(U_0)$ *şi*

$$\frac{\partial f}{\partial x_i}(\mathrm{x}) = -\frac{\dfrac{\partial F}{\partial x_i}(\mathrm{x}, f(\mathrm{x}))}{\dfrac{\partial F}{\partial y}(\mathrm{x}, f(\mathrm{x}))}, \forall\, i \in \overline{1,n}, \forall\, \mathrm{x} \in U_0.$$

iii) *Dacă* $F \in \mathcal{C}^k(U \times V)$, $k \in \mathbb{N}^*$, *rezultă* $f \in \mathcal{C}^k(U_0)$.

Definiţia 11.10 şi Teorema 11.11 pot fi extinse la sisteme de ecuaţii şi de funcţii implicit-definite. Astfel, avem:

Definiţia 11.12 *Un sistem de ecuaţii*

$$\begin{cases} F_1(x_1, x_2, \ldots, x_n; y_1, y_2, \ldots, y_m) & = & 0 \\ F_2(x_1, x_2, \ldots, x_n; y_1, y_2, \ldots, y_m) & = & 0 \\ \vdots & & \\ F_m(x_1, x_2, \ldots, x_n; y_1, y_2, \ldots, y_m) & = & 0, \end{cases}$$

în care $F_k : A \times B \subseteq \mathbb{R}^{n+m} \to \mathbb{R}$ *sunt funcţii de expresii cunoscute, cu* $A \subseteq \mathbb{R}^n$ *şi* $B \subseteq \mathbb{R}^m$, *se numeşte* ***sistem de funcţii implicite***.

O ***soluţie*** *a unui astfel de sistem, cu necunoscutele* $y_1, y_2, \ldots, y_m$, *este un set de* m *funcţii reale (implicit-definite)*

$$\begin{cases} y_1 & = & f_1(x_1, x_2, \ldots, x_n) \\ y_2 & = & f_2(x_1, x_2, \ldots, x_n) \\ \vdots & & \\ y_m & = & f_m(x_1, x_2, \ldots, x_n), \end{cases}$$

cu $f_k : A \to \mathbb{R}$, $\forall\, k \in \overline{1,m}$, *satisfăcând relaţiile:*

$$F_k(x_1, x_2, \ldots, x_n; f_1(x_1, x_2, \ldots, x_n), \ldots, f_m(x_1, x_2, \ldots, x_n)) = 0, \forall\, k \in \overline{1,n}.$$

Teorema 11.13 (a funcţiilor vectoriale implicit-definite)

Fie sistemul de ecuaţii

$$\begin{cases} F_1(x_1, x_2, \ldots, x_n; y_1, y_2, \ldots, y_m) & = & 0 \\ F_2(x_1, x_2, \ldots, x_n; y_1, y_2, \ldots, y_m) & = & 0 \\ \vdots & & \\ F_m(x_1, x_2, \ldots, x_n; y_1, y_2, \ldots, y_m) & = & 0, \end{cases}$$

în care $F_k : A \times B \to \mathbb{R}$, $A \subseteq \mathbb{R}^n$, $B \subseteq \mathbb{R}^m$, $\forall\, k = \overline{1,n}$ *şi* $(x_1^0, x_2^0, \ldots, x_n^0; y_1^0, y_2^0, \ldots, y_m^0)$ *un punct interior al mulţimii* $A \times B$ *(adică* $(x_1^0, x_2^0, \ldots, x_n^0) \in \mathring{A}$ *şi* $(y_1^0, y_2^0, \ldots, y_m^0) \in \mathring{B}$*).*
Dacă

a) $F_k(x_1^0, x_2^0, \ldots, x_n^0; y_1^0, y_2^0, \ldots, y_m^0) = 0$, $\forall\, k \in \overline{1,n}$;

b) există $U \times V \subseteq A \times B$, *o vecinătate a punctului* $(\mathrm{x}_0, \mathrm{y}_0)$ *(cu* $\mathrm{x}_0 = (x_1^0, x_2^0, \ldots, x_n^0)$ *şi* $\mathrm{y}_0 = (y_1^0, y_2^0, \ldots, y_m^0)$*), astfel încât* $F_k \in \mathcal{C}^1(U \times V)$, $\forall\, k = \overline{1,m}$ *şi*

c) jacobianul $\dfrac{D(F_1, F_2, \ldots, F_m)}{D(y_1, y_2, \ldots, y_m)}(\mathrm{x}_0, \mathrm{y}_0) = \det \left[\left(\dfrac{\partial F_j}{\partial y_k}(\mathrm{x}_0, \mathrm{y}_0) \right)_{\substack{1 \leqslant j \leqslant m \\ 1 \leqslant k \leqslant m}} \right]$ *este diferit de zero,*

atunci:

α) există $U_0 \subseteq U \subseteq \mathbb{R}^n$, $U_0 \in \mathcal{V}(\mathrm{x}_0)$, $V_0 \subseteq V \subseteq \mathbb{R}^m$, $V_0 \in \mathcal{V}(\mathrm{y}_0)$ *şi o funcţie unică* $f : U_0 \to V_0$, $f(\mathrm{x}) = (f_1(\mathrm{x}), f_2(\mathrm{x}), \ldots, f_m(\mathrm{x}))$, *astfel încât* $f_j(\mathrm{x}_0) = y_j^0$, $\forall\, j \in \overline{1,m}$ *şi* $F_j(\mathrm{x}, f(\mathrm{x})) = 0$, $\forall\, j \in \overline{1,m}$, $\forall\, \mathrm{x} \in U_0$;

β) *Funcţiile $f_1, f_2, \ldots, f_m \in \mathcal{C}^1(U_0)$, având derivatele parţiale de ordinul întâi, pe U_0, date de formulele:*

$$\frac{\partial f_k}{\partial x_i}(\mathrm{x}) = -\frac{\dfrac{D(F_1, F_2, \ldots, F_m)}{D(y_1, \ldots, y_{k-1}, x_i, y_{k+1}, \ldots, y_m)}(\mathrm{x}, f(\mathrm{x}))}{\dfrac{D(F_1, F_2, \ldots, F_m)}{D(y_1, y_2, \ldots, y_m)}(\mathrm{x}, f(\mathrm{x}))}, \forall\, \mathrm{x} \in U_0, \forall\, k \in \overline{1, m}, i \in \overline{1, n}.$$

γ) *dacă $F_1, F_2, \ldots, F_m \in \mathcal{C}^l(U \times V)$, $l \in \mathbb{N}^*$, atunci $f_1, f_2, \ldots, f_m \in \mathcal{C}^l(U_0; V_0)$.*

Inversabilitatea funcţiilor diferenţiabile de mai multe variabile reale

Se ştie că un sistem algebric liniar de tipul

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= y_2 \\ \vdots & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= y_n \end{cases}, \text{ cu } a_{ij} \in \mathbb{R}, \forall\, i, j = \overline{1, n},$$

are o soluţie unică dacă şi numai dacă $det\left[(a_{ij})_{\substack{1\le i\le n\\ 1\le j\le n}}\right] \neq 0$. Astfel, considerăm funcţia $f : \mathbb{R}^n \to \mathbb{R}^n$, cu expresia

$$f(x_1, x_2, \ldots, x_n) = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = A \cdot \mathrm{x}.$$

Spunem că f este bijectivă, şi deci, ***inversabilă în vecinătatea oricărui punct*** $\mathrm{x}_0 = (x_1^0, x_2^0, \ldots, x_n^0) \in \mathbb{R}^n$ (adică ***adică inversabilă global***), dacă şi numai dacă $\dfrac{D(f_1, f_2, \ldots, f_n)}{D(x_1, x_2, \ldots, x_n)}(\mathrm{x}) = \det A \neq 0$.

Putem extinde un astfel de rezultat la cazul sistemelor neliniare de forma

$$\begin{cases} f_1(x_1, x_2, \ldots, x_n) = y_1 \\ f_2(x_1, x_2, \ldots, x_n) = y_2 \\ \vdots \\ f_n(x_1, x_2, \ldots, x_n) = y_n, \end{cases}$$

conform următoarei teoreme:

Teorema 11.14 (de inversabilitate locală)

Fie $A \subseteq \mathbb{R}^n$ o mulţime nevidă şi deschisă, $f : A \to \mathbb{R}^n$ o funcţie de clasă $\mathcal{C}^1$ pe A, şi fie $\mathrm{x}_0 \in A$. Dacă $(df)(\mathrm{x}_0)$ este o bijecţie de la $\mathbb{R}^n$ la $\mathbb{R}^n$, adică dacă $det(J_f(\mathrm{x}_0)) \neq 0$, atunci există $U \in \mathcal{V}(\mathrm{x}_0)$, $U \subseteq A$ şi există $V \in \mathcal{V}(f(x_0))$, astfel încât $f : U \to V$ să fie de clasă $\mathcal{C}^1$ pe V. În plus, dacă f este de clasă $\mathcal{C}^k$ de la U la V, atunci f^{-1} este tot de clasă $\mathcal{C}^k$, de la V la U.

Observaţii:

i) Teorema 11.14 poate fi reformulată după cum urmează:

"O funcţie $f : A \subseteq \mathbb{R}^n \to R^n$, $f \in \mathcal{C}^1(A)$, este inversabilă în vecinătatea oricărui punct $x_0 \in \mathring{A}$ în care (df) există şi este inversabilă (ca aplicaţie liniară de la R^n la R^n)".

ii) În condiţiile Teoremei 11.14, cu $f \in \mathcal{C}^1(U; V)$ şi $f^{-1} \in \mathcal{C}^1(V; U)$, avem:

$$(df)(\mathrm{x}) \circ \left(d\left(f^{-1}\right)\right)(f(\mathrm{x})) = \left(d\left(f^{-1}\right)\right)(f(\mathrm{x})) \circ (df)(\mathrm{x}) = I(\mathrm{x}), \forall\, \mathrm{x} \in U.$$

Aşadar, putem scrie: $\left(d\left(f^{-1}\right)\right)(f(\mathrm{x})) = ((df)(\mathrm{x}))^{-1}$.

Adică, matricea jacobiană $J_{f^{-1}}$, calculată în punctul $f(\mathrm{x})$, este inversa matricii $J_f(\mathrm{x})$. Acest fapt permite calculul derivatelor parţiale ale lui f^{-1} în funcţie de derivatele parţiale ale lui f.

Dependenţa sau independenţa funcţională a unui ansamblu de funcţii

Definiţia 11.15 *Fie $A \subseteq \mathbb{R}^n$ o mulţime nevidă, iar $f_1, f_2, \ldots, f_m : A \to \mathbb{R}$, cu $m \leq n$.*

1) *Spunem că o funcţie $f_0 : A \to \mathbb{R}$ este* ***dependentă funcţional*** *(sau* ***funcţional-dependentă) de*** *$f_1, f_2, \ldots, f_m$,* ***pe*** *$\widetilde{A} \subseteq A$ (cu $\widetilde{A} \neq \varnothing$), dacă, oricare ar fi $\mathrm{x}_0 \in \widetilde{A}$, există o vecinătate $U \in \mathcal{V}(\mathrm{x}_0)$ şi o funcţie $\varphi : V \to \mathbb{R}$, unde V este o vecinătate a punctului $(f_1(\mathrm{x}_0), f_2(\mathrm{x}_0), \ldots, f_m(\mathrm{x}_0))$, astfel încât:*

$$f_0(\mathrm{x}) = \varphi\left(f_1(\mathrm{x}), f_2(\mathrm{x}), \ldots, f_m(\mathrm{x})\right), \forall\, \mathrm{x} \in U \cap \widetilde{A}.$$

În caz contrar, vom spune că f_0 se numeşte ***funcţional-independentă de*** *$f_1, f_2, \ldots, f_m$, pe $\widetilde{A}$.*

2) *Spunem că un ansamblu finit de funcţii de la A în $\mathbb{R}$ se află în relaţie de* ***dependenţă funcţională*** *atunci când una dintre funcţii este dependentă funcţional de celelalte. În caz contrar, ansamblul respectiv este denumit* ***independent funcţional.***

Propoziţia 11.16 *Fie $A \subseteq \mathbb{R}^n$ o mulţime nevidă şi deschisă, iar $f : A \to \mathbb{R}^m$ (cu $m \leq n$) o funcţie de clasă $\mathcal{C}^1$ pe A, de componente $f_1, f_2, \ldots, f_m : A \to \mathbb{R}$, a cărei matrice jacobiană $J_f = \left(\dfrac{\partial f_i}{\partial x_j}\right)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ are rangul maxim (adică m) în orice punct din A.*

Atunci, o funcţie $f_0 : A \to \mathbb{R}$, $f_0 \in \mathcal{C}^1(A)$, depinde funcţional de $f_1, f_2, \ldots, f_m$ dacă şi numai dacă există m funcţii continue $\psi_1, \psi_2, \ldots, \psi_m : A \to \mathbb{R}$ astfel încât

$$(df_0)(\mathrm{x}) = \sum_{k=1}^{m} \psi_k(\mathrm{x})\,(df_k)(\mathrm{x}), \forall\, \mathrm{x} \in A.$$

Altfel spus, f_0 depinde funcţional de $f_1, \ldots, f_m$, dacă df_0 se poate exprima ca o combinaţie liniară de $df_1, \ldots, df_m$.

Teorema 11.17 (a dependenţei funcţionale)

Fie $A \subseteq \mathbb{R}^n$ o mulţime deschisă şi nevidă, iar $f : A \to \mathbb{R}^n$, cu $m \leq n$, de componente $f_1, f_2, \ldots, f_m : A \to \mathbb{R}$, o funcţie de clasă $\mathcal{C}^1$ pe A, a cărei matrice jacobiană J_j are rang constant – $r \leq m$ – în orice punct $\mathrm{x} \in A$. Atunci funcţiile $f_1, f_2, \ldots, f_m$ satisfac, local, $m - r$ relaţii de dependenţă funcţională.

Cu alte cuvinte, dacă $\operatorname{rang} J_f(\mathrm{x}) = r \leq m$, $\forall\, \mathrm{x} \in A$, atunci doar r dintre funcţiile $f_1, f_2, \ldots, f_m$ sunt funcţional independente, iar celelalte $m - r$ sunt dependente de acele r funcţii independente între ele.

Extreme cu legături. Metoda multiplicatorilor lui Lagrange

Am analizat, până acum, problema determinării punctelor de extrem, fără legături, ale funcţiilor de clasă $\mathcal{C}^1$ (vezi, Exemplele 1 si 2). În practică, apare însă situaţia determinării unor puncte de extrem ale funcţiei, pe anumite submulţimi ale mulţimii de definiţie. Altfel spus, se caută punctele de extrem ale funcţiei, supuse la una sau mai multe condiţii suplimentare. În asemenea cazuri, extremele respective poartă denumirea de ***extreme condiţionate*** sau, echivalent, ***extreme cu legături***.

În acest sens, analizăm următoarea situaţie:

Fie $D \subseteq \mathbb{R}^n \times \mathbb{R}^m$ este o mulţime deschisă, $n, m \in \mathbb{N}^*$, iar $f : D \to \mathbb{R}$ şi $g : D \to \mathbb{R}^m$ sunt funcţii de clasă $\mathcal{C}^1$ pe D, unde $g = (g_1, g_2, \ldots, g_m)$. Se cere să se determine punctele de extrem ale funcţiei f, supuse la condiţia suplimentară $g(\mathrm{x}, \mathrm{y}) = 0$, sau echivalent

$$(\bullet)\ \left\{\begin{array}{c} g_1(\mathrm{x}, \mathrm{y}) = 0 \\ g_2(\mathrm{x}, \mathrm{y}) = 0 \\ \vdots \\ g_m(\mathrm{x}, \mathrm{y}) = 0 \end{array}\right. ,$$

unde $\mathrm{x} \in B \subseteq \mathbb{R}^n$ şi $\mathrm{y} \in E \subseteq \mathbb{R}^m$, cu $B \times E \subseteq D$.

Notând cu $A = \{(\mathrm{x}, \mathrm{y}) \in D \mid g(\mathrm{x}, \mathrm{y}) = \mathbf{0}_{\mathbb{R}^n}\}$ şi ţinând cont că g este continuă ($g \in \mathcal{C}^1(D)$), rezultă că A este o mulţime închisă. Dacă B şi E sunt mărginite, atunci şi A este mărginită. În acest caz, fiind închisă şi mărginită, A este o mulţime compactă. Ca atare, fiind continuă pe A, f va avea, cu siguranţă, puncte de extrem în A. Se mai pune, deci, doar problema determinării lor.

Observăm că, în acest caz, nu putem aplica metoda utilizată în cazul extremelor libere, deoarece punctele lui A vor aparţine interiorului acestei mulţimi dacă şi numai dacă g este identic nulă, ceea ce ar însemna, de fapt, că nu am mai avea o problemă de extrem cu legături.

Pentru tratarea problemei de extrem cu legăturile nenule $g_1, g_2, \ldots, g_m$, am putea încerca să utilizăm teorema funcţiilor implicite (Teorema 11.11 sau Teorema 11.13) pentru a soluţiona mai întâi sistemul $g_k(\mathrm{x}, \mathrm{y}) = 0$, $\forall\, k \in \overline{1, m}$, în raport cu y. Scoţând $\mathrm{y} = \varphi(\mathrm{x})$, am putea reduce problema iniţială de extrem la o problemă de extrem liber pentru funcţia $f(\mathrm{x}, \varphi(\mathrm{x}))$, cu $\mathrm{x} \in B$. În această situaţie am putea studia extremele lui $f(\mathrm{x}, \varphi(\mathrm{x}))$ pe calea prezentată în prima parte a acestui curs. Procedeul ar fi însă dificil de folosit în practică, deoarece, în general, ecuaţia vectorială $g(\mathrm{x}, \mathrm{y}) = \mathbf{0} \in \mathbb{R}^n$ ar fi greu de rezolvat în raport cu y şi, chiar dacă, teoretic, acest lucru ar fi posibil, nu ar fi uşor de ajuns la $f(\mathrm{x}, \varphi(\mathrm{x}))$.

În continuare, vom prezenta o metodă pentru rezolvarea unei probleme de extrem condiţionat, metodă ce se aplică, în special, atunci când nu se poate pune în evidenţă, în mod explicit, forma lui $f(\mathrm{x}, \varphi(\mathrm{x}))$.

Definiţia 11.18 *Fie $D \subseteq \mathbb{R}^n \times \mathbb{R}^m$ este o mulţime deschisă, şi fie $f : D \subseteq \mathbb{R}^{n+m} \to \mathbb{R}$, cu $n, m \in \mathbb{N}^*$.*

a) Spunem că $(\mathrm{x}_0, \mathrm{y}_0) \in A$ este un ***punct de extrem local al funcţiei*** *f,* ***cu restricţiile*** *$(\bullet)$, (sau* ***punct de extrem local condiţionat al lui*** *f), dacă există o vecinătate $V \subseteq D$ a lui $(\mathrm{x}_0, \mathrm{y}_0)$ astfel încât $f(\mathrm{x}, \mathrm{y}) - f(\mathrm{x}_0, \mathrm{y}_0)$ să aibă semn constant $\forall\, (\mathrm{x}, \mathrm{y}) \in V \cap A$.*

b) Spunem că $(\mathrm{x}_0, \mathrm{y}_0) \in A$ este un ***punct de minim local condiţionat al funcţiei*** *f dacă*

$$f(\mathrm{x}, \mathrm{y}) - f(\mathrm{x}_0, \mathrm{y}_0) \geq 0,\ \forall\, (\mathrm{x}, \mathrm{y}) \in V \cap A.$$

Analog, $(\mathrm{x}_0, \mathrm{y}_0)$ este un ***punct de maxim condiţionat pentru*** *f dacă*

$$f(\mathrm{x}, \mathrm{y}) - f(\mathrm{x}_0, \mathrm{y}_0) \leq 0,\ \forall\, (\mathrm{x}, \mathrm{y}) \in V \cap A.$$

Altfel scris, $(\mathrm{x}_0, \mathrm{y}_0)$ este un punct de extrem local condiţionat pentru f dacă el este un punct de extrem (minim şi respectiv maxim) local pentru restricţia funcţiei f la mulţimea de legături (restricţii) A.

c) Un punct $(\mathrm{x}_0, \mathrm{y}_0) \in A$ se numeşte ***punct critic condiţionat*** *(sau* ***punct staţionar condiţionat****) pentru f, dacă el este un punct critic (staţionar) al restricţiei lui f la A.*

Teorema 11.19 (de existenţă a multiplicatorilor lui Lagrange)

Fie D o mulţime nevidă şi deschisă din $\mathbb{R}^n \times \mathbb{R}^m$, $f : D \to \mathbb{R}$ şi $g : D \to \mathbb{R}^m$, $g = (g_1, g_2, \ldots, g_m)$, cu $g_k : D \to \mathbb{R}$, $\forall\, k \in \overline{1, m}$, funcţii de clasă $\mathcal{C}^1$ pe D, iar $(\mathrm{x}_0, \mathrm{y}_0) \in D$ un punct de extrem local condiţionat al lui f cu restricţiile $g_k(\mathrm{x}_0, \mathrm{y}_0) = 0$, $\forall\, k \in \overline{1, m}$.

Dacă $\dfrac{D(g_1, g_2, \ldots, g_m)}{D(y_1, y_2, \ldots, y_m)}(\mathrm{x}_0, \mathrm{y}_0) \neq 0$, adică dacă $g_1, g_2, \ldots, g_m$ sunt funcţional independente, în raport cu $y_1, y_2, \ldots, y_m$, în $(\mathrm{x}_0, \mathrm{y}_0)$, atunci există m numere reale $\lambda_1, \lambda_2, \ldots, \lambda_m$, astfel încât, dacă se consideră funcţia

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

punctul $(\mathrm{x}_0, \mathrm{y}_0)$ să-i fie punct critic, adică să satisfacă sistemul de relaţii:

$$(*) \quad \begin{cases} \dfrac{\partial L}{\partial x_k}(\mathrm{x}, \mathrm{y}) & = & 0, & \forall\, k \in \overline{1, n} \\ \dfrac{\partial L}{\partial y_j}(\mathrm{x}, \mathrm{y}) & = & 0, & \forall\, j \in \overline{1, m} \\ g_i(\mathrm{x}, \mathrm{y}) & = & 0, & \forall\, i \in \overline{1, m}. \end{cases}$$

Demonstraţie: Prin aplicarea teoremei funcţiilor implicite 11.13 asupra sistemului de legături $g_k(\mathrm{x},\mathrm{y}) = 0$, $\forall\, k \in \overline{1,m}$, se poate spune că există $U \in \mathcal{V}(x_0)$, şi o funcţie $\varphi : U \to \mathbb{R}^n$, de clasă $\mathcal{C}^1$ pe U (adică există $\varphi_k : U \to \mathbb{R}$, $\forall\, k \in \overline{1,m}$, $\varphi_k \in \mathcal{C}^1(U)$) aşa încât $\varphi(\mathrm{x}_0) = \mathrm{y}_0$ şi $g_k(\mathrm{x},\varphi(\mathrm{x})) = 0$, $\forall\, k \in \overline{1,m}$, $\forall\, \mathrm{x} \in U$. Aşadar, prin derivare obţinem:

$$(**)\quad \frac{\partial g_k}{\partial x_i}(\mathrm{x},\varphi(\mathrm{x})) + \sum_{j=1}^{m}\left(\frac{\partial g_k}{\partial y_j}\cdot\frac{\partial \varphi_j}{\partial x_i}\right)(\mathrm{x},\varphi(\mathrm{x})) = 0, \forall\, k \in \overline{1,m}, i \in \overline{1,n}.$$

Fie funcţia $F(\mathrm{x}) = f(\mathrm{x},\varphi(\mathrm{x}))$, $\forall\, \mathrm{x} \in U$. Cum $(\mathrm{x}_0,\mathrm{y}_0)$ este un punct de extrem local, cu restricţiile $g_k(\mathrm{x},\mathrm{y}) = 0$, $\forall\, k \in \overline{1,m}$, pentru f, iar $(\mathrm{x},\varphi(\mathrm{x})) \in A$, $\forall\, \mathrm{x} \in U$, rezultă că punctul x_0 este unul de extrem local necondiţionat, pentru F pe U. Atunci, potrivit Teoremei lui Fermat, avem $\dfrac{\partial F}{\partial x_k}(\mathrm{x}_0) = 0$, $\forall\, k \in \overline{1,n}$. Cu alte cuvinte, avem:

$$(!)\quad \frac{\partial f}{\partial x_k}(\mathrm{x}_0,\mathrm{y}_0) + \sum_{j=1}^{m}\frac{\partial f}{\partial y_j}(\mathrm{x}_0,\mathrm{y}_0)\frac{\partial \varphi_j}{\partial x_k}(\mathrm{x}_0) = 0, \forall\, k \in \overline{1,n}.$$

Cum, $\dfrac{D(g_1,g_2,\ldots,g_m)}{D(y_1,y_2,\ldots,y_m)}(\mathrm{x}_0,\mathrm{y}_0) \neq 0$, reiese că sistemul liniar algebric

$$(***)\quad \sum_{k=1}^{m}\frac{\partial g_k}{\partial y_j}(\mathrm{x}_0,\mathrm{y}_0)\alpha_k = -\frac{\partial f}{\partial y_j}(\mathrm{x}_0,\mathrm{y}_0), \forall\, j \in \overline{1,m}$$

are o soluţie unică. Fie aceasta notată cu $(\lambda_1^0,\lambda_2^0,\ldots,\lambda_m^0)$. În raport cu ea, funcţia $L = f + \sum\limits_{k=1}^{m}\lambda_k^0 g_k$ satisface relaţiile din sistemul $(*)$. Într-adevăr, cum $(\mathrm{x}_0,\mathrm{y}_0) \in A$, avem $g_k(\mathrm{x}_0,\mathrm{y}_0) = 0$, $\forall\, k \in \overline{1,m}$.

În plus, calculând derivatele parţiale ale lui L în raport cu x_i, găsim

$$\begin{aligned}
\frac{\partial L}{\partial x_i}(\mathrm{x}_0,\mathrm{y}_0) &= \frac{\partial f}{\partial x_i}(\mathrm{x}_0,\mathrm{y}_0) + \sum_{k=1}^{m}\lambda_k^0\frac{\partial g_k}{\partial x_i}(\mathrm{x}_0,\mathrm{y}_0) = \\
&= \frac{\partial f}{\partial x_i}(\mathrm{x}_0,\mathrm{y}_0) - \sum_{k=1}^{m}\lambda_k^0\cdot\left(\sum_{j=1}^{m}\frac{\partial g_k}{\partial y_j}(\mathrm{x}_0,\mathrm{y}_0)\cdot\frac{\partial \varphi_j}{\partial x_i}(\mathrm{x}_0)\right) = \\
&= \frac{\partial f}{\partial x_i}(\mathrm{x}_0,\mathrm{y}_0) - \sum_{k=1}^{m}\frac{\partial \varphi_j}{\partial x_i}(\mathrm{x}_0)\cdot\left(\sum_{j=1}^{m}\lambda_k^0\cdot\frac{\partial g_k}{\partial y_j}(\mathrm{x}_0,\mathrm{y}_0)\right) = \\
&= \frac{\partial f}{\partial x_i}(\mathrm{x}_0,\mathrm{y}_0) + \sum_{k=1}^{m}\frac{\partial \varphi_j}{\partial x_i}(\mathrm{x}_0)\cdot\frac{\partial f}{\partial y_j}(\mathrm{x}_0,\mathrm{y}_0),
\end{aligned}$$

dacă se ţine seama şi de relaţiile $(**)$ şi $(***)$. Folosind (!), obţinem $\dfrac{\partial L}{\partial x_i}(\mathrm{x}_0,\mathrm{y}_0) = 0$, $\forall\, i \in \overline{1,n}$. Derivând acum L în raport cu y_j şi ţinând cont de faptul că $(\lambda_1^0,\lambda_2^0,\ldots,\lambda_m^0)$ este soluţia sistemului $(***)$, obţinem:

$$\frac{\partial L}{\partial y_j}(\mathrm{x}_0,\mathrm{y}_0) = \frac{\partial f}{\partial y_j} + \sum_{k=1}^{m}\lambda_k^0\frac{\partial g_k}{\partial y_j}(\mathrm{x}_0,\mathrm{y}_0) = 0, \forall j \in \overline{1,m}.$$

◀

Precizări:

1) Numerele $\lambda_1,\lambda_2,\ldots,\lambda_m$ se numesc *multiplicatori Lagrange.*

2) Punctele de extrem condiţionat pentru funcţia f se găsesc printre punctele critice, condiţionate, ale lui f, adică printre punctele critice ale funcţiei asociate $L = f + \sum\limits_{k=1}^{m}\lambda_k g_k$.

3) Dacă, în expresia $f(\mathrm{x},\mathrm{y})+\sum\limits_{k=1}^{m}\lambda_k g_k(\mathrm{x},\mathrm{y})$, admitem că şi $\lambda_1,\lambda_2,\ldots,\lambda_m$ sunt variabile, atunci funcţia
$$(\mathrm{x},\mathrm{y},\lambda)\longmapsto L(\mathrm{x},\mathrm{y},\lambda)=f(\mathrm{x},\mathrm{y})+\sum_{k=1}^{m}\lambda_k g_k(\mathrm{x},\mathrm{y})$$
poartă denumirea de *Lagrangean* asociat lui f şi restricţiilor g_k, $\forall\, k\in\overline{1,m}$.

4) Dacă f, şi g_k sunt funcţii de clasă $\mathcal{C}^2$ pe D, $\forall\, k\in\overline{1,n}$, sau cel puţin pe o vecinătate a unei soluţii $(\mathrm{x}_0,\mathrm{y}_0;\lambda_0)=(\mathrm{x}_1^0,\mathrm{x}_2^0,\ldots,\mathrm{x}_n^0,\mathrm{y}_1^0,\mathrm{y}_2^0,\ldots,\mathrm{y}_m^0,\lambda_1^0,\lambda_2^0,\ldots,\lambda_m^0)$ a sistemului de ecuaţii $(*)$, atunci, în virtutea Teoremei 11.19, punctul $(\mathrm{x}_0,\mathrm{y}_0)$ este unul critic pentru $L(\mathrm{x}_0,\mathrm{y}_0;\lambda_0)$, adică avem $d\,(L(\mathrm{x},\mathrm{y};\lambda_0))\,(\mathrm{x}_0,\mathrm{y}_0)=0$. Totodată, deoarece $L(\mathrm{x},\mathrm{y};\lambda_0)-L(\mathrm{x}_0,\mathrm{y}_0;\lambda_0)=f(\mathrm{x},\mathrm{y})-f(\mathrm{x}_0,\mathrm{y}_0)$, $\forall\,(\mathrm{x},\mathrm{y})\in A$, rezultă că $(\mathrm{x}_0,\mathrm{y}_0)$ este un punct de extrem, condiţionat de legăturile $g_k(\mathrm{x},\mathrm{y})=0$, $\forall\, k\in\overline{1,m}$, pentru f, dacă şi numai dacă el este un punct de extrem liber pentru $L(\cdot\,,\cdot\,;\lambda_0)$. Prin urmare, în conformitate cu Teorema 11.7, este necesară studierea formei pătratice $\left(d^2L(\mathrm{x},\mathrm{y};\lambda_0)\right)(\mathrm{x}_0,\mathrm{y}_0)=\sum\limits_{i,j=1}^{n+m}\dfrac{\partial^2 L}{\partial\widetilde{x}_i\partial\widetilde{x}_j}(\mathrm{x}_0,\mathrm{y}_0)\,d\widetilde{x}_i d\widetilde{x}_j$, unde $\widetilde{\mathrm{x}}=(\widetilde{x}_1,\widetilde{x}_2,\ldots,\widetilde{x}_{n+m})=(\mathrm{x},\mathrm{y})$. Cum diferenţialele $d\widetilde{x}_1,d\widetilde{x}_2,\ldots,d\widetilde{x}_{n+m}$ nu sunt independente, relaţiile de dependenţă dintre ele se obţin prin diferenţierea ecuaţiilor $g_k(\widetilde{\mathrm{x}})=0$, $\forall\, k\in\overline{1,m}$, anume: $\sum\limits_{k=1}^{n+m}\dfrac{\partial g_i}{\partial\widetilde{x}_j}(\widetilde{\mathrm{x}}_0)\,d\widetilde{x}_k=0$. Cum $\dfrac{D(g_1,g_2,\ldots,g_m)}{D(\widetilde{x}_{n+1},\widetilde{x}_{n+2},\ldots,\widetilde{x}_{n+m})}(\mathrm{x}_0,\mathrm{y}_0)\neq 0$, putem găsi $d\widetilde{x}_{n+1},d\widetilde{x}_{n+2},...,d\widetilde{x}_{n+m}$ în funcţie de $d\widetilde{x}_1,d\widetilde{x}_2,\ldots,d\widetilde{x}_n$, prin expresii liniare. Aceste expresii, introduse în $\sum\limits_{i,j=1}^{n+m}\dfrac{\partial^2 L}{\partial\widetilde{x}_i\partial\widetilde{x}_j}(\mathrm{x}_0,\mathrm{y}_0)\,d\widetilde{x}_i d\widetilde{x}_j$ conduc la relaţia $d^2L(\widetilde{\mathrm{x}}_0)=\sum\limits_{i,j=1}^{n}a_{ij}^0\,dx_i dx_j$, reprezentând o formă pătratică pe $\mathbb{R}^n$, în care a_{ij}^0 depind liniar de $\dfrac{\partial^2 g_k}{\partial\widetilde{x}_i\partial\widetilde{x}_j}(\mathrm{x}_0,\mathrm{y}_0)$, $\dfrac{\partial g_i}{\partial x_j}(\mathrm{x}_0,\mathrm{y}_0)$ şi $\dfrac{\partial^2 f}{\partial x_i\partial x_j}(\mathrm{x}_0,\mathrm{y}_0)$. Astfel, pe baza Teoremei 11.7, deducem că, dacă $d^2L(\widetilde{\mathrm{x}}_0)$ este o formă pătratică pozitiv definită (respectiv negativ definită), atunci $\widetilde{\mathrm{x}}_0$ este punct de minim (respectiv maxim) condiţionat, local, pentru f, iar dacă $d^2L(\widetilde{\mathrm{x}}_0)$ reprezintă o formă pătratică nedefinită, atunci $\widetilde{\mathrm{x}}_0$ nu este punct de extrem pentru f, condiţionat de legăturile $g_k(\widetilde{\mathrm{x}})=0$, $\forall\, k=\overline{1,m}$. În fine, dacă $d^2L(\widetilde{\mathrm{x}}_0)$ este semi-definită (pozitiv sau negativ), atunci nu putem stabili natura punctului $\widetilde{\mathrm{x}}_0$ numai pe baza diferenţialei a doua a lui L.

Exemplu: Aplicând acum metoda multiplicatorilor lui Lagrange pentru Exemplul 3, adică în cazul problemei entropiei lui Shannon, adică în cazul problemei de extrem pentru funcţia
$$\widetilde{H}(p_1,p_2,\ldots,p_n)=-\sum_{k=1}^{n}p_k\log_2 p_k,\ \text{unde } p_k\in(0,1),\forall\, k\in\overline{1,n},$$
cu condiţia $\sum\limits_{k=1}^{n}p_k=1$. Observăm că $m=1$, $g(p_1,p_2,\ldots,p_n)=\sum\limits_{k=1}^{n}p_k-1$ şi Lagrangeanul funcţiei este:
$$L(p_1,p_2,\ldots,p_n;\lambda_1)=-\sum_{k=1}^{n}p_k\log_2 p_k+\lambda_1\left(\sum_{k=1}^{n}p_k-1\right).$$
Atunci sistemul de tip $(*)$, este următorul:
$$\begin{cases}\dfrac{\partial L}{\partial p_k}(p_1,p_2,\ldots,p_n;\lambda_1) & = & -\left(\log_2 p_k+\dfrac{1}{p_k\ln 2}\right)+\lambda_1=0,\forall\, k\in\overline{1,n}\\[2ex] \dfrac{\partial L}{\partial\lambda_1}(p_1,p_2,\ldots,p_n;\lambda_1) & = & \sum\limits_{k=1}^{n}p_k-1=0.\end{cases}$$
Din primele n ecuaţii, reiese că $p_1=p_2=\ldots=p_n=\psi(\lambda_1)$, iar din ultima ecuaţie, obţinem:
$$p_1^0=p_2^0=\ldots=p_n^0=\frac{1}{n}=\psi(\lambda_1^0),\lambda_1^0=\log_2\frac{1}{n}+\frac{n}{\ln 2}.$$

Pe baza legăturii $g_1(p_1, p_2, \ldots, p_n) = 0$, deducem că $dp_1 + dp_2 + \cdots + dp_n = 0$. Astfel, avem:

$$\left(\left(d^2 L\right)(p_1^0, p_2^0, \ldots, p_n^0; \lambda_1^0)\right)(\mathrm{u}) < 0, \forall\, \mathrm{u} \in \mathbb{R}^n \setminus \{\mathbf{0}_{\mathbb{R}^n}\},$$

adică punctul $(p_1^0, p_2^0, \ldots, p_n^0) = \left(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\right)$ este unul de maxim pentru $\widetilde{H}$, în condiţia $g_1(p_1, p_2, \ldots, p_n) = 0$.

Prin urmare, entropia informaţională maximă se obţine pentru variabila aleatoare

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

a cărei distribuţie de probabilitate este uniformă.

Bibliografie recomandată

1. Anca Precupanu - *Bazele analizei matematice (§11.6)*, Editura Polirom, Iaşi, 1998.

2. Emil Popescu - *Analiză matematică. Calcul diferenţial (cap. 7 şi 8)*, Editura Matrix Rom, Bucureşti, 2006.

3. Maria Gorunescu - *Lecţii de analiză matematică pentru informaticieni (cap. VII)*, Reprografia Universităţii Craiova, 2000.

4. G. Păltineanu - *Analiză matematică. Calcul diferenţial,* Editura AGIR, Bucureşti, 2002.

5. F. Iacob - *Matematică pentru anul II - ID*, seria 2004-2005 (tema 3, modulul 3).

6. Rodica Luca-Tudorache - *Analiză matematică. Calcul diferenţial (§6.6)*, Editura Tehnopress, Iaşi, 2005.

7. V. Postolică - *Eficienţă prin matematică aplicată. Analiză matematică (§7.4)*, Editura Matrix Rom, Bucureşti, 2006.

8. M. Postolache - *Analiză matematică (teorie şi aplicaţii)*, Editura "Fair Partners", Bucureşti, 2011.

9. R. Larson, B. Edwards - *Calculus (Ninth Edition)*, Brooks/Cole-Cengage Learning, 2014.

10. Claudio Canuto, Anita Tabacco - *Mathematical Analysis II,* Springer, 2015.

11. Niels Jacob, Kristian P. Evans - *A Course in Analysis (Vol. II)*, World Scientific, 2016.