Cursul 11

Aplicaţii ale diferenţiabilităţii funcţiilor. (Extreme fără restricţii. Funcţii implicite. Inversabilitatea funcţiilor de mai multe variabile reale. Dependenţă/independenţă funcţională. Extreme condiţionate.)

Una dintre aplicaţiile de certă importanţă practică 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. După exemplificarea unor astfel de probleme, prezentăm aici chestiunile teoretice necesare abordării şi tratării cazului ce priveşte extremele necondiţionate, iar apoi, pe baza expunerii unor concepte şi rezultate din domeniile funcţiilor implicit-definite, inversabilităţii funcţiilor reale vectorial-vectoriale şi dependenţei funcţionale a unui set de funcţii, considerăm cazul extremelor cu legături.

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

Exemplul 11.1 ***Metoda celor mai mici pătrate pentru minimizarea abaterii unor estimări faţă de determinări***

Admitem că, în urma unor experimente asupra unei anumite mărimi fizice, s-au obţinut valorile $b_1, b_2, \ldots, b_p$, corespunzătoare valorilor (de "intrare") $a_1, a_2, \ldots, a_p$ (unde $p \in \mathbb{N}^*$). Reprezentând punctele (a_k, b_k) $(k = \overline{1,p})$ într-un reper ortogonal din plan, facem o apreciere asupra naturii (formei) expresiei (graficului) funcţiei φ care, necunoscută iniţial, ar avea, în a_k, valoarea b_k, $\forall\, k = \overline{1,p}$. Potrivit acestei aprecieri, estimăm că φ ar avea o expresie de un anumit tip (polinomial, exponenţial, trigonometric etc.), ale cărei caracteristici (parametri) $c_1, c_2, \ldots, c_n$ (din $\mathbb{R}$), nesupuse vreunei condiţii restrictive, se cer a fi identificate. 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 (condiţii), adică prin găsirea soluţiei (când aceasta există şi este unică) $(c_1^0, c_2^0, \ldots, c_n^0) \in \mathbb{R}^n$, 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 $y = \varphi(x; c_1^0, c_2^0, \ldots, c_n^0)$.

Precizăm (în context) că dacă imaginea grafică a mulţimii $\{(a_k, b_k) \mid k = \overline{1,p}\}$ ne sugerează faptul că φ ar fi să aibă 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.

Exemplul 11.2 ***Realizarea unui profit maxim sau a unui cost minim într-o producţie economică***

Într-o teorie economică, spaţiul $\mathbb{R}^n$ se interpretează ca fiind spaţiul complexelor de bunuri de consum, în care fiecare "bun" (produs) este caracterizat de un anumit indice $i \in \{1, 2, \ldots, n\}$, iar un "complex de bunuri" este un vector $\mathrm{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, unde componenta x_i înseamnă cantitatea în care se găseşte "bunul" i. "Unitatea" bunului i este $\mathbf{e}_i = (0, 0, \ldots, 0, 1, 0, \ldots, 0)$. Dreapta reală este interpretată ca mulţimea "valorilor", exprimată în "unităţi de cont". Într-un astfel de context, un "sistem de preţuri" este o funcţie ce asociază fiecărui "complex de bunuri" o anumită valoare. Se consideră, în mod firesc, 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 $< \mathrm{p}, \mathrm{x} >_e$, anume $\sum\limits_{i=1}^{n} p_i x_i$.

Considerând mulţimea $\mathbb{R}^n_+$, a vectorilor $\mathrm{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ cu componente nenegative, comportamentul unui consumator este apreciat în teoria economică printr-o funcţie de utilitate $u : \mathbb{R}^n_+ \to \mathbb{R}$, funcţie ce poate determina o aşa-numită "relaţie de preferinţă" "$\prec$" pe mulţimea $\mathbb{R}^n_+$, definită prin: $\mathrm{y} \prec \mathrm{x} \Leftrightarrow u(\mathrm{x}) \leq u(\mathrm{y})$.

Dacă o firmă (întreprindere) 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 intreprinderi (firme), notată cu K, este o submulţime proprie a lui $\mathbb{R}^n_+$ ce conţine $\mathbf{0}_{\mathbb{R}^n}$, problema optimizării funcţiei de utilitate 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 obiectiv este pătratică sau convexă, iar K este convexă, atunci problema de optimizare considerată se numeşte problemă de optimizare pătratică şi respectiv convexă, cu restricţii.

Exemplul 11.3 ***Problema entropiei informaţionale maxime***

Introdusă, ca noţiune matematică, de Claude E. Shannon (1947), *entropia* reprezintă o funcţie ce corespunde cantităţii de informaţie livrată (oferită) 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 este probabilitatea ($p_k \in (0,1)$) cu care se transmit (se percep) k informaţii transmise de o sursă (sau mai multe surse ce acţionează concomitent), iar $\sum\limits_{k=1}^{n} p_k = 1$.

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 silablei 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 $H = 230$.

Tot o funcţie de entropie de tip Shannon este şi cea de expresie

$$-\sum_{i=0}^{N} \frac{A_i}{4\pi} \cdot \log_2 \left(\frac{A_i}{4\pi}\right),$$

unde A_i este suprafaţa (aria) proiectată a feţei i a unui obiect ce se doreşte observat dintr-un anumit punct P din spaţiu, 4π este măsura unghiului solid (sferic) de rază 1, iar $\frac{A_i}{4\pi}$ reprezintă gradul de vizibilitate, pentru faţa de rang i, din punctul P. O asemenea funcţie se foloseşte în scopul alegerii celei mai bune poziţii a punctului P pentru observarea obiectului cu feţele $1, 2, \ldots, N$.

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

Extreme libere (fără legături)

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

Definiţia 11.1 *a) Punctul* x_0 *se numeşte* ***punct de extrem local*** *(similar spus,* ***relativ****) 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*** *(respectiv* ***punct de minim****)* ***local*** *(sau* ***relativ****) al funcţiei* f *când are loc relaţia:*

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

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

Observaţie: Orice punct de tipul x_0, de extrem (maxim sau minim) absolut pentru o funcţie f, este întotdeauna punct de extrem relativ, pentru f. Nu şi reciproc.

Definiţia 11.2 *a) Problema determinării punctelor şi valorilor de extrem, locale sau globale, ale unei funcţii (funcţionale)* $f : \mathbb{R}^n \to \mathbb{R}$*, 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 (de restricţii) $B \subseteq \mathbb{R}^n$ ($B \neq \varnothing$), se numeşte **problemă de extrem (maxim** sau **minim) cu legături** (sau **problemă de extrem condiţionat**).*

Teorema 11.1 ***(Fermat)***

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

Demonstraţie: Dacă $\mathrm{x}_0 \in A$ şi există $\dfrac{\partial f}{\partial x_k}(\mathrm{x}_0)$, atunci, în conformitate cu Definiţia 10.5 d), avem: $\dfrac{\partial f}{\partial x_k}(\mathrm{x}_0) = \lim\limits_{t\to 0} \dfrac{f(\mathrm{x}_0 + t\mathbf{e}_k) - f(\mathrm{x}_0)}{t}$. În cazul în care x_0 este un punct de maxim al lui f, local sau global, înregistrăm: $f(\mathrm{x}_0 + t\mathbf{e}_k) - f(\mathrm{x}_0) \leq 0$, $\forall\, t \in V_0 \in \mathcal{V}(0)$, 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 11.1, gradientul lui f în x_0, adică $(\nabla f)(\mathrm{x}_0)$ este vectorul nul $\mathbf{0}_{\mathbb{R}^n}$.

Definiţia 11.3 *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: Întrucât $((df)(\mathrm{x}_0))(\mathrm{v}) = \langle (\nabla f)(\mathrm{x}_0), \mathrm{v}\rangle$, $\forall\, \mathrm{v} \in \mathbb{R}^n$,Teorema 11.1 afirmă că orice punct de extrem (local) care aparţine interiorului mulţimii A (de definiţie a funcţionalei $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$) şi în care f este diferenţiabilă Frèchet constituie un punct critic al lui f. Reciproca nu este adevărată, după cum se poate vedea în cazul funcţiei $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x_1, x_2) = x_1^3 + x_2^3 - 3x_1x_2$, cu punctul critic $(1, 1)$ (deoarece $\dfrac{\partial f}{\partial x_1}(1, 1) = 3x_1^2 - 3x_2|_{x_1=1,x_2=1} = 0$ şi $\dfrac{\partial f}{\partial x_2}(1, 1) = 3x_2^2 - 3x_1|_{x_1=1,x_2=1} = 0$). 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 11.1.

Definiţia 11.4 *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 (condiţii suficiente) de discernere a punctelor de extrem sau a punctelor şa, adică criterii de identificare a punctelor de extrem sau şa printre punctele critice ale respectivelor funcţii.

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

Teorema 11.2 *Fie $A \subseteq \mathbb{R}$ un interval, $x_0 \in \mathring{A}$ şi $f : A \to \mathbb{R}$ o funcţie de n ($n \geq 2$) ori derivabilă în x_0. 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, în acest context, n este impar, atunci x_0 nu este punct de extrem local al lui f.*

Demonstraţie: Considerând funcţiile $R_f(\cdot; x_0) : A \to \mathbb{R}$ şi $g_0 : A \to \mathbb{R}$, definite prin $R_f(x; x_0) = f(x) - f(x_0) - \frac{x - x_0}{1!} f'(x_0) - \cdots - \frac{(x - x_0)^n}{n!} f^{(n)}(x_0)$ şi $g_0(x) = (x - x_0)^n$, $\forall\, x \in A$, vedem că $R_f(\cdot; x_0)$ şi g_0 sunt derivabile de n ori în $x_0 \in \mathring{A}$ şi $R_f^{(k)}(x_0; x_0) = g_0^{(k)}(x_0) = 0$, $\forall\, k = \overline{1, n-1}$, 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

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

cu $R_f(\cdot; x_0) : A \to \mathbb{R}$ astfel încât $\lim\limits_{x \to x_0} \frac{R_f(x; x_0)}{(x - x_0)^n} = 0$. Definind $\alpha : A \longrightarrow \mathbb{R}$ prin

$$\alpha(x) = \begin{cases} n! \dfrac{R_f(x; x_0)}{(x - x_0)^n} & , \quad \text{dacă } x \in A \setminus \{x_0\} \\ 0 & , \quad \text{dacă } x = x_0, \end{cases}$$

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, $f'(x_0) = f''(x_0) = \ldots = f^{(n-1)}(x_0) = 0$ 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:

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

Graţie continuităţii lui α în x_0, există $V \in \mathcal{V}(x_0)$ aşa încât, pe baza relaţiei de imediat mai sus, semnul expresiei funcţiei $\alpha(x) + f^{(n)}(x_0)$ este constant şi anume egal cu semnul numărului $f^{(n)}(x_0)$. Atunci $\text{sign}\,(f(x) - f(x_0)) = \text{sign} \left(\left[f^{(n)}(x_0) \right] \cdot \left[\frac{(x - x_0)^n}{n!} \right] \right)$, $\forall\, x \in V \cap A$ şi deci, dacă n este par, avem $(x - x_0)^n > 0$, $\forall\, x \in (V \cap A) \setminus \{x_0\}$, ceea ce implică faptul că $\text{sign}\,(f(x) - f(x_0)) = \text{sign} \left(f^{(n)}(x_0) \right)$. În consecinţă, când $f^{(n)}(x_0) > 0$, obţinem $f(x) > f(x_0)$, $\forall\, x \in (V \cap A) \setminus \{x_0\}$, ceea ce spune că x_0 este un punct de minim al lui f, iar când $f^{(n)}(x_0) < 0$, găsim că $f(x) > f(x_0)$, $\forall\, x \in (V \cap A) \setminus \{x_0\}$, aceasta însemnând că x_0 este punct de maxim al lui f. Dacă n este impar, $(x - x_0)^n$ are semn variabil pe $V \cap A$ şi, la fel, se va petrece atunci cu semnul diferenţei $f(x) - f(x_0)$, ceea ce ne spune că x_0 este un punct şa pentru f şi nu un punct de extrem. ◀

Teorema 11.3 *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 $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v}) \equiv \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} v_i v_j > 0$, $\forall\, \mathrm{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$, 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 $\exists\, v' \neq \mathbf{0}_{\mathbb{R}^n}$ şi $v'' \neq \mathbf{0}_{\mathbb{R}^n}$ din $\mathbb{R}^n$ astfel încât $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v}') < 0$ şi $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v}'') > 0$, x_0 este punct şa pentru f (nefiind punct de extrem);

iii) când $\left(d^2 f\right)(\mathrm{x}_0)$ este o formă pătratică semidefinită pozitiv (sau negativ), 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 $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v}') = 0$ (respectiv $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v}) \leq 0$, $\forall\, \mathrm{v} \in \mathbb{R}^n$ 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 $\left(d^2 f\right)(\mathrm{x}_0)$.

Demonstraţie: Deoarece f are derivate parţiale de ordinul al doilea continue pe o vecinătate V a lui x_0 (vecinătate pe care, fără a restrânge generalitatea prezentului raţionament, o putem consideră de tip sferic, adică, mai precis, $V = S(\mathrm{x}_0; r) \subseteq A$), f este diferenţiabilă (Fréchet) de două ori pe V şi deci există $\alpha : V \to \mathbb{R}$, aşa încât $\lim\limits_{x \to x_0} \alpha(\mathrm{x}) = \alpha(\mathrm{x}_0) = 0$ şi

$$f(\mathrm{x}) = f(\mathrm{x}_0) + \frac{1}{1!}\left(df(\mathrm{x}_0)\right)(\mathrm{x} - \mathrm{x}_0) + \frac{1}{2!}\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{x} - \mathrm{x}_0) + \frac{\alpha(x)}{2!}\|\mathrm{x} - \mathrm{x}_0\|^2, \forall\, \mathrm{x} \in V,$$

în conformitate cu formula lui Taylor pentru funcţii de mai multe variabile.

Cum x_0 este punct staţionar pentru f, avem $\left(df(\mathrm{x}_0)\right)(\mathrm{x} - \mathrm{x}_0) = 0$, $\forall\, \mathrm{x} \in V$. Atunci:

$$f(\mathrm{x}) - f(\mathrm{x}_0) = \frac{\|\mathrm{x} - \mathrm{x}_0\|^2}{2!}\left[\left(\left(d^2 f\right)(\mathrm{x}_0)\right)\left(\frac{\mathrm{x} - \mathrm{x}_0}{\|\mathrm{x} - \mathrm{x}_0\|}\right) + \alpha(\mathrm{x})\right], \forall\, \mathrm{x} \in V \setminus \{\mathrm{x}_0\}.$$

Totodată, mulţimea $\mathcal{M} = \left\{\mathrm{y} = \frac{\mathrm{x} - \mathrm{x}_0}{\|\mathrm{x} - \mathrm{x}_0\|} \mid \mathrm{x} \in V \setminus \{\mathrm{x}_0\}\right\} \subseteq \{\mathrm{z} \in \mathbb{R}^n \mid \|\mathrm{z}\| = 1\}$ este mărginită şi închisă (în raport cu topologia uzuală pe $\mathbb{R}^n$), fiind deci compactă. Teorema lui Weierstrass, aplicată funcţiei $f \in \mathcal{C}^2(V)$, ne asigură că $\mathrm{y} \in \mathcal{M} \overset{\varphi}{\longmapsto} \left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{y})$ este o funcţie mărginită, care îşi atinge marginile $m = \inf\limits_{\mathrm{y} \in \mathcal{M}} \varphi(\mathrm{y})$ şi $M = \sup\limits_{\mathrm{y} \in \mathcal{M}} \varphi(\mathrm{y})$. În cazul i), ipoteza $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v}) > 0$, $\forall\, \mathrm{v} \in \mathbb{R}^n$, cu $\mathrm{v} \neq \mathbf{0}_{\mathbb{R}^n}$, ne asigură că $m > 0$. Cum α este continuă în x_0 şi $\alpha(\mathrm{x}_0) = 0$, există $\delta(m) > 0$ astfel încât $|\alpha(\mathrm{x})| < m$, $\forall\, \mathrm{x} \in S(\mathrm{x}_0; \delta(m))$. În consecinţă, avem

$$f(\mathrm{x}) - f(\mathrm{x}_0) = \frac{\|\mathrm{x} - \mathrm{x}_0\|^2}{2!}\left[\left(\left(d^2 f\right)(\mathrm{x}_0)\right)\left(\frac{\mathrm{x} - \mathrm{x}_0}{\|\mathrm{x} - \mathrm{x}_0\|}\right) + \alpha(\mathrm{x})\right] >$$

$$> \frac{\|\mathrm{x} - \mathrm{x}_0\|^2}{2!}\left[m + \alpha(\mathrm{x})\right] > 0, \forall\, \mathrm{x} \in \left(S(\mathrm{x}_0, \delta(m)) \cap V\right) \setminus \{\mathrm{x}_0\},$$

ceea ce înseamnă că x_0 este un punct de minim local al lui f.

În cazul i), atunci când $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v}) < 0$, $\forall\, \mathrm{v} \in \mathbb{R}^n \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$, avem $M < 0$ şi, ca mai sus, pe baza continuităţii lui α în x_0 şi a faptului că $\alpha(\mathrm{x}_0) = 0$, deducem că există $S(\mathrm{x}_0; \eta(M))$ astfel încât

$$f(\mathrm{x}) - f(\mathrm{x}_0) = \frac{\|\mathrm{x} - \mathrm{x}_0\|^2}{2!}\left[\left(\left(d^2 f\right)(\mathrm{x}_0)\right)\left(\frac{\mathrm{x} - \mathrm{x}_0}{\|\mathrm{x} - \mathrm{x}_0\|}\right) + \alpha(\mathrm{x})\right] <$$

$$< \frac{\|\mathrm{x}-\mathrm{x}_0\|^2}{2!}\left[M+\alpha(\mathrm{x})\right] > 0, \forall\, \mathrm{x} \in \left(S\left(\mathrm{x}_0, \eta(M)\right) \cap V\right) \setminus \{\mathrm{x}_0\},$$

ceea ce înseamnă că x_0 este punct de maxim local pentru f.

În cazul ii), când $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v})$ este nedefinită, existând $v' \in \mathbb{R}^n$ şi $v'' \in \mathbb{R}^n$, vectori nenuli, pentru care $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v}') < 0$ şi respectiv $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\mathrm{v}'') > 0$, putem vedea că funcţiile scalare $t \longmapsto f(\mathrm{x}_0 + t\mathrm{v}')$ şi $t \longmapsto f(\mathrm{x}_0 + t\mathrm{v}'')$ au punctul critic $t = 0$ şi derivata de ordinul al doilea în $t = 0$ strict negativă, respectiv strict pozitivă. Prin aplicarea Teoremei 11.2 , rezultă că $t = 0$ este punct de maxim pentru prima dintre funcţii şi punct de minim pentru cea de-a doua. Aşadar, punctul x_0, corespunzător lui $t = 0$, este de maxim pentru f, pe direcţia v' şi, simultan, punct de minim, pe direcţia v''. Deducem astfel că x_0 nu este punct de extrem al funcţiei f, ci punct şa. În fine, în situaţia iii), când $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)$ este o formă pătratică semi-definită pozitiv sau negativ, există $\widetilde{v} \in \mathbb{R}^n \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$, astfel încât $\left(\left(d^2 f\right)(\mathrm{x}_0)\right)(\widetilde{v}) = 0$ şi, atunci, pentru $\mathrm{x} = \mathrm{x}_0 + \widetilde{v}$, avem:

$$f(\mathrm{x}) - f(\mathrm{x}_0) = \frac{1}{2}\|\mathrm{x}-\mathrm{x}_0\|^2 \alpha(\mathrm{x}).$$

Cum semnul lui $\alpha(\mathrm{x})$ nu este cunoscut, nu putem stabili, pe baza acestei relaţii, natura punctului x_0 (de a fi sau nu punct de extrem). ◄

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 $\left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}\right)_{\substack{1\le i\le n\\ 1\le j\le n}}$, ***denumită hessiana lui*** f ***în*** x_0 şi notată cu $H_f(\mathrm{x}_0)$, este, pe baza criteriului lui Schwarz (v. Teorema 10.8), simetrică, se poate spune că toate valorile proprii ale acesteia sunt reale. Atunci, ţinând seama de Teorema 8.5 şi de Teorema de inerţie a lui Sylvester (Teorema 8.2), 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. În fine, 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 11.3, putem formula următorul rezultat.

Propoziţia 11.1 *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 x_0 este un punct de maxim local al lui f , când $H_f(\mathrm{x}_0)$ are toate valorile proprii negative şi punct de minim local pentru f, când $H_f(\mathrm{x}_0)$ are toate valorile proprii pozitive. Când $H_f(\mathrm{x}_0)$ are cel puţin două valori proprii nenule şi de semne contrare, atunci punctul critic x_0 este un punct şa pentru f . Dacă toate valorile proprii ale lui $H_f(\mathrm{x}_0)$ sunt nule, nu putem decide natura lui x_0.*

În mod asemănător, ţinând seama de Teorema 8.12, cât şi de observaţia ce imediat o succede, putem formula, în virtutea Teoremei 11.3, următoarea propoziţie:

Propoziţia 11.2 *Fie A, x_0, V şi f ca în enunţul Propoziţiei 11.1, iar $\Delta_0 = 1$, $\Delta_1 = \det[a_{11}]$, $\Delta_2 = \det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{bmatrix}, \ldots, \Delta_n = \det\left[(a_{ij})_{\substack{1\le i\le n\\ 1\le j\le n}}\right]$ minorii principali ai matricii $(a_{ij})_{\substack{1\le i\le n\\ 1\le j\le n}}$, unde $a_{ij} = \dfrac{\partial^2 f}{\partial x_i \partial x_j}(\mathrm{x}_0)$, $\forall\, i, j = \overline{1,n}$. Avem:*

i) Dacă $\Delta_j > 0$, $\forall\, j = \overline{1,n}$, 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 $(-1)^{j+1}\Delta_j < 0$, $\forall j = \overline{1,n}$, *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 $\Delta_j \geq 0$, $\forall j = \overline{1,n}$ *sau* $(-1)^{j+1}\Delta_j \leq 0$, $\forall j = \overline{1,n}$ *şi există cel puţin un rang* $i \in \{1, 2, \ldots, n\}$ *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 $(\Delta_j)_{j=\overline{1,n}}$ *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 11.2 revine la a spune că, 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) $\mathrm{x}_0 \in \mathring{A}$ şi adoptăm notaţiile $p = \dfrac{\partial^2 f}{\partial x_1^2}(\mathrm{x}_0)$, $q = \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(\mathrm{x}_0)$, $r = \dfrac{\partial^2 f}{\partial x_2^2}(\mathrm{x}_0)$, atunci:

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.

De pildă, revenind la Exemplul 11.1, relativ la metoda celor mai mici pătrate, cu φ de expresie liniară, se vede 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 ∇f nul în acel punct $\left(c_1^0, c_2^0\right)$ pentru care:

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

Rezolvând acest sistem algebric liniar, echivalent cu sistemul

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

al cărui determinant

$$\begin{vmatrix} \sum_{k=1}^{l} a_k^2 & \sum_{k=1}^{l} a_k \\ \sum_{k=1}^{l} a_k & l \end{vmatrix}$$

are valoarea $l\sum_{k=1}^{l} a_k^2 - \left(\sum_{k=1}^{l} a_k\right)^2$ − nenulă, când nu suntem în situaţia $a_1 = a_2 = \ldots = a_l$, pe baza inegalităţii lui Cauchy-Schwarz-Buniakowski − obţinem:

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

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

Cum, în acest caz, avem $p = \dfrac{\partial^2 f}{\partial c_1^2}\left(c_1^0, c_2^0\right) = 2\sum_{k=1}^{l} a_k^2$, $q = \dfrac{\partial^2 f}{\partial c_1 \partial c_2}\left(c_1^0, c_2^0\right) = 2\sum_{k=1}^{l} a_k$ 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 $pr - q^2 = 4\left[l\sum_{k=1}^{l} a_k^2 - \left(\sum_{k=1}^{l} a_k\right)^2\right] > 0$, ceea ce înseamnă că punctul $\left(c_1^0, c_2^0\right)$ este unul de minim pentru actuala funcţie f.

În continuare, deoarece pentru abordarea teoretică a unei probleme de extrem cu restricţii (legături) avem nevoie de elemente relative la noţiunile de funcţie implicită, inversabilitatea unei funcţii de mai multe variabile reale şi dependenţa funcţională a unui set de funcţii, prezentăm, pe scurt, astfel de noţiuni şi, fără demonstraţii, rezultate de bază care le privesc.

Funcţii implicite

Definiţia 11.5 *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 $(x_1, x_2, \ldots, x_n)$ cu x, 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 11.4 ***(a 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 $U_0 \subseteq U$ - vecinătate deschisă a punctului a şi o funcţie unică $f : U_0 \to V_0$ astfel încât $f(\mathrm{a}) = b$ şi $F(\mathrm{x}, f(\mathrm{x})) = 0$, $\forall\, \mathrm{x} \in U_0$.

ii) $f \in \mathcal{C}^1(U_0)$ şi

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

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

Definiţia 11.5 şi Teorema 11.4 pot fi extinse la sisteme de ecuaţii şi de funcţii implicit-definite. Astfel, avem:

Definiţia 11.6 *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 $f_k : A \to \mathbb{R}$, $\forall\, k = \overline{1, m}$, satisfăcând relaţiile:

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

Teorema 11.5 *(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

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

2) *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*

3) *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 o vecinătate deschisă* U_0 *a punctului* x_0, *cu* $U_0 \subseteq U \subseteq \mathbb{R}^n$, *o vecinătate* V_0, *deschisă, a punctului* y_0, *cu* $V_0 \subseteq V \subseteq \mathbb{R}^m$ *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* $f_j(\mathrm{x}_0) = y_j^0$, $\forall\, j = \overline{1,m}$ *şi* $F_j(\mathrm{x}, f(\mathrm{x})) = 0$, $\forall\, j = \overline{1,m}$, $\forall\, \mathrm{x} = (x_1, x_2, \ldots, x_n) \in U_0$;

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

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

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 variable 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 $det\left[(a_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq n}}\right]\neq 0$. Altfel spus, funcţia inversă $f:\mathbb{R}^n\to\mathbb{R}^n$, cu expresia

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

adică, matriceal, $f(\mathrm{x})=A\mathrm{x}$, unde $\mathrm{x}=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n$ şi $A=(a_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq n}}$ – este bijectivă (şi ca atare, ***inversabilă în vecinătatea oricărui punct*** $\mathrm{x}_0=(x_1^0,x_2^0,\ldots,x_n^0)\in\mathbb{R}^n$; ***adică inversabilă global***) dacă şi numai dacă $\dfrac{D(f_1,f_2,\ldots,f_n)}{D(x_1,x_2,\ldots,x_n)}(\mathrm{x})=\det A\neq 0$.

Extinderea unui asemenea rezultat la cazul sistemelor neliniare de forma

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

se poate face, de regulă local, în virtutea următoarei teoreme:

Teorema 11.6 ***(de inversabilitate locală)***
Fie $A\subseteq\mathbb{R}^n$ o mulţime nevidă şi deschisă, iar $f:A\to\mathbb{R}^n$ o funcţie de clasă $\mathcal{C}^1$ pe A. De asemenea, fie $\mathrm{x}_0\in A$. Dacă $(df)(\mathrm{x}_0)$ este o bijecţie (liniară) de la $\mathbb{R}^n$ la $\mathbb{R}^n$, adică dacă $det(J_f(\mathrm{x}_0))\neq 0$, atunci există o vecinătate deschisă U a lui x_0, $U\subseteq A$ şi o vecinătate deschisă V a lui $f(\mathrm{x}_0)$, 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 11.6 poate fi reformulată după cum urmează: "*O funcţie $f:A\subseteq\mathbb{R}^n\to R^n$, $f\in C^1(A)$, 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 11.6,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 11.7 *Fie $A\subseteq\mathbb{R}^n$ o mulţime nevidă, iar $f_1,f_2,\ldots,f_m$, cu $m\leq n$, funcţii definite pe A şi cu valori în $\mathbb{R}$. Se spune despre o funcţie $f_0:A\to\mathbb{R}$ că* ***depinde (este dependentă) funcţional*** *(sau că este* ***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:*

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

Altminteri, f_0 *se numeşte* ***funcţional-independentă de*** $f_1, f_2, \ldots, f_m$, *pe* $\widetilde{A}$.

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

Propoziţia 11.3 *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* $f_0 : A \to \mathbb{R}$, *de clasă* $\mathcal{C}^1$ *pe* 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, în ipotezele Propoziţiei 11.3, f_0 *depinde funcţional de* $f_1, f_2, \ldots, f_m$ *dacă* df_0 *se poate exprima, pe* A, *ca o combinaţie liniară de* $df_1, df_2, \ldots, df_m$.

Teorema 11.7 ***(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 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

În realitatea obiectivă, pe lângă situaţiile în care se cer rezolvate probleme de extrem libere, fără restricţii, precum cele din Exemplele 11.1 şi 11.2, în parte, se întâlnesc şi cazuri de probleme în care extremele unei funcţii se caută în condiţiile satisfacerii, de către argumentele funcţiei respective, a unor relaţii de legătură între ele. În asemenea cazuri, extremele respective poartă denumirea de ***extreme condiţionate*** sau, echivalent, ***extreme cu legături***, legături ce se exprimă, adeseori, prin relaţii de egalitate.

Analizăm aici următoarea situaţie, în care D este o mulţime deschisă, în raport cu topologia uzuală, din $\mathbb{R}^n \times \mathbb{R}^m$ $(n, m \in \mathbb{N}^*)$, iar $f : D \to \mathbb{R}$ şi $g : F \to \mathbb{R}^m$ sunt funcţii de clasă $\mathcal{C}^1$ pe D, astfel încât se caută punctele de extrem ale funcţiei f, supuse la condiţia suplimentară $g(\mathrm{x}, \mathrm{y}) = 0$, 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 mulţimea $\{(\mathrm{x}, \mathrm{y}) \in D \mid g(\mathrm{x}, \mathrm{y}) = \mathbf{0}_{\mathbb{R}^n}\}$ şi cu $g_1, g_2, \ldots, g_m$ componentele funcţiei g, se vede că problema de mai sus cere, de fapt, determinarea punctelor de extrem ale restricţiei lui f la A, adică acele puncte de extrem ale lui f care satisfac, simultan, relaţiile $g_1(\mathrm{x}, \mathrm{y}) = 0, g_2(\mathrm{x}, \mathrm{y}) = 0, \ldots, g_m(\mathrm{x}, \mathrm{y}) = 0$. Cum $g_1, g_2, \ldots, g_m$ sunt continue, rezultă că g este continuă şi deci A este o mulţime închisă, în raport cu topologia amintită. 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. În acest scop, este de observat că metoda de evidenţiere a acestor puncte, folosită în cazul extremelor libere, nu este de utilizat, deoarece punctele lui A vor aparţine interiorului acestei mulţimi dacă şi numai dacă $g_1 \equiv g_2 \equiv \ldots \equiv g_m \equiv 0$, 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 11.4 sau Teorema 11.5) pentru a soluţiona mai întâi sistemul $g_k(\mathrm{x}, \mathrm{y}) = 0$, $\forall\, k = \overline{1, m}$, 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 obiectiv $f(\mathrm{x}, \varphi(\mathrm{x}))$, 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 materialului de faţă. 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 $f(\mathrm{x}, \varphi(\mathrm{x}))$.

Iată de ce, în realitate, se foloseşte aşa-numita ***metodă a multiplicatorilor lui Lagrange***, descrisă în teorema care urmează după definiţia dată în continuare.

Definiţia 11.8 *a) Se spune că* $(\mathrm{x}_0, \mathrm{y}_0) \in A$ *este un* ***punct de extrem local al funcţiei*** $f : D \subseteq \mathbb{R}^{n+m} \to \mathbb{R}$, ***cu restricţiile*** $g_1(\mathrm{x}, \mathrm{y}) = 0, g_2(\mathrm{x}, \mathrm{y}) = 0, \ldots, g_m(\mathrm{x}, \mathrm{y}) = 0$, *unde* $g_k : D \to \mathbb{R}$, $\forall\, k = \overline{1, m}$, *(sau că* $(\mathrm{x}_0, \mathrm{y}_0)$ *este* ***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* $\forall\, (\mathrm{x}, \mathrm{y}) \in V \cap A$.

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 exprimat, $(\mathrm{x}_0, \mathrm{y}_0)$ *este un punct de extrem (minim sau maxim) 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.

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

Teorema 11.8 ***(de existenţă a multiplicatorilor Lagrange)***

Fie D *o mulţime nevidă şi deschisă din* $\mathbb{R}^n \times \mathbb{R}^m$, $f : D \to \mathbb{R}$ *şi* $g : D \to \mathbb{R}^m$, *adică* $g = (g_1, g_2, \ldots, g_m)$, *cu* $g_k : D \to \mathbb{R}$, $\forall\, k = \overline{1, m}$, *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* $g_k(\mathrm{x}_0, \mathrm{y}_0) = 0$, $\forall\, k = \overline{1, m}$.

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:*

$$(*) \quad \begin{cases} \dfrac{\partial L}{\partial x_k}(\mathrm{x}, \mathrm{y}) = 0, & \forall\, k = \overline{1, n} \\[2ex] \dfrac{\partial L}{\partial y_j}(\mathrm{x}, \mathrm{y}) = 0, & \forall\, j = \overline{1, m} \\[2ex] g_i(\mathrm{x}, \mathrm{y}) = 0, & \forall\, i = \overline{1, m}. \end{cases}$$

Precizări:

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

2) Punctele de extrem condiţionat (în raport cu restricţiile $g_k(\mathrm{x}, \mathrm{y}) = 0$, $\forall\, k = \overline{1, m}$) pentru funcţia f se găsesc printre punctele critice, condiţionate, ale lui f , adică printre punctele critice ale funcţiei asociate $L = f + \sum_{k=1}^{m} \lambda_k g_k$.

3) Dacă, în expresia $f(\mathrm{x}, \mathrm{y}) + \sum_{k=1}^{m} \lambda_k g_k(\mathrm{x}, \mathrm{y})$, 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, $\forall\, k = \overline{1, m}$.

4) Dacă f şi g_k sunt funcţii de clasă $\mathcal{C}^2$ pe D, $\forall\, k = \overline{1, n}$, 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 11.8, punctul $(\mathrm{x}_0, \mathrm{y}_0)$ este unul critic pentru funcţia $L(\mathrm{x}_0, \mathrm{y}_0; \lambda_0)$, 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 $g_k(\mathrm{x}, \mathrm{y}) = 0$, $\forall\, k = \overline{1, m}$, 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 11.3, este necesară studierea formei pătratice $\left(d^2 L(\mathrm{x}, \mathrm{y}; \lambda_0)\right)(\mathrm{x}_0, \mathrm{y}_0)$. Altfel spus, trebuie studiată expresia $\sum_{i,j=1}^{n+m} \frac{\partial^2 L}{\partial \widetilde{x}_i \partial \widetilde{x}_j}(\mathrm{x}_0, \mathrm{y}_0)\, d\widetilde{x}_i d\widetilde{x}_j$, 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 $g_k(\widetilde{\mathrm{x}}) = 0$, $\forall\, k = \overline{1, m}$, anume: $\sum_{k=1}^{n+m} \frac{\partial g_i}{\partial \widetilde{x}_j}(\widetilde{\mathrm{x}}_0)\, d\widetilde{x}_k = 0$. Cum $\frac{D(g_1, g_2, \ldots, g_m)}{D(\widetilde{x}_{n+1}, \widetilde{x}_{n+2}, \ldots, \widetilde{x}_{n+m})}(\mathrm{x}_0, \mathrm{y}_0) \neq 0$, de aici putem găsi $d\widetilde{x}_{n+1}, d\widetilde{x}_{n+2}, \ldots, d\widetilde{x}_{n+m}$ în funcţie de $d\widetilde{x}_1, d\widetilde{x}_2, \ldots, d\widetilde{x}_n$, prin expresii liniare. Aceste expresii, introduse în $\sum_{i,j=1}^{n+m} \frac{\partial^2 L}{\partial \widetilde{x}_i \partial \widetilde{x}_j}(\mathrm{x}_0, \mathrm{y}_0)\, d\widetilde{x}_i d\widetilde{x}_j$ conduc la relaţia $d^2 L(\widetilde{\mathrm{x}}_0) = \sum_{i,j=1}^{n} a_{ij}^0\, dx_i dx_j$, reprezentând o formă pătratică pe $\mathbb{R}^n$, în care a_{ij}^0 depind liniar de $\frac{\partial^2 g_k}{\partial \widetilde{x}_i \partial \widetilde{x}_j}(\mathrm{x}_0, \mathrm{y}_0)$, $\frac{\partial g_i}{\partial x_j}(\mathrm{x}_0, \mathrm{y}_0)$ şi $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathrm{x}_0, \mathrm{y}_0)$. Astfel, pe baza Teoremei 11.3, deducem că, dacă $d^2 L(\widetilde{\mathrm{x}}_0)$ 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 $d^2 L(\widetilde{\mathrm{x}}_0)$ 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 $d^2 L(\widetilde{\mathrm{x}}_0)$ 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.

Aceasta este *metoda multiplicatorilor lui Lagrange* pentru tratarea problemelor de extrem condiţionat.

Demonstraţie: *(pentru Teorema 11.8)*
Prin aplicarea teoremei funcţiilor implicite pentru sisteme de ecuaţii, adică prin aplicarea Teoremei 11.5 asupra sistemului de legături $g_k(\mathrm{x}, \mathrm{y}) = 0$, $\forall\, k = \overline{1, m}$, se poate spune că există U, o vecinătate a

lui x_0, şi o funcţie $\varphi : U \to \mathbb{R}^n$, de clasă $\mathcal{C}^1$ pe U (adică există $\varphi_k : U \to \mathbb{R}$, $\forall\, k = \overline{1,m}$, $\varphi_k \in \mathcal{C}^1(U)$) aşa încât $\varphi(\mathrm{x}_0) = \mathrm{y}_0$ şi $g_k(\mathrm{x}, \varphi(\mathrm{x})) = 0$, $\forall\, k = \overline{1,m}$, $\forall\, \mathrm{x} \in U$. De aici, rezultă că avem (în virtutea regulii "lanţului"), prin derivare:

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

Considerând funcţia $F(\mathrm{x}) = f(\mathrm{x}, \varphi(\mathrm{x}))$, $\forall\, \mathrm{x} \in U$, putem zice că, deoarece $(\mathrm{x}_0, \mathrm{y}_0)$ este un punct de extrem local, cu restricţiile $g_k(\mathrm{x}, \mathrm{y}) = 0$, $\forall\, k = \overline{1,m}$, pentru f, iar $(\mathrm{x}, \varphi(\mathrm{x})) \in A$, $\forall\, \mathrm{x} \in U$, punctul x_0 este unul de extrem local, necondiţionat, pentru F pe U. Atunci, potrivit Teoremei 11.1 (Fermat), avem $\dfrac{\partial F}{\partial x_k}(\mathrm{x}_0) = 0$, $\forall\, k = \overline{1,n}$. Cu alte cuvinte, avem:

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

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

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

are o soluţie unică. Fie aceasta notată cu $(\lambda_1^0, \lambda_2^0, \ldots, \lambda_m^0)$. În raport cu ea, funcţia $L = f + \displaystyle\sum_{k=1}^{m} \lambda_k^0 g_k$ satisface relaţiile din sistemul $(*)$.

Într-adevăr, cum $(\mathrm{x}_0, \mathrm{y}_0) \in A$, avem $g_k(\mathrm{x}_0, \mathrm{y}_0) = 0$, $\forall\, k = \overline{1,m}$.

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 $(***)$. De aici, folosind (!), se deduce că $\dfrac{\partial L}{\partial x_i}(\mathrm{x}_0, \mathrm{y}_0) = 0$, $\forall\, i = \overline{1,n}$

În fine, derivând L în raport cu y_j şi ţinând, iarăşi, cont de faptul că $(\lambda_1^0, \lambda_2^0, \ldots, \lambda_m^0)$ este soluţia sistemului $(***)$, obţinem:

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

◀

Aplicând acum metoda multiplicatorilor lui Lagrange în cazul problemei entropiei lui Shannon, adică în cazul problemei de extrem pentru funcţia

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

cu condiţia $\sum_{k=1}^{n} p_k = 1$, ne dăm seama că $m = 1$, $g(p_1, p_2, \ldots, p_n) = \sum_{k=1}^{n} p_k - 1$ şi Lagrangeanul în cauza prezentă 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:

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

Din primele n ecuaţii, reiese că $p_1 = p_2 = \ldots = p_n = \psi(\lambda_1)$, iar dacă se ţine seama şi de ultima ecuaţie, rezultă soluţia:

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

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:

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

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

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

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

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