

Cursul 11

Aplicații ale diferențiabilității funcțiilor.

(Extreme fără restricții. Funcții implicite. Inversabilitatea funcțiilor de mai multe variabile reale. Dependență/independență funcțională. Extreme condiționate.)

Una dintre aplicațiile de certă importanță practică a diferențiabilității funcțiilor reale (de una sau mai multe variabile și cu valori scalare) este aceea relativă la stabilirea elementelor de extrem pentru funcțiile implicate în anumite probleme de optimizare, adică în probleme care vizează minimizarea sau maximizarea unei așa-numite funcționale de cost, în absența sau în prezența unor condiții (restricții) precizate. După exemplificarea unor astfel de probleme, prezentăm aici chestiunile teoretice necesare abordării și tratării cazului ce privește extremele necondiționate, iar apoi, pe baza expunerii unor concepte și rezultate din domeniile funcțiilor implicit-definite, inversabilității funcțiilor reale vectorial-vectoriale și dependenței funcționale a unui set de funcții, 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 față de determinări*

Admitem că, în urma unor experimente asupra unei anumite mărimi fizice, s-au obținut valorile b_1, b_2, \dots, b_p , corespunzătoare valorilor (de "intrare") a_1, a_2, \dots, 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 (forme) expresiei (graficului) funcției φ care, necunoscută inițial, 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, exponențial, trigonometric etc.), ale cărei caracteristici (parametri) c_1, c_2, \dots, c_n (din \mathbb{R}), nesupuse vreunei condiții restrictive, se cer a fi identificate. În acest scop, folosind așa-numita **metoda a celor mai mici pătrate**, considerăm problema minimizării expresiei

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

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

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

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

Precizăm (în context) că dacă imaginea grafică a mulțimii $\{(a_k, b_k) \mid k = \overline{1, p}\}$ ne sugerează faptul că φ ar fi să aibă o expresie liniară, atunci putem lua $n = 2$ și $\varphi(x) = c_1x + c_2$. Prin aceasta, metoda celor mai mici pătrate va consta în determinarea parametrilor c_1 și c_2 așa 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 producție economică*

Într-o teorie economică, spațiul \mathbb{R}^n se interpretează ca fiind spațiul complexelor de bunuri de consum, în care fiecare "bun" (produs) este caracterizat de un anumit indice $i \in \{1, 2, \dots, n\}$, iar un "complex de bunuri" este un vector $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, unde componenta x_i înseamnă cantitatea în care se găsește "bunul" i . "Unitatea" bunului i este $e_i = (0, 0, \dots, 0, 1, 0, \dots, 0)$. Dreapta reală este interpretată ca mulțimea "valorilor", exprimată în "unități de cont". Într-un astfel de context, un "sistem de prețuri" este o funcție ce asociază fiecărui "complex de bunuri" o anumită valoare. Se consideră, în mod firesc, că un "sistem de prețuri" este un element din dualul lui \mathbb{R}^n , adică o aplicație liniară $p = (p_1, p_2, \dots, p_n)$, unde p_i este prețul unitar al "bunului" i . Astfel, pentru complexul de bunuri $x \in \mathbb{R}^n$, valoarea sa în raport cu sistemul de prețuri p este dată de $\langle p, x \rangle_e$, anume $\sum_{i=1}^n p_i x_i$.

Considerând mulțimea \mathbb{R}_+^n , a vectorilor $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ cu componente nenegative, comportamentul unui consumator este apreciat în teoria economică printr-o funcție de utilitate $u : \mathbb{R}_+^n \rightarrow \mathbb{R}$, funcție ce poate determina o așa-numită "relație de preferință" " \prec " pe mulțimea \mathbb{R}_+^n , definită prin: $y \prec x \Leftrightarrow u(x) \leq u(y)$.

Dacă o firmă (întreprindere) produce un anumit "complex de bunuri", se poate pune problema realizării respectivei producții astfel încât cheltuielile de producție să fie minime sau/și profitul de producție să fie maxim. Prin studii economice adecvate, se utilizează, în acest sens, o funcție de cost corespunzătoare contextului și o funcție de profit convenabil stabilită.

Când optimul funcției obiectiv (de cost sau /și profit) se cere a fi găsit, în situația în care mulțimea complexelor de bunuri este \mathbb{R}^n , spunem că avem de-a face cu o problemă de extrem fără restricții. În caz contrar, când mulțimea producțiilor nete ale unei întreprinderi (firme), notată cu K , este o submulțime proprie a lui \mathbb{R}_+^n ce conține $0_{\mathbb{R}^n}$, problema optimizării funcției de utilitate u , pe K , este una de extrem condiționat. Problema în cauză este de "programare liniară" ori de câte ori atât funcția de optimizat, cât și relațiile ce definesc mulțimea K sunt liniare. Dacă funcția obiectiv este pătratică sau convexă, iar K este convexă, atunci problema de optimizare considerată se numește problemă de optimizare pătratică și respectiv convexă, cu restricții.

Exemplul 11.3 *Problema entropiei informaționale maxime*

Introdusă, ca noțiune matematică, de Claude E. Shannon (1947), *entropia* reprezintă o funcție ce corespunde cantității de informație livrată (oferită) de o anumită sursă, prin intermediul unui anumit limbaj, semnal electric sau fișier (informatic) de date. Funcția respectivă, notată cu H , este definită pe mulțimea variabilelor aleatoare

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

și are expresia

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

unde p_k este probabilitatea ($p_k \in (0, 1)$) cu care se transmit (se percep) k informații transmise de o sursă (sau mai multe surse ce acționează concomitent), iar $\sum_{k=1}^n p_k = 1$.

O astfel de funcție (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) și

p_k egală cu raportul dintre frecvența globală a silabei aflate pe poziția k (în clasamentul obținut prin ordonarea descrescătoare a silabelor după frecvența lor în limba română) și numărul total de apariții 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 componența silabelor) și p_k este probabilitatea de apariție a tipului k (în ordinea frecvenței de apariție), se găsește $H = 230$.

Tot o funcție de entropie de tip Shannon este și 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 suprafața (aria) proiectată a feței i a unui obiect ce se dorește observat dintr-un anumit punct P din spațiu, 4π este măsura unghiului solid (sferic) de rază 1, iar $\frac{A_i}{4\pi}$ reprezintă gradul de vizibilitate, pentru fața de rang i , din punctul P . O asemenea funcție se folosește în scopul alegerii celei mai bune poziții a punctului P pentru observarea obiectului cu fețele $1, 2, \dots, N$.

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

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

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

Extreme libere (fără legături)

Fie $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ o funcțională reală pe $A \neq \emptyset$ și $x_0 \in A$.

Definiția 11.1 a) Punctul x_0 se numește **punct de extrem local** (similar spus, **relativ**) al funcției f dacă există o vecinătate $V \in \mathcal{V}(x_0)$ așa încât $(f(x) - f(x_0))$ are semn constant sau este nulă pe $V \cap A$. Atunci $f(x_0)$ se numește **valoare extremă locală a lui f** (pe $V \cap A$).

b) Punctul x_0 se numește **punct de maxim** (respectiv **punct de minim**) **local** (sau **relativ**) al funcției f când are loc relația:

$$f(x) - f(x_0) \leq 0 \text{ (respectiv } f(x) - f(x_0) \geq 0), \forall x \in V \cap A.$$

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

c) Dacă $f(x) - f(x_0) \leq 0$ (respectiv $f(x) - f(x_0) \geq 0$), $\forall x \in A$, atunci x_0 se numește **punct de maxim** (respectiv **minim**) **absolut** (sau **global**) pentru f . În acest caz, "**extrema**" $f(x_0)$ este **valoarea maximă** (respectiv **minimă**) **absolută (globală)** a lui f .

Observație: Orice punct de tipul x_0 , de extrem (maxim sau minim) absolut pentru o funcție f , este întotdeauna punct de extrem relativ, pentru f . Nu și reciproc.

Definiția 11.2 a) Problema determinării punctelor și valorilor de extrem, locale sau globale, ale unei funcții (funcționale) $f : \mathbb{R}^n \rightarrow \mathbb{R}$, în absența oricărei condiții restrictive asupra argumentului lui f , se numește **problemă de extrem liber** (**necondiționat** sau **fără legături**).

b) Problema determinării elementelor de extrem ale unei funcții $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, în condițiile în care se cere punctului de extrem să aparțină unei anumite mulțimi (de restricții) $B \subseteq \mathbb{R}^n$ ($B \neq \emptyset$), se numește **problemă de extrem (maxim sau minim) cu legături** (sau **problemă de extrem condiționat**).

Teorema 11.1 (Fermat)

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

Demonstrație: Dacă $x_0 \in \mathring{A}$ și există $\frac{\partial f}{\partial x_k}(x_0)$, atunci, în conformitate cu Definiția 10.5 d), avem:

$\frac{\partial f}{\partial x_k}(x_0) = \lim_{t \rightarrow 0} \frac{f(x_0 + t\mathbf{e}_k) - f(x_0)}{t}$. În cazul în care x_0 este un punct de maxim al lui f , local sau global, înregistrăm: $f(x_0 + t\mathbf{e}_k) - f(x_0) \leq 0, \forall t \in V_0 \in \mathcal{V}(0)$, cu $V_0 = (a, b)$, $a < 0 < b$. Astfel, $\lim_{\substack{t \rightarrow 0 \\ t < 0}} \frac{f(x_0 + t\mathbf{e}_k) - f(x_0)}{t} \geq 0$ și $\lim_{\substack{t \rightarrow 0 \\ t > 0}} \frac{f(x_0 + t\mathbf{e}_k) - f(x_0)}{t} \leq 0$. Deci $\frac{\partial f}{\partial x_k}(x_0) = 0$. Când x_0 este punct de minim al lui f , avem $\lim_{\substack{t \rightarrow 0 \\ t < 0}} \frac{f(x_0 + t\mathbf{e}_k) - f(x_0)}{t} \leq 0$ și $\lim_{\substack{t \rightarrow 0 \\ t > 0}} \frac{f(x_0 + t\mathbf{e}_k) - f(x_0)}{t} \geq 0$. Prin urmare, rezultă și atunci că $\frac{\partial f}{\partial x_k}(x_0) = 0$. ◀

Observație: Dacă $x_0 \in \mathring{A}$ este un punct de extrem al unei funcții $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, cu gradient în x_0 , atunci, potrivit Teoremei 11.1, gradientul lui f în x_0 , adică $(\nabla f)(x_0)$ este vectorul nul $\mathbf{0}_{\mathbb{R}^n}$.

Definiția 11.3 Fie $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ și $x_0 \in \mathring{A}$, astfel încât f este diferențiabilă Fréchet (de ordinul întâi) în x_0 . Punctul x_0 se numește **punct critic** (sau **punct staționar**) al funcției f dacă $(df)(x_0) = \mathbf{0}_{\mathcal{L}(\mathbb{R}^n; \mathbb{R})}$, adică $(\nabla f)(x_0) = \mathbf{0}_{\mathbb{R}^n}$.

Observație: Întrucât $((df)(x_0))(v) = \langle (\nabla f)(x_0), v \rangle, \forall v \in \mathbb{R}^n$, Teorema 11.1 afirmă că orice punct de extrem (local) care aparține interiorului mulțimii A (de definiție a funcționalei $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$) și în care f este diferențiabilă Fréchet constituie un punct critic al lui f . Reciproca nu este adevărată, după cum se poate vedea în cazul funcției $f : \mathbb{R}^2 \rightarrow \mathbb{R}, f(x_1, x_2) = x_1^3 + x_2^3 - 3x_1x_2$, cu punctul critic $(1, 1)$ (deoarece $\frac{\partial f}{\partial x_1}(1, 1) = 3x_1^2 - 3x_2|_{x_1=1, x_2=1} = 0$ și $\frac{\partial f}{\partial x_2}(1, 1) = 3x_2^2 - 3x_1|_{x_1=1, x_2=1} = 0$). Se constată că diferența $f(x_1, x_2) - f(1, 1) = x_1^3 + x_2^3 - 3x_1x_2 + 1$ este, pentru $x_2 = 1$, egală cu $(x_1 - 1)(x_1 - 2)$ ceea ce înseamnă că $f(1 - a, 1) - f(1, 1) = (-a)(-1 - a) = a(1 + 1) \geq 0, \forall a \geq 0$ și $f(1 + b, 1) - f(1, 1) = b(b - 1) \leq 0, \forall b \in [0, 1]$. Deci, în punctul $(1, 1)$, funcția f din acest caz nu are nici minim și nici maxim local, nesatisfăcând Definiția 11.1.

Definiția 11.4 Un punct critic al unei funcții (diferențiabile) $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ care nu este punct de extrem al lui f se numește **punct șa** pentru funcția f .

Pentru funcțiile diferențiabile de cel puțin ordinul al doilea într-un punct critic există criterii (condiții suficiente) de discernere a punctelor de extrem sau a punctelor șa, adică criterii de identificare a punctelor de extrem sau șa printre punctele critice ale respectivelor funcții.

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

Teorema 11.2 Fie $A \subseteq \mathbb{R}$ un interval, $x_0 \in \mathring{A}$ și $f : A \rightarrow \mathbb{R}$ o funcție de n ($n \geq 2$) ori derivabilă în x_0 . Dacă $f'(x_0) = f''(x_0) = \dots = f^{(n-1)}(x_0) = 0$ și $f^{(n)}(x_0) \neq 0$, iar n este par, atunci x_0 este punct

de extrem local pentru f și 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 .

Demonstrație: Considerând funcțiile $R_f(\cdot; x_0) : A \rightarrow \mathbb{R}$ și $g_0 : A \rightarrow \mathbb{R}$, definite prin $R_f(x; x_0) = f(x) - f(x_0) - \frac{x - x_0}{1!} f'(x_0) - \dots - \frac{(x - x_0)^n}{n!} f^{(n)}(x_0)$ și $g_0(x) = (x - x_0)^n$, $\forall x \in A$, vedem că $R_f(\cdot; x_0)$ și g_0 sunt derivabile de n ori în $x_0 \in \mathring{A}$ și $R_f^{(k)}(x_0; x_0) = g_0^{(k)}(x_0) = 0$, $\forall k = \overline{1, n-1}$, iar $R_f^{(n)}(x_0; x_0) = 0$ și $g_0^{(n)}(x_0) = n!$. Atunci: $\lim_{x \rightarrow x_0} \frac{R_f(x; x_0)}{(x - x_0)^n} = \lim_{x \rightarrow x_0} \frac{R_f'(x; x_0)}{n(x - x_0)^{n-1}} = \dots = \lim_{x \rightarrow x_0} \frac{R_f^{(n)}(x; x_0)}{g_0^{(n)}(x)} = \frac{0}{n!} = 0$. În consecință, avem

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

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

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

putem scrie

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

unde α este continuă și 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, $f'(x_0) = f''(x_0) = \dots = f^{(n-1)}(x_0) = 0$ și $f^{(n)}(x_0) \neq 0$, rezultă că avem:

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

În același timp, deoarece $\lim_{x \rightarrow x_0} \alpha(x) = 0$, avem:

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

Grație continuității lui α în x_0 , există $V \in \mathcal{V}(x_0)$ așa încât, pe baza relației de imediat mai sus, semnul expresiei funcției $\alpha(x) + f^{(n)}(x_0)$ este constant și anume egal cu semnul numărului $f^{(n)}(x_0)$. Atunci $\text{sign}(f(x) - f(x_0)) = \text{sign} \left([f^{(n)}(x_0)] \cdot \left[\frac{(x - x_0)^n}{n!} \right] \right)$, $\forall x \in V \cap A$ și 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 $\text{sign}(f(x) - f(x_0)) = \text{sign}(f^{(n)}(x_0))$. În consecință, când $f^{(n)}(x_0) > 0$, obținem $f(x) > f(x_0)$, $\forall x \in (V \cap A) \setminus \{x_0\}$, 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 $V \cap A$ și, la fel, se va petrece atunci cu semnul diferenței $f(x) - f(x_0)$, ceea ce ne spune că x_0 este un punct șa pentru f și nu un punct de extrem. ◀

Teorema 11.3 Fie $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ o funcție ce are pe $x_0 \in \mathring{A}$ ca punct critic. Dacă f are derivate parțiale de ordinul al doilea continue într-o vecinătate a lui x_0 , atunci:

- i) când $((d^2 f)(x_0))(v) \equiv \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) v_i v_j > 0, \forall v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$, punctul x_0 este unul de minim pentru funcția f , iar când $((d^2 f)(x_0))(v) < 0, \forall v \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$, punctul x_0 este unul de maxim pentru f ;
- ii) când $(d^2 f)(x_0)$ este o formă pătratică nedefinită (adică $\exists v' \neq 0_{\mathbb{R}^n}$ și $v'' \neq 0_{\mathbb{R}^n}$ din \mathbb{R}^n astfel încât $((d^2 f)(x_0))(v') < 0$ și $((d^2 f)(x_0))(v'') > 0$, x_0 este punct șa pentru f (nefiind punct de extrem);
- iii) când $(d^2 f)(x_0)$ este o formă pătratică semidefinită pozitiv (sau negativ),adică $((d^2 f)(x_0))(v) \geq 0, \forall v \in \mathbb{R}^n$ și există $v' \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$, astfel încât $((d^2 f)(x_0))(v') = 0$ (respectiv $((d^2 f)(x_0))(v) \leq 0, \forall v \in \mathbb{R}^n$ și există $v'' \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$, astfel încât $((d^2 f)(x_0))(v'') = 0$), nu putem stabili natura punctului staționar x_0 cu ajutorul diferențialei $(d^2 f)(x_0)$.

Demonstrație: Deoarece f are derivate parțiale de ordinul al doilea continue pe o vecinătate V a lui x_0 (vecinătate pe care, fără a restrânge generalitatea prezentului raționament, o putem considera de tip sferic, adică, mai precis, $V = S(x_0; r) \subseteq A$), f este diferențiabilă (Fréchet) de două ori pe V și deci există $\alpha : V \rightarrow \mathbb{R}$, așa încât $\lim_{x \rightarrow x_0} \alpha(x) = \alpha(x_0) = 0$ și

$$f(x) = f(x_0) + \frac{1}{1!} (df(x_0))(x - x_0) + \frac{1}{2!} ((d^2 f)(x_0))(x - x_0) + \frac{\alpha(x)}{2!} \|x - x_0\|^2, \forall x \in V,$$

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

Cum x_0 este punct staționar pentru f , avem $(df(x_0))(x - x_0) = 0, \forall x \in V$. Atunci:

$$f(x) - f(x_0) = \frac{\|x - x_0\|^2}{2!} \left[((d^2 f)(x_0)) \left(\frac{x - x_0}{\|x - x_0\|} \right) + \alpha(x) \right], \forall x \in V \setminus \{x_0\}.$$

Totodată, mulțimea $\mathcal{M} = \left\{ y = \frac{x - x_0}{\|x - x_0\|} \mid x \in V \setminus \{x_0\} \right\} \subseteq \{z \in \mathbb{R}^n \mid \|z\| = 1\}$ este mărginită și închisă (în raport cu topologia uzuală pe \mathbb{R}^n), fiind deci compactă. Teorema lui Weierstrass, aplicată funcției $f \in \mathcal{C}^2(V)$, ne asigură că $y \in \mathcal{M} \mapsto ((d^2 f)(x_0))(y)$ este o funcție mărginită, care își atinge marginile $m = \inf_{y \in \mathcal{M}} \varphi(y)$ și $M = \sup_{y \in \mathcal{M}} \varphi(y)$. În cazul i), ipoteza $((d^2 f)(x_0))(v) > 0, \forall v \in \mathbb{R}^n$, cu $v \neq 0_{\mathbb{R}^n}$, ne asigură că $m > 0$. Cum α este continuă în x_0 și $\alpha(x_0) = 0$, există $\delta(m) > 0$ astfel încât $|\alpha(x)| < m, \forall x \in S(x_0; \delta(m))$. În consecință, avem

$$\begin{aligned} f(x) - f(x_0) &= \frac{\|x - x_0\|^2}{2!} \left[((d^2 f)(x_0)) \left(\frac{x - x_0}{\|x - x_0\|} \right) + \alpha(x) \right] > \\ &> \frac{\|x - x_0\|^2}{2!} [m + \alpha(x)] > 0, \forall x \in (S(x_0, \delta(m)) \cap V) \setminus \{x_0\}, \end{aligned}$$

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

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

$$f(x) - f(x_0) = \frac{\|x - x_0\|^2}{2!} \left[((d^2 f)(x_0)) \left(\frac{x - x_0}{\|x - x_0\|} \right) + \alpha(x) \right] <$$

$$< \frac{\|x - x_0\|^2}{2!} [M + \alpha(x)] > 0, \forall x \in (S(x_0, \eta(M)) \cap V) \setminus \{x_0\},$$

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

În cazul ii), când $((d^2 f)(x_0))(v)$ este nedefinită, existând $v' \in \mathbb{R}^n$ și $v'' \in \mathbb{R}^n$, vectori nenuli, pentru care $((d^2 f)(x_0))(v') < 0$ și respectiv $((d^2 f)(x_0))(v'') > 0$, putem vedea că funcțiile scalare $t \mapsto f(x_0 + tv')$ și $t \mapsto f(x_0 + tv'')$ au punctul critic $t = 0$ și 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 funcții și punct de minim pentru cea de-a doua. Așadar, punctul x_0 , corespunzător lui $t = 0$, este de maxim pentru f , pe direcția v' și, simultan, punct de minim, pe direcția v'' . Deducem astfel că x_0 nu este punct de extrem al funcției f , ci punct ș.a. În fine, în situația iii), când $((d^2 f)(x_0))$ este o formă pătratică semi-definită pozitiv sau negativ, există $\tilde{v} \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$, astfel încât $((d^2 f)(x_0))(\tilde{v}) = 0$ și, atunci, pentru $x = x_0 + \tilde{v}$, avem:

$$f(x) - f(x_0) = \frac{1}{2} \|x - x_0\|^2 \alpha(x).$$

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

Observație: Având în vedere că, în ipoteza $f \in \mathcal{C}^2(V)$, matricea formei pătratice $((d^2 f)(x_0))$, adică matricea $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$, **denumită hessiană lui f în x_0** și notată cu $H_f(x_0)$, este, pe

baza criteriului lui Schwarz (v. Teorema 10.8), simetrică, se poate spune că toate valorile proprii ale acesteia sunt reale. Atunci, ținând seama de Teorema 8.5 și de Teorema de inerție a lui Sylvester (Teorema 8.2), se poate afirma că forma pătratică $((d^2 f)(x_0))(v)$, adică $\langle H_f(x_0)(v), v \rangle_e$, este pozitiv-definită atunci când toate valorile proprii ale matricii $H_f(x_0)$ sunt pozitive. Analog, forma pătratică $((d^2 f)(x_0))$ este negativ definită când toate valorile proprii ale matricii $H_f(x_0)$ sunt negative. În fine, dacă matricea $H_f(x_0)$ are valori proprii atât din \mathbb{R}_+ , cât și din \mathbb{R}_- , atunci $((d^2 f)(x_0))$ este o formă pătratică nedefinită. Combinând aceste remarci cu Teorema 11.3, putem formula următorul rezultat.

Propoziția 11.1 Fie $A \neq \emptyset$, $A \subseteq \mathbb{R}^n$, $x_0 \in \mathring{A}$, $V \in \mathcal{V}(x_0)$ și $f : A \rightarrow \mathbb{R}$ astfel încât $f \in \mathcal{C}^2(V)$. Dacă $\nabla f(x_0) = 0_{\mathbb{R}^n}$, atunci x_0 este un punct de maxim local al lui f , când $H_f(x_0)$ are toate valorile proprii negative și punct de minim local pentru f , când $H_f(x_0)$ are toate valorile proprii pozitive. Când $H_f(x_0)$ are cel puțin două valori proprii nenule și de semne contrare, atunci punctul critic x_0 este un punct ș.a pentru f . Dacă toate valorile proprii ale lui $H_f(x_0)$ sunt nule, nu putem decide natura lui x_0 .

În mod asemănător, ținând seama de Teorema 8.12, cât și de observația ce imediat o succede, putem formula, în virtutea Teoremei 11.3, următoarea propoziție:

Propoziția 11.2 Fie A , x_0 , V și f ca în enunțul Propoziției 11.1, iar $\Delta_0 = 1$, $\Delta_1 = \det[a_{11}]$, $\Delta_2 = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, ..., $\Delta_n = \det \left[(a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \right]$ minorii principali ai matricii $(a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$, unde $a_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0)$, $\forall i, j = \overline{1, n}$. Avem:

i) Dacă $\Delta_j > 0$, $\forall j = \overline{1, n}$, atunci $((d^2 f)(x_0))$ este pozitiv definită și 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 $((d^2f)(x_0))$ este negativ definită și deci x_0 este punct de maxim al lui f .

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

iv) Dacă șirul $(\Delta_j)_{j=\overline{1, n}}$ nu este în nici unul dintre cazurile de la i), ii) sau iii), atunci forma pătratică $(d^2f)(x_0)$ este nedefinită și deci x_0 nu este punct de extrem, ci punct șa al lui f .

Observație: În cazul particular în care $n = 2$, Propoziția 11.2 revine la a spune că, dacă $f : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ este de clasă \mathcal{C}^2 pe o vecinătate a unui punct critic (pentru f) $x_0 \in A$ și adoptăm notațiile $p = \frac{\partial^2 f}{\partial x_1^2}(x_0)$, $q = \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_0)$, $r = \frac{\partial^2 f}{\partial x_2^2}(x_0)$, atunci:

- i) când $p > 0$ și $pr - q^2 > 0$, punctul $x_0 = (x_1^0, x_2^0)$ este de minim pentru f ;
- ii) când $p < 0$ și $pr - q^2 > 0$, 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 diferențialei 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 funcția $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, definită prin

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

este, evident, de clasă \mathcal{C}^2 pe \mathbb{R}^2 și are gradientul ∇f nul în acel punct (c_1^0, c_2^0) pentru care:

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

Rezolvând acest sistem algebric liniar, echivalent cu sistemul

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

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 situația $a_1 = a_2 = \dots = a_l$, pe baza inegalității lui Cauchy-Schwarz-Buniakowski – obținem:

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

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

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

În continuare, deoarece pentru abordarea teoretică a unei probleme de extrem cu restricții (legături) avem nevoie de elemente relative la noțiunile de funcție implicită, inversabilitatea unei funcții de mai multe variabile reale și dependența funcțională a unui set de funcții, prezentăm, pe scurt, astfel de noțiuni și, fără demonstrații, rezultate de bază care le privesc.

Funcții implicite

Definiția 11.5 Fie ecuația $F(x_1, x_2, \dots, x_n, y) = 0$, unde $F : A \times B \subseteq \mathbb{R}^{n+1} \rightarrow \mathbb{R}$, $A \subseteq \mathbb{R}^n$ și $B \subseteq \mathbb{R}$. O funcție $f : A \rightarrow \mathbb{R}$ este o **soluție**, în raport cu y , a ecuației $F(x_1, x_2, \dots, x_n, y) = 0$, pe mulțimea A , dacă, pentru orice $(x_1, x_2, \dots, x_n) \in A$, avem:

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

O asemenea funcție $y = f(x_1, x_2, \dots, x_n)$, definită prin intermediul ecuației $F(x_1, x_2, \dots, x_n, y) = 0$ se numește **funcție implicită** sau **funcție definită implicit**.

Observații:

- a) Dacă notăm (x_1, x_2, \dots, x_n) cu \mathbf{x} , ecuația $F(x_1, x_2, \dots, x_n, y) = 0$ se poate rescrie sub forma $F(\mathbf{x}, y) = 0$, iar soluția $y = f(x_1, x_2, \dots, x_n)$ se redă prin $y = f(\mathbf{x})$.
- b) O ecuație $F(\mathbf{x}, y) = 0$ poate să aibă, pe A , mai multe soluții sau nici una. De exemplu, ecuația $F(x, y) = 0$, în care $F : A \times \mathbb{R} \rightarrow \mathbb{R}$ este definită prin $F(\mathbf{x}, y) = y^2 - x$, unde $A \subseteq \mathbb{R}_+$, nu are nici o soluție, pe când dacă $A \subseteq \mathbb{R}_+^*$ ea are două soluții: $y = -\sqrt{x}$ și $y = \sqrt{x}$.

Teorema 11.4 (a funcțiilor implicite)

Fie $F : A \times B \subseteq \mathbb{R}^{n+1} \rightarrow \mathbb{R}$, $A \subseteq \mathbb{R}^n$, $B \subseteq \mathbb{R}$, $\mathbf{a} = (a_1, a_2, \dots, a_n) \in \mathring{A}$ și $b \in \mathring{B}$.

Dacă

- i) $F(\mathbf{a}, b) = 0$,
- ii) există $U \times V \subseteq A \times B$, vecinătate a punctului (\mathbf{a}, b) , astfel încât $F \in \mathcal{C}^1(U \times V)$,
- iii) $\frac{\partial F}{\partial y}(\mathbf{a}, b) \neq 0$,

atunci:

- i) există $U_0 \subseteq U$ - vecinătate deschisă a punctului \mathbf{a} și o funcție unică $f : U_0 \rightarrow V_0$ astfel încât $f(\mathbf{a}) = b$ și $F(\mathbf{x}, f(\mathbf{x})) = 0$, $\forall \mathbf{x} \in U_0$.
- ii) $f \in \mathcal{C}^1(U_0)$ și

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

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

Definiția 11.5 și Teorema 11.4 pot fi extinse la sisteme de ecuații și de funcții implicit-definite. Astfel, avem:

Definiția 11.6 Un sistem de ecuații

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

în care $F_k : A \times B \subseteq \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ sunt funcții de expresii cunoscute, cu $A \subseteq \mathbb{R}^n$ și $B \subseteq \mathbb{R}^m$, se numește **sistem de funcții implicite**.

O **soluție** a unui astfel de sistem, cu necunoscutele y_1, y_2, \dots, y_m , este un set de m funcții reale (implicit-definite)

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

cu $f_k : A \rightarrow \mathbb{R}$, $\forall k = \overline{1, m}$, satisfăcând relațiile:

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

Teorema 11.5 (a funcțiilor vectoriale implicit-definite)

Fie sistemul de ecuații

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

în care $F_k : A \times B \rightarrow \mathbb{R}$, $A \subseteq \mathbb{R}^n$, $B \subseteq \mathbb{R}^m$, $\forall k = \overline{1, n}$ și $(x_1^0, x_2^0, \dots, x_n^0; y_1^0, y_2^0, \dots, y_m^0)$ un punct interior al mulțimii $A \times B$ (adică $(x_1^0, x_2^0, \dots, x_n^0) \in \overset{\circ}{A}$ și $(y_1^0, y_2^0, \dots, y_m^0) \in B$).

Dacă

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

2) există $U \times V \subseteq A \times B$, o vecinătate a punctului (x_0, y_0) (cu $x_0 = (x_1^0, x_2^0, \dots, x_n^0)$ și $y_0 = (y_1^0, y_2^0, \dots, y_m^0)$), astfel încât $F_k \in C^1(U \times V)$, $\forall k = \overline{1, m}$ și

3) jacobianul $\frac{D(F_1, F_2, \dots, F_m)}{D(y_1, y_2, \dots, y_m)}(x_0, y_0) = \det \left[\left(\frac{\partial F_j}{\partial y_k}(x_0, y_0) \right)_{\substack{1 \leq j \leq m \\ 1 \leq k \leq m}} \right]$ 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 $V_0 \subseteq V \subseteq \mathbb{R}^m$ și o funcție unică $f : U_0 \rightarrow V_0$, $f(x) = (f_1(x), f_2(x), \dots, f_m(x))$, astfel încât $f_j(x_0) = y_j^0$, $\forall j = \overline{1, m}$ și $F_j(x, f(x)) = 0$, $\forall j = \overline{1, m}$, $\forall x = (x_1, x_2, \dots, x_n) \in U_0$;

β) Funcțiile f_1, f_2, \dots, f_m sunt de clasă C^1 pe U_0 , având derivatele parțiale de ordinul întâi, pe U_0 , date de formulele:

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

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

**Inversabilitatea funcțiilor diferențiabile, de mai multe
variable reale**

Se știe că un sistem algebric liniar de tipul

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

are o soluție unică dacă și numai dacă $\det \begin{bmatrix} (a_{ij})_{1 \leq i \leq n} \\ 1 \leq j \leq n \end{bmatrix} \neq 0$. Altfel spus, funcția inversă $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, cu expresia

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

adică, matriceal, $f(x) = Ax$, unde $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ și $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$ - este bijectivă (și ca atare, **inversabilă în vecinătatea oricărui punct** $x_0 = (x_1^0, x_2^0, \dots, x_n^0) \in \mathbb{R}^n$; **adică inversabilă global**) dacă și numai dacă $\frac{D(f_1, f_2, \dots, f_n)}{D(x_1, x_2, \dots, x_n)}(x) = \det A \neq 0$.

Extinderea unui asemenea rezultat la cazul sistemelor neliniare de forma

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

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

Teorema 11.6 (de inversabilitate locală)

Fie $A \subseteq \mathbb{R}^n$ o mulțime nevidă și deschisă, iar $f : A \rightarrow \mathbb{R}^n$ o funcție de clasă \mathcal{C}^1 pe A . De asemenea, fie $x_0 \in A$. Dacă $(df)(x_0)$ este o bijecție (liniară) de la \mathbb{R}^n la \mathbb{R}^n , adică dacă $\det(J_f(x_0)) \neq 0$, atunci există o vecinătate deschisă U a lui x_0 , $U \subseteq A$ și o vecinătate deschisă V a lui $f(x_0)$, astfel încât $f : U \rightarrow V$ 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 .

Observații:

- i) Teorema 11.6 poate fi reformulată după cum urmează: "O funcție $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$, $f \in \mathcal{C}^1(A)$, este inversabilă în vecinătatea oricărui punct $x_0 \in \overset{\circ}{A}$ în care (df) există și este inversabilă (ca aplicație liniară de la \mathbb{R}^n la \mathbb{R}^n)".
- ii) În condițiile Teoremei 11.6, cu $f \in \mathcal{C}^1(U; V)$ și $f^{-1} \in \mathcal{C}^1(V; U)$, avem: $(df)(x) \circ (d(f^{-1}))(f(x)) = (d(f^{-1}))(f(x)) \circ (df)(x) = I(x)$, $\forall x \in U$. Așadar, putem scrie: $(d(f^{-1}))(f(x)) = ((df)(x))^{-1}$. Adică, matricea jacobiană $J_{f^{-1}}$, calculată în punctul $f(x)$, este inversa matricii $J_f(x)$. Acest fapt permite calculul derivatelor parțiale ale lui f^{-1} în funcție de derivatele parțiale ale lui f .

Dependența sau independența funcțională a unui ansamblu de funcții

Definiția 11.7 Fie $A \subseteq \mathbb{R}^n$ o mulțime nevidă, iar f_1, f_2, \dots, f_m , cu $m \leq n$, funcții definite pe A și cu valori în \mathbb{R} . Se spune despre o funcție $f_0 : A \rightarrow \mathbb{R}$ că **depinde (este dependentă) funcțional** (sau că este **funcțional-dependentă**) de f_1, f_2, \dots, f_m , **pe** $\tilde{A} \subseteq A$ (cu $\tilde{A} \neq \emptyset$), dacă, oricare ar fi

$x_0 \in \tilde{A}$, există o vecinătate $U \in \mathcal{V}(x_0)$ și o funcție $\varphi : V \rightarrow \mathbb{R}$, unde V este o vecinătate a punctului $(f_1(x_0), f_2(x_0), \dots, f_m(x_0))$, astfel încât:

$$f_0(x) = \varphi(f_1(x), f_2(x), \dots, f_m(x)), \forall x \in U \cap \tilde{A}.$$

Altminteri, f_0 se numește **funcțional-independentă de** f_1, f_2, \dots, f_m , pe \tilde{A} .

În context, despre un ansamblu finit de funcții de la A în \mathbb{R} se spune că este **în dependență funcțională** atunci când una dintre funcții este dependentă funcțional de celelalte. În caz contrar, ansamblul respectiv este denumit **independent funcțional**.

Propoziția 11.3 Fie $A \subseteq \mathbb{R}^n$ o mulțime nevidă și deschisă, iar $f : A \rightarrow \mathbb{R}^m$ (cu $m \leq n$) o funcție de clasă \mathcal{C}^1 pe A , de componente $f_1, f_2, \dots, f_m : A \rightarrow \mathbb{R}$, a cărei matrice jacobiană $J_f = \left(\frac{\partial f_i}{\partial x_j} \right)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ are rangul maxim (adică m) în orice punct din A . Atunci, o funcție $f_0 : A \rightarrow \mathbb{R}$, de clasă \mathcal{C}^1 pe A , depinde funcțional de f_1, f_2, \dots, f_m dacă și numai dacă există m funcții continue $\psi_1, \psi_2, \dots, \psi_m : A \rightarrow \mathbb{R}$ astfel încât

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

Altfel spus, în ipotezele Propoziției 11.3, f_0 depinde funcțional de f_1, f_2, \dots, f_m dacă df_0 se poate exprima, pe A , ca o combinație liniară de df_1, df_2, \dots, df_m .

Teorema 11.7 (a dependenței funcționale)

Fie $A \subseteq \mathbb{R}^n$ o mulțime deschisă și nevidă, iar $f : A \rightarrow \mathbb{R}^n$, cu $m \leq n$, de componente $f_1, f_2, \dots, f_m : A \rightarrow \mathbb{R}$, o funcție de clasă \mathcal{C}^1 pe A , a cărei matrice jacobiană J_f are rang constant – $r \leq m$ – în orice punct $x \in A$. Atunci funcțiile f_1, f_2, \dots, f_m satisfac, local, $m - r$ relații de dependență funcțională.

Cu alte cuvinte, dacă $\text{rang } J_f(x) = r \leq m$, $\forall x \in A$, atunci doar r dintre funcțiile f_1, f_2, \dots, f_m sunt funcțional independente, iar celelalte $m - r$ sunt dependente de acele r funcții independente între ele.

Extreme cu legături

În realitatea obiectivă, pe lângă situațiile în care se cer rezolvate probleme de extrem libere, fără restricții, precum cele din Exemplele 11.1 și 11.2, în parte, se întâlnesc și cazuri de probleme în care extremele unei funcții se caută în condițiile satisfacerii, de către argumentele funcției respective, a unor relații de legătură între ele. În asemenea cazuri, extremele respective poartă denumirea de **extreme condiționate** sau, echivalent, **extreme cu legături**, legături ce se exprimă, adeseori, prin relații de egalitate.

Analizăm aici următoarea situație, în care D este o mulțime deschisă, în raport cu topologia uzuală, din $\mathbb{R}^n \times \mathbb{R}^m$ ($n, m \in \mathbb{N}^*$), iar $f : D \rightarrow \mathbb{R}$ și $g : F \rightarrow \mathbb{R}^m$ sunt funcții de clasă \mathcal{C}^1 pe D , astfel încât se caută punctele de extrem ale funcției f , supuse la condiția suplimentară $g(x, y) = 0$, unde $x \in B \subseteq \mathbb{R}^n$ și $y \in E \subseteq \mathbb{R}^m$, cu $B \times E \subseteq D$.

Notând cu A mulțimea $\{(x, y) \in D \mid g(x, y) = \mathbf{0}_{\mathbb{R}^m}\}$ și cu g_1, g_2, \dots, g_m componentele funcției g , se vede că problema de mai sus cere, de fapt, determinarea punctelor de extrem ale restricției lui f la A , adică acele puncte de extrem ale lui f care satisfac, simultan, relațiile $g_1(x, y) = 0, g_2(x, y) = 0, \dots, g_m(x, y) = 0$. Cum g_1, g_2, \dots, g_m sunt continue, rezultă că g este continuă și deci A este o mulțime închisă, în raport cu topologia amintită. Dacă B și E sunt mărginite, atunci și A este mărginită. În acest caz, fiind închisă și mărginită, A este o mulțime compactă. Ca atare, fiind

continuă pe A , f va avea, cu siguranță, puncte de extrem în A . Se mai pune, deci, doar problema determinării lor. În acest scop, este de observat că metoda de evidențiere a acestor puncte, folosită în cazul extremelor libere, nu este de utilizat, deoarece punctele lui A vor aparține interiorului acestei mulțimi dacă și numai dacă $g_1 \equiv g_2 \equiv \dots \equiv g_m \equiv 0$, 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, \dots, g_m , am putea încerca să utilizăm teorema funcțiilor implicite (Teorema 11.4 sau Teorema 11.5) pentru a soluționa mai întâi sistemul $g_k(x, y) = 0$, $\forall k = \overline{1, m}$, în raport cu y . Scoțând $y = \varphi(x)$, am putea reduce problema inițială de extrem la o problemă de extrem liber pentru funcția obiectiv $f(x, \varphi(x))$, cu $x \in B$. În această situație am putea studia extremele lui $f(x, \varphi(x))$ pe calea prezentată în prima parte a materialului de față. Procedul ar fi însă dificil de folosit în practică, deoarece, în general, ecuația vectorială $g(x, y) = \mathbf{0} \in \mathbb{R}^n$ ar fi greu de rezolvat în raport cu y și, chiar dacă, teoretic, acest lucru ar fi posibil, nu ar fi ușor de ajuns la $f(x, \varphi(x))$.

Iată de ce, în realitate, se folosește așa-numita **metodă a multiplicatorilor lui Lagrange**, descrisă în teorema care urmează după definiția dată în continuare.

Definiția 11.8 a) Se spune că $(x_0, y_0) \in A$ este un **punct de extrem local al funcției $f : D \subseteq \mathbb{R}^{n+m} \rightarrow \mathbb{R}$, cu restricțiile $g_1(x, y) = 0, g_2(x, y) = 0, \dots, g_m(x, y) = 0$, unde $g_k : D \rightarrow \mathbb{R}$, $\forall k = \overline{1, m}$, (sau că (x_0, y_0) este punct de extrem local, condiționat, al lui f)**, dacă există o vecinătate $V \subseteq D$ a lui (x_0, y_0) astfel încât $f(x, y) - f(x_0, y_0)$ să aibă semn constant $\forall (x, y) \in V \cap A$.

Spunem că $(x_0, y_0) \in A$ este un **punct de minim local, condiționat, al funcției f** dacă $f(x, y) - f(x_0, y_0) \geq 0$, $\forall (x, y) \in V \cap A$. Analog, (x_0, y_0) este un **punct de maxim, condiționat, pentru f** dacă $f(x, y) - f(x_0, y_0) \leq 0$, $\forall (x, y) \in V \cap A$.

Altfel exprimat, (x_0, y_0) este un punct de extrem (minim sau maxim) local, condiționat, pentru f dacă el este un punct de extrem (minim și respectiv maxim) local pentru restricția funcției f la mulțimea de legături (restricții) A .

b) Un punct (x_0, y_0) este un **punct critic, condiționat (de A) (sau punct staționar, condiționat)**, pentru f , dacă el este un punct critic (staționar) al restricției lui f la A .

Teorema 11.8 (de existență a multiplicatorilor Lagrange)

Fie D o mulțime nevidă și deschisă din $\mathbb{R}^n \times \mathbb{R}^m$, $f : D \rightarrow \mathbb{R}$ și $g : D \rightarrow \mathbb{R}^m$, adică $g = (g_1, g_2, \dots, g_m)$, cu $g_k : D \rightarrow \mathbb{R}$, $\forall k = \overline{1, m}$, funcții de clasă \mathcal{C}^1 pe D , iar $(x_0, y_0) \in D$ un punct de extrem local, condiționat, al lui f , cu restricțiile $g_k(x_0, y_0) = 0$, $\forall k = \overline{1, m}$.

Dacă $\frac{D(g_1, g_2, \dots, g_m)}{D(y_1, y_2, \dots, y_m)}(x_0, y_0) \neq 0$, adică dacă g_1, g_2, \dots, g_m sunt funcțional independente, în raport cu y_1, y_2, \dots, y_m , în (x_0, y_0) , atunci există m numere reale $\lambda_1, \lambda_2, \dots, \lambda_m$, astfel încât, dacă se consideră funcția

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

punctul (x_0, y_0) să-i fie punct critic, adică să satisfacă sistemul de relații:

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

Precizări:

- 1) Numerele $\lambda_1, \lambda_2, \dots, \lambda_m$ se numesc *multiplicatori Lagrange*.
- 2) Punctele de extrem condiționat (în raport cu restricțiile $g_k(x, y) = 0, \forall k = \overline{1, m}$) pentru funcția f se găsesc printre punctele critice, condiționate, ale lui f , adică printre punctele critice ale funcției asociate $L = f + \sum_{k=1}^m \lambda_k g_k$.
- 3) Dacă, în expresia $f(x, y) + \sum_{k=1}^m \lambda_k g_k(x, y)$, admitem că și $\lambda_1, \lambda_2, \dots, \lambda_m$ sunt variabile, atunci funcția $(x, y, \lambda) \mapsto L(x, y, \lambda) = f(x, y) + \sum_{k=1}^m \lambda_k g_k(x, y)$ poartă denumirea de *Lagrangean* asociat lui f și restricțiilor $g_k, \forall k = \overline{1, m}$.
- 4) Dacă f și g_k sunt funcții de clasă \mathcal{C}^2 pe $D, \forall k = \overline{1, m}$, sau cel puțin pe o vecinătate a unei soluții $(x_0, y_0; \lambda_0) = (x_1^0, x_2^0, \dots, x_n^0, y_1^0, y_2^0, \dots, y_m^0, \lambda_1^0, \lambda_2^0, \dots, \lambda_m^0)$ a sistemului de ecuații (*), atunci, în virtutea Teoremei 11.8, punctul (x_0, y_0) este unul critic pentru funcția $L(x_0, y_0; \lambda_0)$, adică avem $d(L(x, y; \lambda_0))(x_0, y_0) = 0$. Totodată, deoarece $L(x, y; \lambda_0) - L(x_0, y_0; \lambda_0) = f(x, y) - f(x_0, y_0), \forall (x, y) \in A$, rezultă că (x_0, y_0) este un punct de extrem, condiționat de legăturile $g_k(x, y) = 0, \forall k = \overline{1, m}$, pentru f , dacă și 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 $(d^2 L(x, y; \lambda_0))(x_0, y_0)$. Altfel spus, trebuie studiată expresia $\sum_{i,j=1}^{n+m} \frac{\partial^2 L}{\partial \tilde{x}_i \partial \tilde{x}_j}(x_0, y_0) d\tilde{x}_i d\tilde{x}_j$, unde $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_{n+m}) = (x, y)$. Cum diferențialele $d\tilde{x}_1, d\tilde{x}_2, \dots, d\tilde{x}_{n+m}$ nu sunt independente, relațiile de dependență dintre ele se obțin prin diferențierea ecuațiilor $g_k(\tilde{x}) = 0, \forall k = \overline{1, m}$, anume: $\sum_{k=1}^{n+m} \frac{\partial g_i}{\partial \tilde{x}_j}(\tilde{x}_0) d\tilde{x}_k = 0$. Cum $\frac{D(g_1, g_2, \dots, g_m)}{D(\tilde{x}_{n+1}, \tilde{x}_{n+2}, \dots, \tilde{x}_{n+m})}(x_0, y_0) \neq 0$, de aici putem găsi $d\tilde{x}_{n+1}, d\tilde{x}_{n+2}, \dots, d\tilde{x}_{n+m}$ în funcție de $d\tilde{x}_1, d\tilde{x}_2, \dots, d\tilde{x}_n$, prin expresii liniare. Aceste expresii, introduse în $\sum_{i,j=1}^{n+m} \frac{\partial^2 L}{\partial \tilde{x}_i \partial \tilde{x}_j}(x_0, y_0) d\tilde{x}_i d\tilde{x}_j$ conduc la relația $d^2 L(\tilde{x}_0) = \sum_{i,j=1}^n a_{ij}^0 dx_i dx_j$, reprezentând o formă pătratică pe \mathbb{R}^n , în care a_{ij}^0 depind liniar de $\frac{\partial^2 g_k}{\partial \tilde{x}_i \partial \tilde{x}_j}(x_0, y_0), \frac{\partial g_i}{\partial x_j}(x_0, y_0)$ și $\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0, y_0)$. Astfel, pe baza Teoremei 11.3, deducem că, dacă $d^2 L(\tilde{x}_0)$ este o formă pătratică pozitiv definită (respectiv negativ definită), atunci \tilde{x}_0 este punct de minim (respectiv maxim) condiționat, local, pentru f , iar dacă $d^2 L(\tilde{x}_0)$ reprezintă o formă pătratică nedefinită, atunci \tilde{x}_0 nu este punct de extrem pentru f , condiționat de legăturile $g_k(\tilde{x}) = 0, \forall k = \overline{1, m}$. În fine, dacă $d^2 L(\tilde{x}_0)$ este semi-definită (pozitiv sau negativ), atunci nu putem stabili natura punctului \tilde{x}_0 numai pe baza diferențialei a doua a lui L .

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

Demonstrație: (pentru Teorema 11.8)

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

lui x_0 , și o funcție $\varphi : U \rightarrow \mathbb{R}^n$, de clasă \mathcal{C}^1 pe U (adică există $\varphi_k : U \rightarrow \mathbb{R}$, $\forall k = \overline{1, m}$, $\varphi_k \in \mathcal{C}^1(U)$) așa încât $\varphi(x_0) = y_0$ și $g_k(x, \varphi(x)) = 0$, $\forall k = \overline{1, m}$, $\forall x \in U$. De aici, rezultă că avem (în virtutea regulii "lanțului"), prin derivare:

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

Considerând funcția $F(x) = f(x, \varphi(x))$, $\forall x \in U$, putem zice că, deoarece (x_0, y_0) este un punct de extrem local, cu restricțiile $g_k(x, y) = 0$, $\forall k = \overline{1, m}$, pentru f , iar $(x, \varphi(x)) \in A$, $\forall x \in U$, punctul x_0 este unul de extrem local, necondiționat, pentru F pe U . Atunci, potrivit Teoremei 11.1 (Fermat), avem $\frac{\partial F}{\partial x_k}(x_0) = 0$, $\forall k = \overline{1, n}$. Cu alte cuvinte, avem:

$$(!) \quad \frac{\partial f}{\partial x_k}(x_0, y_0) + \sum_{j=1}^m \frac{\partial f}{\partial x_j}(x_0, y_0) \frac{\partial \varphi_j}{\partial x_k}(x_0) = 0, \forall k = \overline{1, n}.$$

Cum, $\frac{D(g_1, g_2, \dots, g_m)}{D(y_1, y_2, \dots, y_m)}(x_0, y_0) \neq 0$, reiese că sistemul liniar algebric

$$(* *) \quad \sum_{k=1}^m \frac{\partial g_k}{\partial y_j}(x_0, y_0) \alpha_k = - \frac{\partial f}{\partial y_j}(x_0, y_0), \forall j = \overline{1, m}$$

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

Într-adevăr, cum $(x_0, y_0) \in A$, avem $g_k(x_0, y_0) = 0$, $\forall k = \overline{1, m}$.

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

$$\begin{aligned} \frac{\partial L}{\partial x_i}(x_0, y_0) &= \frac{\partial f}{\partial x_i}(x_0, y_0) + \sum_{k=1}^m \lambda_k^0 \frac{\partial g_k}{\partial x_i}(x_0, y_0) = \\ &= \frac{\partial f}{\partial x_i}(x_0, y_0) - \sum_{k=1}^m \lambda_k^0 \cdot \left(\sum_{j=1}^m \frac{\partial g_k}{\partial y_j}(x_0, y_0) \cdot \frac{\partial \varphi_j}{\partial x_i}(x_0) \right) = \\ &= \frac{\partial f}{\partial x_i}(x_0, y_0) - \sum_{k=1}^m \frac{\partial \varphi_j}{\partial x_i}(x_0) \cdot \left(\sum_{j=1}^m \lambda_k^0 \cdot \frac{\partial g_k}{\partial y_j}(x_0, y_0) \right) = \\ &= \frac{\partial f}{\partial x_i}(x_0, y_0) + \sum_{k=1}^m \frac{\partial \varphi_j}{\partial x_i}(x_0) \cdot \frac{\partial f}{\partial y_j}(x_0, y_0), \end{aligned}$$

dacă se ține seama și de relațiile (**) și (* * *). De aici, folosind (!), se deduce că $\frac{\partial L}{\partial x_i}(x_0, y_0) = 0$, $\forall i = \overline{1, n}$

În fine, derivând L în raport cu y_j și ținând, iarăși, cont de faptul că $(\lambda_1^0, \lambda_2^0, \dots, \lambda_m^0)$ este soluția sistemului (* * *), obținem:

$$\frac{\partial L}{\partial y_j}(x_0, y_0) = \frac{\partial f}{\partial y_j} + \sum_{k=1}^m \lambda_k^0 \frac{\partial g_k}{\partial y_j}(x_0, y_0) = 0, \forall j = \overline{1, m}.$$

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

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

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

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

Atunci sistemul de tip (*), este următorul:

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

Din primele n ecuații, reiese că $p_1 = p_2 = \dots = p_n = \psi(\lambda_1)$, iar dacă se ține seama și de ultima ecuație, rezultă soluția:

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

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

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

ceea ce înseamnă că punctul $(p_1^0, p_2^0, \dots, p_n^0) = \left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n} \right)$ este unul de maxim pentru \tilde{H} , în condiția $g_1(p_1, p_2, \dots, p_n) = 0$.

Prin urmare, entropia informațională maximă se obține pentru variabila aleatoare

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

a cărei distribuție de probabilitate este uniformă.

Bibliografie recomandată

1. Anca Precupanu - *Bazele analizei matematice (§11.6)*, Editura Polirom, Iași, 1998.
2. Emil Popescu - *Analiză matematică. Calcul diferențial (cap. 7 și 8)*, Editura Matrix Rom, București, 2006.
3. Maria Gorunescu - *Lecții de analiză matematică pentru informaticieni (cap. VII)*, Reprografia Universității Craiova, 2000.

4. F. Iacob - *Matematică pentru anul II - ID*, seria 2004-2005 (tema 3, modulul 3).
5. Rodica Luca-Tudorache - *Analiză matematică. Calcul diferențial (§6.6)*, Editura Tehnopress, Iași, 2005.
6. V. Postolică - *Eficiență prin matematică aplicată. Analiză matematică (§7.4)*, Editura Matrix Rom, București, 2006.
7. M. Postolache - *Analiză matematică (teorie și aplicații)*, Editura "Fair Partners", București, 2011.