Viazané lokálne extrémy funkcií viac premenných

Peter Šepitka

jaro 2016

Obsah

Obsah

Motivácia

Pojem viazaného lokálneho extrému funkcie viac premenných súvisí s hľadaním globálnych extrémov na podmnožinách v $\mathbb{R}^n$. Konkrétne, pri ich určovaní na kompaktnej množine $N \subseteq \mathbb{R}^n$ sme postupovali tak, že sme našli

(i) jednak všetky lokálne extrémy danej funkcie vo vnútri N,

(ii) a jednak všetky jej extremálne hodnoty na hranici ∂N.

V kroku (ii) teda hľadáme lokálne extrémy danej funkcie, ktoré sú viazané na hranicu ∂N. Uvažujme napríklad nejakú funkciu $f(x, y, z)$ na množine

$$N = \{[x, y, z] \in \mathbb{R}^3 \,:\, x^2 + y^2 + z^2 \leq 1,\ x, y, z \geq 0\}.$$

Pri vyšetrovaní funkcie $f(x, y, z)$ na časti hranice ∂N tvorenej guľovou plochou $x^2 + y^2 + z^2 = 1$ postupujeme tak, že vyjadríme $z = \sqrt{1 - x^2 - y^2}$, a následne hľadáme lokálne extrémy funkcie $f(x, y, \sqrt{1 - x^2 - y^2})$ na množine

$$\tilde{N} = \{[x, y] \in \mathbb{R}^3 \,:\, x^2 + y^2 \leq 1,\ x, y \geq 0\}.$$

Podobne, na časti hranice $\partial\tilde{N}$ tvorenej štvrťkružnicou $x^2 + y^2 = 1$ vyjadríme $y = \sqrt{1 - x^2}$ a vyšetrujeme funkciu $f(x, \sqrt{1 - x^2}, 0)$ pre $x \in [0, 1]$.

Týmto spôsobom prevedieme pôvodný problém vyšetrenia funkcie $f(x, y, z)$ troch premenných na hranici množiny N na úlohu hľadania extrémov funkcie $f(x, \sqrt{1-x^2}, 0)$ jednej premennej na intervale $[0, 1]$. Je však zrejmé, že tento postup je zdĺhavý a pri väčšom počte premenných i nepraktický. Okrem toho postupná eliminácia premenných nemusí byť pri všeobecnom zadaní množiny N vôbec riešiteľný problém.

Definícia 1 (Lokálny extrém funkcie vzhľadom na množinu)

Nech f je funkcia n premenných a $M \subseteq \mathcal{D}(f)$ je neprázdna množina. Povieme, že funkcia f má v bode $x^* \in M$ lokálne minimum (maximum) vzhľadom na množinu M, ak existuje okolie $\mathcal{O}(x^*)$ také, že pre každé $x \in \mathcal{O}(x^*) \cap M$ platí nerovnosť $f(x) \geq f(x^*)$ $(f(x) \leq f(x^*))$. V prípade, ak sú dané nerovnosti pre $x \neq x^*$ ostré, hovoríme o ostrých lokálnych extrémoch funkcie f vzhľadom na množinu M.

V praktických úlohách sa obvykle študujú situácie, v ktorých je množina M daná systémom nerovností a rovností

$$h_1(x) \leq 0,\ h_2(x) \leq 0,\ \ldots,\ h_k(x) \leq 0, \qquad g_1(x) = 0,\ g_2(x) = 0,\ \ldots,\ g_m(x) = 0.$$

V tomto prípade sa namiesto pomenovania lokálny extrém vzhľadom na M používa termín lokálny extrém viazaný danými podmienkami, resp. väzbami.

Obsah

Optimalizačná úloha

V nasledujúcom výklade sa budeme zaoberať optimalizačnou úlohou typu

$$f(x) \;\to\; \max(\min), \quad x \in M, \tag{1}$$

$$M \;:\; g_1(x) = 0, \quad g_2(x) = 0, \quad \dots, \quad g_m(x) = 0, \tag{2}$$

kde $x = [x_1, \dots, x_n]$, $f, g_k : \mathbb{R}^n \to \mathbb{R}$, $k = 1, \dots, m$, sú funkcie n premenných a $1 \leq m < n$. Rovnosti v (2), udávajúce množinu M, sa nazývajú väzbové podmienky danej optimalizačnej úlohy. Ďalej budeme predpokladať, že funkcie f, g_k majú spojité parciálne derivácie prvého rádu podľa premenných $x_1, \dots, x_n$ na otvorenej množine $U \subseteq \mathbb{R}^n$ a matica

$$G(x) = \begin{pmatrix} \frac{\partial g_1(x)}{\partial x_1} & \dots & \frac{\partial g_1(x)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m(x)}{\partial x_1} & \dots & \frac{\partial g_m(x)}{\partial x_n} \end{pmatrix} \quad \text{má v každom } x \in U \text{ hodnosť } m. \tag{3}$$

Poznámka 1

Poznamenajme, že podmienka (3) je ekvivalentná so skutočnosťou, že vektory $\operatorname{grad} g_k(x)$, $k = 1, \dots, m$, sú lineárne nezávislé v každom bode x množiny U.

Viazaný lokálny extrém – nutná podmienka

Veta 1 (Nutná podmienka existencie viazaného extrému)

Nech platí podmienka (3) a nech v bode $x^* \in M$ má funkcia f lokálny extrém vzhľadom na množinu M, t.j., x^* spĺňa optimalizačnú úlohu (1) a (2). Potom existujú reálne čísla $\lambda_1, \dots, \lambda_m$ tak, že pre každé $i = 1, \dots, n$ platia rovnosti

$$\frac{\partial f}{\partial x_i}(x^*) - \sum_{k=1}^{m} \lambda_k \frac{\partial g_k}{\partial x_i}(x^*) = 0. \tag{4}$$

Poznámka 2

Rovnica (4) je ekvivalentná so skutočnosťou, že gradienty funkcií f, g_k spĺňajú v bode x^* identitu

$$\mathrm{grad}\, f(x^*) = \lambda_1 \cdot \mathrm{grad}\, g_1(x^*) + \lambda_2 \cdot \mathrm{grad}\, g_2(x^*) + \dots + \lambda_m \cdot \mathrm{grad}\, g_m(x^*),$$

t.j., vektor $\mathrm{grad}\, f(x^*)$ je lineárnou kombináciou vektorov $\mathrm{grad}\, g_k(x^*)$.

Lagrangeova funkcia a multiplikátory

Reálne konštanty $\lambda_1, \ldots, \lambda_m$ vo Vete 1 sa nazývajú Lagrangeove multiplikátory a funkcia $n+m$ premenných

$$L(x_1, \ldots, x_n, \lambda_1, \ldots \lambda_m) := f(x_1, \ldots, x_n) - \sum_{k=1}^{m} \lambda_k \, g_k(x_1, \ldots, x_n) \qquad (5)$$

sa označuje ako Lagrangeova funkcia optimalizačnej úlohy (1) a (2).

Definícia 2 (Stacionárny bod funkcie vzhľadom na množinu)

Nech množina $M \subseteq \mathbb{R}^n$ je daná systémom rovníc (2). Bod $x^* \in M$ sa nazýva stacionárny bod funkcie f vzhľadom na množinu M, ak existujú Lagrangeove multiplikátory $\lambda_1, \ldots, \lambda_m$ také, že sú splnené rovnosti (4).

Poznámka 3

Bod $x^* \in M$ je zrejme stacionárnym bodom funkcie f vzhľadom na množinu M práve vtedy, keď je stacionárnym bodom Lagrangeovej funkcie (5) pre istú voľbu Lagrangeových multiplikátorov $\lambda = (\lambda_1, \ldots, \lambda_m)$, t.j., platí

$$L'_{x_i}(x^*, \lambda) = 0 \quad \text{pre každé } i = 1, \ldots, n.$$

Príklad 1

Určme stacionárne body a Lagrangeove multiplikátory funkcie

$$f(x, y) = xy - x + y - 1$$

vzhľadom na množinu $M : x + y = 1$. V súlade s (2) má väzbová podmienka tvar $g(x, y) = x + y - 1 = 0$. Príslušná Lagrangeova funkcia v (5) potom je

$$L(x, y, \lambda) = f(x, y) - \lambda g(x, y) = xy - x + y - 1 - \lambda(x + y - 1).$$

Podľa Poznámky 3 nájdeme jej stacionárne body ležiace v M, t.j.,

$$L'_x(x, y, \lambda) = 0, \quad L'_y(x, y, \lambda) = 0, \quad x + y - 1 = 0.$$

Dostávame sústavu troch rovníc s tromi neznámymi x, y a λ

$$y - 1 - \lambda = 0, \quad x + 1 - \lambda = 0, \quad x + y - 1 = 0.$$

Máme jediné riešenie $x = -1/2$, $y = 3/2$ a $\lambda = 1/2$. Funkcia $f(x, y)$ má teda vzhľadom na množinu M jeden stacionárny bod $[-1/2, 3/2]$ s odpovedajúcim Lagrangeovým multiplikátorom $\lambda = 1/2$.

Príklad 2

Určme stacionárne body a Lagrangeove multiplikátory funkcie $f(x, y, z) = xyz$ vzhľadom na množinu M určenú rovnosťami

$$x^2 + y^2 + z^2 = 1, \quad x + y + z = 0.$$

V tomto prípade máme predpísané dve väzbové podmienky

$$g_1(x, y, z) = x^2 + y^2 + z^2 - 1 = 0, \quad g_2(x, y, z) = x + y + z = 0,$$

a teda odpovedajúca Lagrangeova funkcia má tvar

$$L(x, y, z, \lambda_1, \lambda_2) = xyz - \lambda_1(x^2 + y^2 + z^2 - 1) - \lambda_2(x + y + z).$$

Tri súradnice x, y a z hľadaných stacionárnych bodov a dva Lagrangeove multiplikátory λ_1, λ_2 stanovíme riešením systému piatich rovníc

$$L'_x = 0, \quad L'_y = 0, \quad L'_z = 0, \quad x^2 + y^2 + z^2 = 1, \quad x + y + z = 0,$$

Príklad 2

t.j., po výpočte príslušných parciálnych derivácií máme

$$yz - 2\lambda_1 x = \lambda_2, \quad xz - 2\lambda_1 y = \lambda_2, \quad xy - 2\lambda_1 z = \lambda_2,$$

$$x^2 + y^2 + z^2 = 1, \quad x + y + z = 0.$$

Dostaneme celkovo šesť stacionárnych bodov, konkrétne

$$\left.\begin{array}{l} \left[\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right], \\ \left[\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right], \\ \left[-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right] \end{array}\right\} \text{ pre } \quad \lambda_1 = -\frac{1}{2\sqrt{6}} \quad \text{a} \quad \lambda_2 = -\frac{1}{6},$$

$$\left.\begin{array}{l} \left[-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right], \\ \left[-\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right], \\ \left[\frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right] \end{array}\right\} \text{ pre } \quad \lambda_1 = \frac{1}{2\sqrt{6}} \quad \text{a} \quad \lambda_2 = -\frac{1}{6}.$$

Lokálny extrém a druhý diferenciál funkcie

Existenciu lokálneho extrému funkcie $f(x_1, \ldots, x_n)$ v jej stacionárnom bode x^*, t.j., v bode s $\operatorname{grad} f(x^*) = 0$, sme za predpokladu spojitosti parciálnych derivácií druhého rádu funkcie f a bez prítomnosti väzbových podmienok skúmali prostredníctvom definitnosti príslušnej Hessovej matice $H_f(x^*)$, t.j.,

$$H_f(x^*) = \begin{pmatrix} f''_{x_1x_1}(x^*) & \cdots & f''_{x_1x_n}(x^*) \\ \vdots & \ddots & \vdots \\ f''_{x_nx_1}(x^*) & \cdots & f''_{x_nx_n}(x^*) \end{pmatrix}. \tag{6}$$

Vyšetrovali sme vlastne definitnosť kvadratickej formy, ktorá odpovedá matici $H_f(x^*)$, konkrétne $h^T H_f(x^*)\, h$, kde $h = (h_1, \ldots, h_n)^T \in \mathbb{R}^n$. Táto forma je podľa definície rovná druhému diferenciálu funkcie f v bode x^*, teda

$$\mathrm{d}^2 f(x^*, h) = \sum_{i,j=1}^{n} f''_{x_ix_j}(x^*)\, h_i h_j \tag{7}$$

Ak kvadratická forma (7) je pozitívne (negatívne) definitná, potom funkcia f má v bode x^* lokálne minimum (maximum). V prípade indefinitnej formy (7) funkcia f nenadobúda v stacionárnom bode x^* lokálny extrém.

Viazaný lokálny extrém – postačujúca podmienka

Charakter viazaného stacionárneho bodu x^* funkcie f vzhľadom na množinu M v (2) budeme preto vyšetrovať pomocou definitnosti druhého diferenciálu Lagrangeovej funkcie $L(x, \lambda)$ v (5), t.j., pomocou kvadratickej formy

$$\mathrm{d}^2 L(x^*, \lambda, h) = \sum_{i,j=1}^{n} L''_{x_i x_j}(x^*, \lambda)\, h_i h_j, \tag{8}$$

kde $h = (h_1, \ldots, h_n)^T \in \mathbb{R}^n$ a $\lambda = (\lambda_1, \ldots, \lambda_m)$ sú Lagrangeove multiplikátory odpovedajúce stacionárnemu bodu x^*. Na rozdiel od lokálnych extrémov sme však teraz obmedzení väzbovými podmienkami (2). To sa odzrkadlí v tom, že definitnosť formy (8) nebudeme skúmať na celom priestore $\mathbb{R}^n$, ale iba pre

vektory h kolmé na vektory $\operatorname{grad} g_k(x^*)$ pre každé $k = 1, \ldots, m$.

Podpriestor všetkých takýchto vektorov $h \in \mathbb{R}^n$ sa nazýva dotykový priestor množiny M v bode x^* a označuje sa

$$\mathcal{T}_M(x^*) := \mathrm{Lin}\{\operatorname{grad} g_k(x^*),\ k = 1, \ldots, m\}^{\perp} \tag{9}$$

Poznámka 4

Z podmienky (3) a Poznámky 1 vyplýva, že dimenzia dotykového priestoru $\mathcal{T}_M(x^*)$ je rovná $n - m$. Okrem toho vektory $h \in \mathcal{T}_M(x^*)$ sú práve riešenia homogénneho systému $G(x^*)\,h = 0$ s maticou $G(x)$ definovanou v (3).

Veta 2 (Postačujúca podmienka existencie viazaného extrému)

Nech platí podmienka (3) *a nech* x^* *je stacionárny bod funkcie* f *vzhľadom na množinu* M *v* (2) *s Lagrangeovými multiplikátormi* $\lambda = (\lambda_1, \ldots, \lambda_m)$. *Ďalej nech funkcie* f, g_k, $k = 1, \ldots, m$, *majú spojité parciálne derivácie druhého rádu v bode* x^*. *Ak pre každý nenulový vektor* $h \in \mathcal{T}_M(x^*)$ *platí*

$$\mathrm{d}^2 L(x^*, \lambda, h) > 0, \quad \textit{resp.} \quad \mathrm{d}^2 L(x^*, \lambda, h) < 0, \tag{10}$$

potom funkcia f *má v bode* x^* *ostré lokálne minimum, resp. ostré lokálne maximum vzhľadom na množinu* M. *Ak pre nejaké nenulové* $\tilde{h}, \bar{h} \in \mathcal{T}_M(x^*)$ *je*

$$\mathrm{d}^2 L(x^*, \lambda, \tilde{h}) > 0 \quad \textit{a} \quad \mathrm{d}^2 L(x^*, \lambda, \bar{h}) < 0, \tag{11}$$

potom funkcia f *nemá v bode* x^* *lokálny extrém vzhľadom na množinu* M.

Poznámka 5

Nerovnosti v (10) znamenajú, že druhý diferenciál $\mathrm{d}^2 L(x^*, \lambda, h)$ Lagrangeovej funkcie je pozitívne, resp. negatívne definitný na podpriestore $\mathcal{T}_M(x^*)$, kým relácie v (11) vyjadrujú indefinitnosť diferenciálu $\mathrm{d}^2 L(x^*, \lambda, h)$ na $\mathcal{T}_M(x^*)$.

Predchádzajúci výklad popisuje hľadanie viazaných lokálnych extrémov funkcie f pomocou metódy Lagrangeových multiplikátorov. Hlavný princíp spočíva v zabudovaní väzbových podmienok (2) do samotného procesu hľadania extrému

- jednak prostredníctvom Lagrangeovej funkcie (Veta 1 a Poznámka 3),
- a jednak pri zisťovaní existencie/neexistencie extrému v stacionárnom bode (Veta 2 a Poznámky 4 a 5).

Hrubo povedané, namiesto hľadania viazaných lokálnych extrémov funkcie f na množine M zisťujeme lokálne extrémy príslušnej Lagrangeovej funkcie L, avšak už bez obmedzujúcich podmienok.

Viazaný lokálny extrém – praktický návod

Metódu Lagrangeových multiplikátorov na vyšetrovanie viazaných lokálnych extrémov funkcií možno zhrnúť do nasledujúceho praktického návodu.

1. Vytvoríme Lagrangeovu funkciu (5) a nájdeme jej stacionárne body x^* ležiace v M, t.j., riešime sústavu $n+m$ rovníc

$$L'_{x_i}(x, \lambda) = 0, \quad g_k(x) = 0, \qquad i = 1, \dots, n, \quad k = 1, \dots, m,$$

   s $n+m$ neznámymi $x = [x_1, \dots, x_n]$ a $\lambda = (\lambda_1, \dots, \lambda_m)$.

2. Pre daný stacionárny bod x^* a príslušné multiplikátory $\lambda = (\lambda_1, \dots, \lambda_m)$ zostavíme druhý diferenciál $\mathrm{d}^2 L(x^*, \lambda, h)$ Lagrangeovej funkcie, t.j.,

$$\mathrm{d}^2 L(x^*, \lambda, h) = \sum_{i,j=1}^{n} L''_{x_i x_j}(x^*, \lambda)\, h_i h_j, \quad h = (h_1, \dots, h_n)^T \in \mathbb{R}^n.$$

   V tejto kvadratickej forme znížime počet n premenných $h_1, \dots, h_n$ na počet $n-m$ premenných pomocou podmienky $G(x^*)\, h = 0$ pre $G(x)$ z (3) (tzv. diferencovanie väzbových podmienok v stacionárnom bode x^*).

3. Vyšetríme definitnosť vzniknutej kvadratickej formy s $n-m$ premennými.

Príklad 3

Nájdime všetky lokálne extrémy funkcie $f(x, y) = xy - x + y - 1$ vzhľadom na množinu $M : x + y = 1$. V Príklade 1 sme ukázali, že funkcia f má vzhľadom na množinu M iba jeden stacionárny bod $x^* = [-1/2, 3/2]$ s multiplikátorom $\lambda = 1/2$. Vyšetríme druhý diferenciál funkcie L v bode x^*. Platí

$$L''_{xx}(x^*, \lambda) = 0 = L''_{yy}(x^*, \lambda), \quad L''_{xy}(x^*, \lambda) = 1 = L''_{yx}(x^*, \lambda).$$

Potom $\mathrm{d}^2L(x^*, \lambda) = 2h_1h_2$, $h_1, h_2 \in \mathbb{R}^n$. Matica G v (3) má v tomto prípade v bode x^* tvar $G(x^*) = (\operatorname{grad} g(x^*)) = (1, 1)$. Vektory $h = (h_1, h_2)^T$ priestoru $\mathcal{T}_M(x^*)$ preto spĺňajú podmienku

$$(1, 1) \cdot (h_1, h_2)^T = 0 \quad \Longleftrightarrow \quad h_2 = -h_1.$$

Diferenciál $\mathrm{d}^2L(x^*, \lambda) = 2h_1h_2$ sa nám preto zredukuje na kvadratickú formu s jednou premennou h_1, konkrétne $\mathrm{d}^2L(x^*, \lambda) = -2h_1^2$. Táto forma je zrejme negatívne definitná, a preto v súlade s Vetou 2 má funkcia f v bode x^* ostré lokálne maximum s hodnotou $f(x^*) = 1/4$. Poznamenajme ešte, že pôvodný, neredukovaný diferenciál $\mathrm{d}^2L(x^*, \lambda) = 2h_1h_2$ je ako kvadratická forma s dvomi premennými indefinitný, a teda funkcia L nemá v bode x^* lokálny extrém. To ilustruje nevyhnutnosť redukcie diferenciálu $\mathrm{d}^2L(x^*, \lambda)$ vzhľadom na predpísané väzbové podmienky, ak nechceme stratiť viazané lokálne extrémy funkcie f.