

MTMM.00.202 Optimiseerimismeetodid

Evely Kirsiaed

`evely.kirsiaed@ut.ee`

sügis 2020

- 3 Mitme muutuja funktsioonide ekstreemumülesanded
- Mitme muutuja funktsioonide diferentseeruvus
 - Kitsendusteta ülesanded
 - Kitsendustega ülesanded, Lagrange'i kordajate meetod

Mitme muutuja funktsioonide diferentseeruvus

Olgu $\Omega \subset \mathbb{R}^n$ lahtine hulk (erijuhul $\Omega = \mathbb{R}^n$), $f : \Omega \rightarrow \mathbb{R}$.

Definitsioon

Öeldakse, et f on **diferentseeruv** punktis $x \in \Omega$, kui eksisteerib vektor $f'(x) \in \mathbb{R}^n$ (üldisemalt lineaarne operaator $f'(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$) nii, et

$$f(x+h) - f(x) = (f'(x), h) + \alpha(x; h),$$

kus $\frac{\alpha(x; h)}{\|h\|} \rightarrow 0$ ehk $\alpha(x; h) = o(\|h\|)$ protsessis $\|h\| \rightarrow 0$. Seejuures $h \in \mathbb{R}^n \setminus \{0\}$ on selline, et $x+h \in \Omega$.

Kui f on diferentseeruv punktis x , siis

$$f'(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right).$$

Kasutatakse tähistust $f'(x) = \text{grad } f(x) = \nabla f(x)$.

Mitme muutuja funktsioonide diferentseeruvus

Näide 1

Uurime funktsiooni $f : \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$f(x, y) = \begin{cases} \frac{y^3}{x^2 + y^2}, & (x, y) \neq (0, 0), \\ 0, & (x, y) = (0, 0), \end{cases}$$

diferentseeruvust punktis $(0, 0)$.

Funktsioon on pidev punktis $(0, 0)$, sest

$$\lim_{(x,y) \rightarrow (0,0)} f(x, y) = \lim_{r \rightarrow 0} \frac{r^3 \sin^3 \varphi}{r^2} = 0 \quad (x = r \cos \varphi, y = r \sin \varphi).$$

Funktsioonil leiduvad lõplikud osatuletised punktis $(0, 0)$

$$\begin{aligned} \frac{\partial f}{\partial x}(0, 0) &= \lim_{h \rightarrow 0} \frac{f(0 + h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0 - 0}{h} = 0, \\ \frac{\partial f}{\partial y}(0, 0) &= \lim_{h \rightarrow 0} \frac{f(0, 0 + h) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{h - 0}{h} = 1. \end{aligned}$$

Mitme muutuja funktsioonide diferentseeruvus

Näide 1 (jätkub)

Selleks, et f oleks diferentseeruv punktis 0 , peab leiduma vektor (M, N) nii, et

$$\underbrace{f(0+h) - f(0)}_{= \frac{h_2^3}{h_1^2 + h_2^2}} = Mh_1 + Nh_2 + \alpha(x; h),$$

kus $\alpha(x; h) = o(\|h\|)$, kui $\|h\| \rightarrow 0$.

Olgu $h_1 = 0$, $h_2 > 0$, $h_2 \rightarrow 0$, siis $\alpha(x; h) = h_2 - Nh_2 = o(\|h\|) \Rightarrow N = 1$.

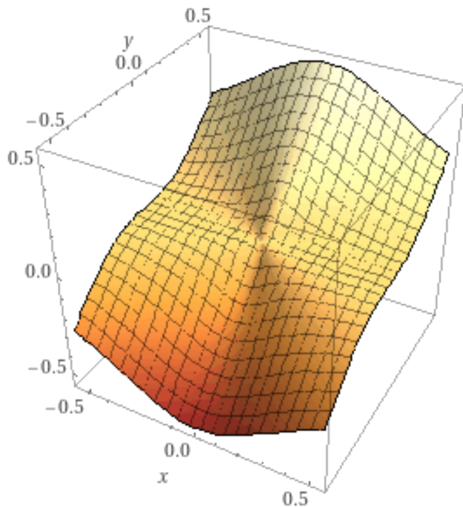
Olgu $h_2 = 0$, $h_1 > 0$, $h_1 \rightarrow 0$, siis $\alpha(x; h) = -Mh_1 = o(\|h\|) \Rightarrow M = 0$.

(Kui funktsioon on diferentseeruv, siis on tuletisvektori komponentideks osatuletised.)

Olgu $h_1 = h_2$, siis $\alpha(x; h) = \frac{1}{2}h_2 - (M + N)h_2 = o(\|h\|)$ parajasti siis, kui $M + N = \frac{1}{2}$, mis on vastuolus eelnevaga. Funktsioon f ei ole diferentseeruv punktis $(0, 0)$. (Pidevus ja lõplike osatuletiste olemasolu on küll tarvilik diferentseeruvuseks, kuid ei ole piisav.)

Mitme muutuja funktsioonide diferentseeruvus

Näide 1 (jätkub)



Mitme muutuja funktsioonide diferentseeruvus

Kui $f : \Omega \rightarrow \mathbb{R}$ on diferentseeruv igas punktis $x \in \Omega$, siis võib vaadelda funktsiooni $f' : \Omega \rightarrow \mathbb{R}^n$, kus elemendile $x \in \Omega$ seatakse vastavusse $f'(x) \in \mathbb{R}^n$. Funktsioon f' on diferentseeruv punktis $x \in \Omega$, kui eksisteerib maatriks $f''(x)$ (üldisemalt lineaarne operaator $f''(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$) nii, et

$$f'(x+h) - f'(x) = f''(x)h + \beta(x; h),$$

kus $\frac{\beta(x;h)}{\|h\|} \rightarrow 0$ protsessis $\|h\| \rightarrow 0$.

Näide 2

Vaatleme funktsiooni $f : \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$f(x, y) = 2x^4 - 3x^2y + y^2 = (y - x^2)(y - 2x^2).$$

Funktsioon f on kõikjal diferentseeruv, sest tal leiduvad pidevad osatuletised $f_x(x, y) = 8x^3 - 6xy$ ja $f_y(x, y) = -3x^2 + 2y$. (Pidevate osatuletiste olemasolu on piisav diferentseeruvuseks.)

Mitme muutuja funktsioonide diferentseeruvus

Näide 2 (jätkub)

Funktsiooni $f' : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $f'(x, y) = (8x^3 - 6xy; -3x^2 + 2y)$ diferentseeruvuse näitamiseks kasutame definitsiooni. Leidub lineaarne operaator $f''(x, y) \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ nii, et

$$f'(x + h_1, y + h_2) - f'(x, y) = f''(x, y)h + \beta((x, y); h),$$

kus $\beta((x, y); h) = o(\|h\|)$. Operaatori $f''(x, y)$ rolli sobib Hesse maatriks

$$f''(x, y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 24x^2 - 6y & -6x \\ -6x & 2 \end{pmatrix},$$

sest siis

$$f'(x + h_1, y + h_2) - f'(x, y) - f''(x, y)h = \dots$$

Mitme muutuja funktsioonide diferentseeruvus

Näide 2 (jätkub)

$$\begin{aligned} f'(x + h_1, y + h_2) - f'(x, y) - f''(x, y)h &= \\ &= (8(x + h_1)^3 - 6(x + h_1)(y + h_2); -3(x + h_1)^2 + 2(y + h_2)) \\ &\quad - (8x^3 - 6xy; -3x^2 + 2y) \\ &\quad - (24x^2h_1 - 6yh_1 - 6xh_2; -6xh_1 + 2h_2) \\ &= (24xh_1^2 + 8h_1^3 - 6h_1h_2; -3h_1^2) =: \beta((x, y); h). \end{aligned}$$

Saab näidata, et $\frac{\beta((x,y);h)}{\|h\|} \rightarrow 0$ protsessis $\|h\| \rightarrow 0$ (näiteks minnes üle polaarkoordinaatidele).

Diferentseeruvust oleks saanud näidata ka osatuletiste pidevuse abil: f' on diferentseeruv, sest tema komponendid on diferentseeruvad (viimane järelneb sellest, et neil leiduvad pidevad osatuletised).

Kitsendusteta ülesanded

Näide 3

Vaatleme funktsiooni $f : \mathbb{R}^2 \rightarrow \mathbb{R}$

$$f(x, y) = 2x^4 - 3x^2y + y^2 = (y - x^2)(y - 2x^2).$$

ning ülesannet

$$\min_{(x,y) \in \mathbb{R}^2} f(x, y).$$

Lähenedes punktile $(0, 0)$ mööda sirget $x = 0$ või $y = 0$ on tegemist miinimumpunktiga, sest $f(0, 0) = 0$, kuid $f(0, y) = y^2 > 0$ iga $y \neq 0$ korral ning $f(x, 0) = 2x^4 > 0$ iga $x \neq 0$ korral.

Lähenedes punktile $(0, 0)$ mööda sirget $y = kx$, $k \in \mathbb{R} \setminus \{0\}$, näeme, et $f(x, kx) = x^2(k - x)(k - 2x) > 0$, kui $0 < |x| < \frac{|k|}{2}$, st x on küllalt lähedal nullile.

Kas $(0, 0)$ võiks olla funktsiooni f lokaalne miinimumpunkt?

Kitsendusteta ülesanded

Näide 3 (jätkub)

Vastus on ei, sest punkti $(0,0)$ igas ε -ümbruses leidub punkte, kus $f(x,y) < 0$.

Olgu $0 < \varepsilon < 1$, valime $x < \frac{\varepsilon}{2}$ ning $y = \frac{3}{2}x^2$, siis $x^2 + y^2 < \frac{\varepsilon^2}{4} + \frac{9\varepsilon^4}{4 \cdot 2^4} < \varepsilon^2$, st $\sqrt{x^2 + y^2} < \varepsilon$ ning

$$f\left(x, \frac{3}{2}x^2\right) = \left(\frac{3}{2}x^2 - x^2\right)\left(\frac{3}{2}x^2 - 2x^2\right) = -\frac{1}{4}x^4 < 0, \quad x \neq 0.$$

Kitsendusteta ülesanded

Lause

Kui funktsioonil f on punktis $x \in \Omega$ lokaalne miinimum või lokaalne maksimum ja f on punktis x diferentseeruv, siis $f'(x) = 0$ (ehk $\frac{\partial f}{\partial x_i}(x) = 0, i = 1, \dots, n$).

Lause

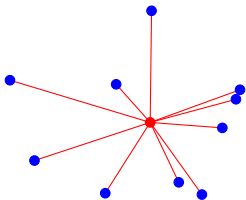
Kui f on diferentseeruv punkti $x \in \Omega$ ümbruses, f' on diferentseeruv punktis x , $f'(x) = 0$ ja $(f''(x)h, h) > 0$ iga $h \in \mathbb{R}^n \setminus \{0\}$ korral, siis funktsioonil f on punktis x lokaalne miinimum. Kui lisaks diferentseeruvuse eeldustele $f'(x) = 0$ ja $(f''(x)h, h) < 0$ iga $h \in \mathbb{R}^n \setminus \{0\}$ korral, siis funktsioonil f on punktis x lokaalne maksimum.

Näite 3 korral $f'(0) = 0$, kuid $f''(0) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$ ning $(f''(x)h, h) = 2h_2^2 \geq 0$ iga $h \in \mathbb{R}^n \setminus \{0\}$ korral.

Kitsendusteta ülesanded

Näide 4

Antud on $x^i \in \mathbb{R}^n$, $i = 1, \dots, m$. Leida $x \in \mathbb{R}^n$ nii, et $\sum_{i=1}^m \|x - x^i\|^2$ oleks minimaalne. Siin $\|x\| = \|x\|_2 = \sqrt{\sum_{k=1}^n x_k^2} = \sqrt{(x, x)}$.



Vt funktsiooni

$$f(x) = \sum_{i=1}^m \|x - x^i\|^2, x \in \mathbb{R}^n.$$

Leiame $f(x + h) - f(x)$.

Kitsendusteta ülesanded

Näide 4 (jätkub)

$$\begin{aligned}f(x + h) - f(x) &= \sum_{i=1}^m \|x + h - x^i\|^2 - \sum_{i=1}^m \|x - x^i\|^2 \\&= \sum_{i=1}^m (x - x^i + h, x - x^i + h) - \sum_{i=1}^m \|x - x^i\|^2 \\&= \sum_{i=1}^m (\|x - x^i\|^2 + 2(x - x^i, h) + \|h\|^2) - \sum_{i=1}^m \|x - x^i\|^2 \\&= \left(2 \sum_{i=1}^m (x - x^i), h \right) + m\|h\|^2.\end{aligned}$$

Kitsendusteta ülesanded

Näide 4 (jätkub)

Võrdusest $f(x+h) - f(x) = \left(2 \sum_{i=1}^m (x - x^i), h\right) + m\|h\|^2$ saame, et

$$f'(x) = 2 \sum_{i=1}^m (x - x^i), x \in \mathbb{R}^n,$$

$$\alpha(x; h) = m\|h\|^2 = o(\|h\|).$$

Ekstreemumi olemasoluks on tarvilik, et

$$f'(x) = 0 \Leftrightarrow \sum_{i=1}^m (x - x^i) = 0 \Leftrightarrow mx - \sum_{i=1}^m x^i = 0 \Leftrightarrow x = \frac{1}{m} \sum_{i=1}^m x^i =: x^0.$$

Kitsendusteta ülesanded

Näide 4 (jätkub)

Leiame ka teise tuletise. Et

$$f'(x+h) - f'(x) = 2 \sum_{i=1}^m (x+h-x^i) - 2 \sum_{i=1}^m (x-x^i) = 2mh$$

siis $f''(x) = 2mI$ ja $\beta(x; h) = 0$. Seega $(f''(x)h, h) = 2m\|h\|^2 > 0$ iga $h \in \mathbb{R}^n \setminus \{0\}$ korral. Seega x^0 on lokaalne miinimumpunkt.

Tegelikult on x^0 globaalne miinimumpunkt. Taylori arendise abil ($f''' = 0$)

$$\begin{aligned} f(x^0 + h) &= f(x^0) + (f'(x^0), h) + \frac{1}{2}(f''(x^0)h, h) \\ &= f(x^0) + \frac{1}{2}(f''(x^0)h, h) > f(x^0) \quad \forall h \in \mathbb{R}^n \setminus \{0\}. \end{aligned}$$

Lagrange'i kordajate meetod

Olgu $f: \mathbb{R}^n \rightarrow \mathbb{R}$. Vaatleme ülesannet

$$\min_{x \in \Omega} f(x) \quad \text{või} \quad \max_{x \in \Omega} f(x), \quad (1)$$

kus

$$\Omega = \{x \in \mathbb{R}^n \mid g_j(x) = 0, j = 1, \dots, m\}, \quad g_j: \mathbb{R}^n \rightarrow \mathbb{R}, j = 1, \dots, m.$$

Moodustame funktsiooni

$$\mathcal{L}(x, \lambda) = \lambda_0 f(x) + \sum_{j=1}^m \lambda_j g_j(x),$$

kus $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_m)$. Funktsiooni $\mathcal{L}: \mathbb{R}^{n+m+1} \rightarrow \mathbb{R}$ nimetatakse **Lagrange'i funktsiooniks**. Vektori λ komponente nimetatakse **Lagrange'i kordajateks** (λ on Lagrange'i kordajate vektor).

Lagrange'i kordajate meetod

Teoreem 1

Olgu $x^ \in \mathbb{R}^n$ ülesande (1) lahend, $f, g_j, j = 1, \dots, m$, pidevalt diferentseeruvad x^* ümbruses. Siis leidub $\lambda^* \neq 0$ nii, et*

$$\frac{\partial \mathcal{L}}{\partial x_i}(x^*, \lambda^*) = 0, \quad i = 1, \dots, n \quad (\text{lühidalt } \mathcal{L}_x(x^*, \lambda^*) = 0). \quad (2)$$

Lahendi (x^*, λ^*) otsimiseks (kokku $n + m + 1$ komponenti) on meil $n + m$ võrrandit

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x_i}(x, \lambda) = 0, & i = 1, \dots, n, \\ g_j(x) = 0 \quad \left(\Leftrightarrow \frac{\partial \mathcal{L}}{\partial \lambda_j}(x, \lambda) = 0 \right), & j = 1, \dots, m. \end{cases} \quad (3)$$

Lagrange'i kordajate meetod

Kui (x^*, λ^*) on süsteemi (3) lahend, kusjuures $\lambda^* \neq 0$, siis on sobivaks lahendiks ka $(x^*, c\lambda^*)$, $c \neq 0$. Lahendi võimaliku paljususe vähendamiseks lisatakse tavaliselt veel mingi normeerimistingimus, näiteks $\|\lambda\| = 1$ (levinuim 2-normis, s.t. $\sum_{j=0}^m \lambda_j^2 = 1$). Kui on teada, et lahendi korral $\lambda_0^* \neq 0$, siis on üks loomulik lisatingimus $\lambda_0^* = 1$.

Süsteemi (3) lahendamise tulemusena saab leida punktid x , mille hulgas on teoreemi 1 põhjal ülesande (1) kõik lahendid.

Lagrange'i kordajate meetod

Teoreem 2

Eeldame, et

- ① *f ja g_j , $j = 1, \dots, m$, on kaks korda diferentseeruvad punktis x^* ;*
- ② *(x^*, λ^*) rahuldab süsteemi (3), kusjuures $\lambda_0^* \geq 0$;*
- ③ *kehtib $(\mathcal{L}_{xx}(x^*, \lambda^*)h, h) > 0$ iga $h \in \mathbb{R}^n \setminus \{0\}$ korral, mis rahuldab tingimusi*

$$(f'(x^*), h) \leq 0, \quad (g'_j(x^*), h) = 0, \quad j = 1, \dots, m. \quad (4)$$

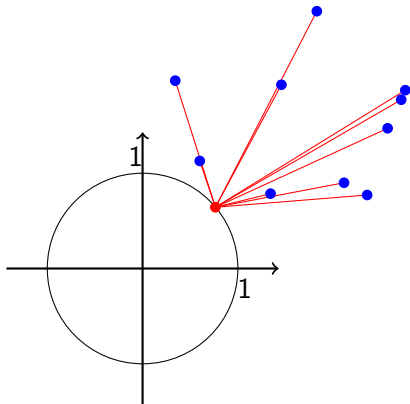
Siis x^ on miinimumülesande (1) lokaalne lahend ja ta on lokaalselt ühene miinimumpunkt.*

Siin $\mathcal{L}_{xx}(x, \lambda) = \left(\frac{\partial^2 \mathcal{L}}{\partial x_i \partial x_j}(x, \lambda) \right)_{i,j=1}^n$ on sümmeetriline $n \times n$ maatriks.

Lagrange'i kordajate meetod

Näide 5

Antud on punktid $x^i \in \mathbb{R}^n$, $i = 1, \dots, m$. Leida $x \in \Omega = \{x \in \mathbb{R}^n : \|x\| = 1\}$ nii, et $\sum_{i=1}^m \|x - x^i\|^2$ oleks minimaalne (kasutame 2-normi).



Lagrange'i kordajate meetod

Näide 5

Minimiseerime funktsiooni $f(x) = \sum_{i=1}^m \|x - x^i\|^2$ tingimusel $g_1(x) := (x, x) - 1 = 0$. Moodustame Lagrange'i funktsiooni

$$L(x, \lambda) = \lambda_0 f(x) + \lambda_1 ((x, x) - 1), \quad x \in \mathbb{R}^n, \lambda_0, \lambda_1 \in \mathbb{R}.$$

Näite 4 põhjal $f'(x) = 2 \sum_{i=1}^m (x - x^i)$ ning $g_1'(x) = 2x$. Süsteem (3) on kujul

$$\begin{cases} \mathcal{L}_x(x, \lambda) = 2\lambda_0 \sum_{i=1}^m (x - x^i) + 2\lambda_1 x = 0, \\ (x, x) - 1 = 0. \end{cases} \quad (6)$$

Kasutades tähistust $x^0 := \frac{1}{m} \sum_{i=1}^m x^i$ saame (6) esitada kujul

$$\lambda_0 m(x - x^0) + \lambda_1 x = 0, \quad \|x\| = 1.$$

Lagrange'i kordajate meetod

Näide 5 (jätkub)

Süsteemi

$$\lambda_0 m(x - x^0) + \lambda_1 x = 0, \quad \|x\| = 1,$$

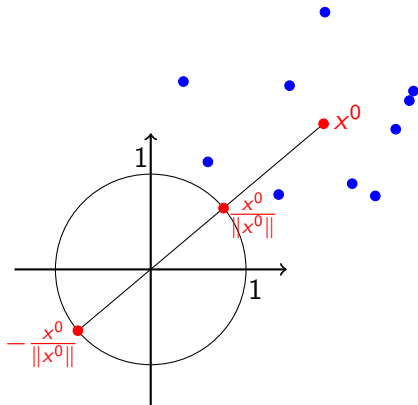
korral, kui $\lambda_0 = 0$, siis ka $\lambda_1 = 0$, st süsteemil ei oleks lahendit (x, λ) , kus $\lambda \neq 0$. Seetõttu võtame $\lambda_0 = 1$ ning saame süsteemi kujul

$$(m + \lambda_1)x = mx^0, \quad \|x\| = 1. \quad (8)$$

Lisaks $\lambda_0 = 1$ tõttu $\mathcal{L}_{xx} = 2mI + 2\lambda_1 I = 2(m + \lambda_1)I$. Tingimustest (8) esimene ütleb, et x ja x^0 on lineaarselt sõltuvad. Kui $x^0 \neq 0$, siis on $\|x\| = 1$ tõttu kaks võimalust: a) $x = \frac{x^0}{\|x^0\|}$ või b) $x = -\frac{x^0}{\|x^0\|}$. Juhul a) $m + \lambda_1 = m\|x^0\|$ ning $\mathcal{L}_{xx} = 2m\|x^0\|I$, mis on positiivselt määratud maatriks, seega $x = \frac{x^0}{\|x^0\|}$ on (lokaalne) miinimumpunkt. Juhul b) $m + \lambda_1 = -m\|x^0\|$ ning $\mathcal{L}_{xx} = -2m\|x^0\|I$, mis on negatiivselt määratud maatriks, seega $x = -\frac{x^0}{\|x^0\|}$ on (lokaalne) maksimumpunkt.

Lagrange'i kordajate meetod

Näide 5 (jätkub)



Lagrange'i kordajate meetod

Näide 5 (jätkub)

Kui $x^0 = 0$, siis (8) põhjal $\lambda_1 = -m$ ja x võib olla suvaline $\mathbb{R}^n$ element, mille korral $\|x\| = 1$, sest

$$\begin{aligned} f(x) &= \sum_{i=1}^m \|x - x^i\|^2 = \sum_{i=1}^m (x - x^i, x - x^i) \\ &= \sum_{i=1}^m (\|x\|^2 - 2(x, x^i) + \|x^i\|^2) \\ &= m - 2 \left(x, \underbrace{\sum_{i=1}^m x^i}_{=0} \right) + \sum_{i=1}^m \|x^i\|^2 = \text{const.} \end{aligned}$$