

MTMM.00.202 Optimiseerimismeetodid

Evely Kirsiaed

`evely.kirsiaed@ut.ee`

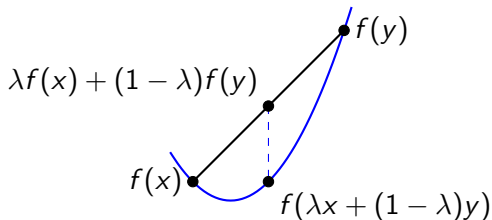
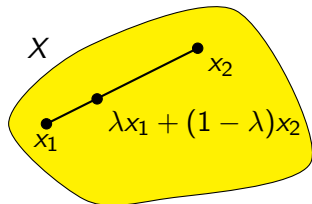
sügis 2020

6 Kumer planeerimine

- Kumerad funktsioonid
- Kumerate hulkade eraldusteoreemid
- Kumera planeerimise põhiteoreem

Kumerad funktsioonid

Olgu E vektorruum, $X \subset E$ kumer, st $x_1, x_2 \in X$, $x_1 \neq x_2$, $\lambda \in (0, 1)$ korral $\lambda x_1 + (1 - \lambda)x_2 \in X$.



Definitsioon

Funktsiooni $f: X \rightarrow \mathbb{R}$ nimetatakse kumeraks, kui iga $x, y \in X$, $x \neq y$, ja iga $\lambda \in (0, 1)$ korral

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Lause 1 (ül 34)

Kui $f, g: X \rightarrow \mathbb{R}$ on kumerad, siis $f + g$ ja cf , $c \geq 0$, on kumerad.

Lause 2

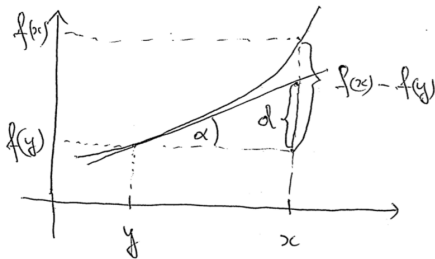
Funktsioon $f: X \rightarrow \mathbb{R}$ on kumer parajasti siis, kui iga $x, y \in X$, $x \neq y$, korral funktsioon $g(t) = f(tx + (1 - t)y)$, $t \in [0, 1]$, on kumer.

Kumerad funktsioonid

Lause 3

Kui $X \subset \mathbb{R}^n$ on kumer, siis diferentseeruv funktsioon $f: X \rightarrow \mathbb{R}$ on kumer parajasti siis, kui iga $x, y \in X$, $x \neq y$, korral

$$f'(y) \cdot (x - y) \leq f(x) - f(y). \quad (1)$$



Juhul $n = 1$ $f'(y) = \tan \alpha = \frac{d}{x-y}$, seega $f'(y) \cdot (x - y) = d \leq f(x) - f(y)$.

Lause (Ül* 10)

Kui $X \subset \mathbb{R}^n$ on kumer, siis diferentseeruv funktsioon $f: X \rightarrow \mathbb{R}$ on kumer parajasti siis, kui iga $x, y \in X$, $x \neq y$, korral

$$(f'(x) - f'(y)) \cdot (x - y) \geq 0, \quad (2)$$

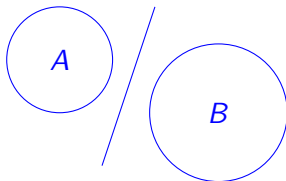
st funktsioon $f': X \rightarrow \mathbb{R}^n$ on monotoonne.

Lause 4

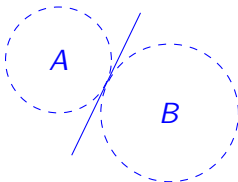
Kui $X \subset \mathbb{R}^n$ on kumer, lahtine ja $f: X \rightarrow \mathbb{R}$ on kaks korda diferentseeruv, siis f on kumer parajasti siis, kui iga $x \in X$, korral $f''(x)$ on positiivne, st $(f''(x)h, h) \geq 0$ iga $h \in \mathbb{R}^n$ korral.

Kumerate hulkade eraldusteoreemid

Põhiidee. Olgu $A, B \subset \mathbb{R}^2$ kumerad, kinnised, $A \cap B = \emptyset$. Siis saab A ja B eraldada sirgega.



Kui A ja B ei ole kinnised (on kumerad ja lõikumatud), siis saab nad samuti eraldada sirgega.



Kumerate hulkade eraldusteoreemid

Kasutame tähiseid $\overline{X}$ – hulga X sulund, X° – hulga X sisepunktide hulk, ∂X – hulga X raja.

Lause (Ül 37)

Normeeritud ruumi kumera osahulga sulund on kumer.

Teoreem 1

Olgu $X \subset \mathbb{R}^n$ kumer ning $y \notin \overline{X}$. Siis leidub $a \in \mathbb{R}^n$ nii, et

$$\inf_{x \in X} a \cdot x > a \cdot y.$$

Fikseeritud $a \in \mathbb{R}^n \setminus \{0\}$ ja $c \in \mathbb{R}$ korral on hulk $\{x \in \mathbb{R}^n \mid a \cdot x = c\}$ hüpertasand, $\{x \in \mathbb{R}^n \mid a \cdot x > c\}$ poolruum. Teoreem 1 väidab, et X asub hüpertasandist $\{x \in \mathbb{R}^n \mid a \cdot x = a \cdot y\}$ rangelt ühel pool.

Kumerate hulkade eraldusteoreemid

Tõestuse skeem:

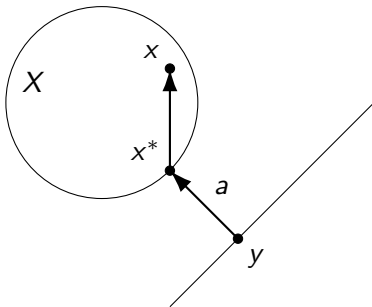
$$1) \exists x^* \in \overline{X}: \|x^* - y\| = \inf_{x \in X} \|x - y\|$$

$$2) x^* \neq y \rightarrow \|x^* - y\| > 0$$

$$3) \forall x \in X \quad (x^* - y) \cdot (x - x^*) \geq 0 \text{ ehk } a \cdot x \geq a \cdot x^*, \text{ kus } a = x^* - y$$

$$4) a \cdot x^* > a \cdot y$$

$$\Rightarrow \inf_{x \in X} a \cdot x > a \cdot y$$



Ül* 11. Tõestada, et kui $X \subset \mathbb{R}^n$ on kumer ja $y \in \partial X$, siis $y \in \partial \overline{X}$. Leida näide, kus mittekumera hulga X korral väide ei kehti.

Järeldus

Olgu $X \subset \mathbb{R}^n$ kumer ning $y \notin X^\circ$. Siis leidub $a \in \mathbb{R}^n$, $a \neq 0$, nii, et

$$\inf_{x \in X} a \cdot x \geq a \cdot y.$$

Tõestuse skeem:

$$y \notin X^\circ \Rightarrow y \in \partial X \vee y \notin \overline{X}$$

Juhul $y \notin \overline{X}$ rakendub Teoreem 1, juhul $y \in \partial X$ ül* 11 põhjal $y \in \partial \overline{X}$

$\exists y_k \notin \overline{X} : y_k \rightarrow y$, kasutame (a_k) tõkestatust ja skalaarkorrutise pidevust

Teoreem 2

Olgu hulgad $X, Y \subset \mathbb{R}^n$ kumerad, $X^\circ \neq \emptyset$, $Y^\circ \neq \emptyset$, $X^\circ \cap Y^\circ = \emptyset$. Siis leidub $a \in \mathbb{R}^n$, $a \neq 0$ nii, et

$$a \cdot x \geq a \cdot y \quad \forall x \in X, \forall y \in Y.$$

Järeldus

Teoreemi 2 eeldustel kehtib

$$a \cdot x \geq a \cdot y \quad \forall x \in \overline{X}, \forall y \in \overline{Y}$$

ehk

$$\inf_{x \in \overline{X}} a \cdot x \geq \sup_{y \in \overline{Y}} a \cdot y,$$

mis tähendab, et eraldatakse (mitterangelt) hulkade X ja Y sulundid.

Teoreemi 2 tõestuse skeem:

$$Z := X - Y = \{x - y \mid x \in X, y \in Y\}$$

1) Z on kumer

2) $0 \notin Z^\circ \Leftarrow 0 \notin Z$

a) Lemma: hulga sisepunkti ja suvalist punkti ühendava lõigu "sisemised" punktid kuuluvad hulga sisemusse

b) Kui $0 \in Z$, siis leidub kuitahes lähedal punkt $y_1 - x_1 \notin Z$

3) $\exists a \in \mathbb{R}^n, a \neq 0 : a \cdot (x - y) \geq a \cdot 0 = 0 \quad \forall x \in X, \forall y \in Y$

Olgu X hulk, sihifunktsioon $f: X \rightarrow \mathbb{R}$ ning tõkkefunktsioonid $f_1, \dots, f_m: X \rightarrow \mathbb{R}$. Tõkkefunktsioonina võime vaadleda ka funktsiooni $F: X \rightarrow \mathbb{R}^m$, $F(x) = (f_1(x), \dots, f_m(x))$.

Üldülesanne:

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

kus $\Omega = \{x \in X \mid F(x) \leq 0\}$.

Põhiülesanne:

$$\min_{x \in \Omega} f(x), \quad (2)$$

kus $f, f_1, \dots, f_m: \mathbb{R}^n \rightarrow \mathbb{R}$, $\Omega = \{x \in \mathbb{R}^n \mid F(x) \leq 0, x \geq 0\}$.

Põhiülesanne on erijuht üldülesandest, kui võtta $X = \{x \in \mathbb{R}^n \mid x \geq 0\} = \mathbb{R}_+^n$.

Näide

Valime ülesandes (2) $f(x) = -c \cdot x$ ja $F(x) = Ax - b$, saame põhikujul oleva lineaarse planeerimise ülesande. Seega põhikujul lineaarse planeerimise ülesanne on erijuht põhiülesandest (2).

Definitsioon

Üldülesande (1) **Lagrange'i funktsiooniks** nimetatakse funktsiooni $\mathcal{L}: X \times \mathbb{R}_+^m \rightarrow \mathbb{R}$,

$$\mathcal{L}(x, y) = f(x) + \sum_{i=1}^m y_i f_i(x) = f(x) + y \cdot F(x), \quad x \in X, y \in \mathbb{R}_+^m.$$

Sadulpunkt

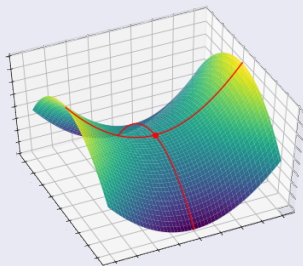
Definitsioon

Punkti $(x^*, y^*) \in X \times \mathbb{R}_+^m$ nimetatakse funktsiooni $\mathcal{L}$ sadulpunktiks, kui

$$\mathcal{L}(x^*, y) \leq \mathcal{L}(x^*, y^*) \leq \mathcal{L}(x, y^*) \quad \forall x \in X, \forall y \in \mathbb{R}_+^m.$$

Näide

Vaatleme juhtu, kus $X = \mathbb{R}$, $Y = \mathbb{R}_+$. Sadulpunkti tingimus nõuab funktsioonilt teatud käitumist ristsirgetel (x^*, y) ja (x, y^*) , mujal mingeid piiranguid ei ole.



Teoreem

Kui (x^, y^*) on üldülesande (1) Lagrange'i funktsiooni sadulpunkt, siis x^* on ülesande (1) lahend.*

Kas iga ülesande (1) saab taandada tema Lagrange'i funktsiooni sadulpunkti leidmisele? St, kui ülesandel (1) on olemas lahend x^* , kas siis leidub $y^* \geq 0$ nii, et (x^*, y^*) on ülesande (1) Lagrange'i funktsiooni sadulpunkt?

Näide

Vaatleme ülesannet

$$\min \left\{ - \sum_{i=1}^n x_i \mid \sum_{i=1}^n x_i^2 \leq 0, x \geq 0 \right\}.$$

Tegemist on põhiülesandega (2), kus $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f(x) = -\sum_{i=1}^n x_i$, $m = 1$, $f_1: \mathbb{R}^n \rightarrow \mathbb{R}$, $f_1(x) = \sum_{i=1}^n x_i^2$. Siin $\Omega = \{0\}$ ja $x^* = 0$ on optimaalne lahend. Ülesande Lagrange'i funktsioon on

$$\mathcal{L}(x, y) = - \sum_{i=1}^n x_i + y \sum_{i=1}^n x_i^2, \quad x \in \mathbb{R}_+^n, y \in \mathbb{R}_+.$$

Näitame, et ei leidu sadulpunkti $(0, y^*)$, kus $y^* \geq 0$.

Näide

Vasakpoolne sadulpunkti võrratus kehtib, sest iga $y \in \mathbb{R}_+$ korral $\mathcal{L}(x^*, y) = \mathcal{L}(x^*, y^*) = 0$. Parempoolne sadulpunkti võrratus

$$0 \leq \mathcal{L}(x, y^*) = - \sum_{i=1}^n x_i + y^* \sum_{i=1}^n x_i^2 \quad \forall x \in \mathbb{R}_+^n$$

ei saa kehtida juhul $y^* = 0$, sest $x \neq 0$ korral $-\sum_{i=1}^n x_i < 0$. Kui $y^* > 0$, siis valime ε , $0 < \varepsilon < \frac{1}{y^*}$, ja $x = \left(\frac{1}{y^*} - \varepsilon, 0, \dots, 0\right)$ ning

$$\mathcal{L}(x, y^*) = - \left(\frac{1}{y^*} - \varepsilon\right) + y^* \left(\frac{1}{y^*} - \varepsilon\right)^2 = \left(\frac{1}{y^*} - \varepsilon\right) (-1 + 1 - \varepsilon y^*) < 0,$$

st sadulpunkti ei leidu.

Põhikujul lineaarse planeerimise ülesannete korral sadulpunkt alati leidub (kui lineaarse planeerimise ülesanne on lahenduv) (ÜI* 12).

Definitsioon

Eeldame, et f ja f_i , $i = 1, \dots, m$, on kumerad, siis ülesannet (1) nimetatakse **kumera planeerimise ülesandeks**.

Kui funktsioonid f_i on kumerad ($X \subset E$ kumer, E vektorruum), siis ülesande (1) lubatav hulk Ω on kumer; kui funktsioonid f_i on pidevad ($X \subset E$, E normeeritud ruum), siis Ω on kinnine (ÜI 38).

Kumera planeerimise põhiteoreem

Teoreem (Kuhn-Tuckeri teoreem, kumera planeerimise põhiteoreem, 1951)

Olgu üldülesandes (1) funktsioonid $f, f_1, \dots, f_m$ kumerad (seega on hulk $X \subset E$ kumer, E vektorruum) ning eksisteerigu $\bar{x} \in X$ nii, et $F(\bar{x}) < 0$ (st $f_i(\bar{x}) < 0, i = 1, \dots, m$). Kui x^ on ülesande (1) lahend, siis on olemas $y^* \in \mathbb{R}_+^m$ nii, et (x^*, y^*) on ülesande (1) Lagrange'i funktsiooni sadulpunkt.*

Tingimust $F(\bar{x}) < 0$ mingi $\bar{x} \in X$ korral nimetatakse **Slateri** ehk **regulaarsuse tingimuseks**. Viimases näites ei ole see tingimus rahuldatud.

Kumera planeerimise põhiteoreem

Teoreem

Olgu $f, f_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, m$, kumerad ja diferentseeruvad. Siis $(x^*, y^*) \in \mathbb{R}_+^{n+m}$ on põhiülesande (2) Lagrange'i funktsiooni sadupunkt parajasti siis, kui

$$\begin{aligned} \mathcal{L}_x(x^*, y^*) &\geq 0, & \mathcal{L}_y(x^*, y^*) &\leq 0, \\ x^* \cdot \mathcal{L}_x(x^*, y^*) &= 0, & y^* \cdot \mathcal{L}_y(x^*, y^*) &= 0. \end{aligned} \tag{8}$$

Siin $\mathcal{L}_x(x, y) = \left(\frac{\partial \mathcal{L}}{\partial x_1}(x, y), \dots, \frac{\partial \mathcal{L}}{\partial x_n}(x, y) \right)$ ja $\mathcal{L}_y(x, y) = \left(\frac{\partial \mathcal{L}}{\partial y_1}(x, y), \dots, \frac{\partial \mathcal{L}}{\partial y_m}(x, y) \right) = (f_1(x), \dots, f_m(x))$.

Olgu täidetud teoreemi eeldused ning (8), siis (x^*, y^*) on $\mathcal{L}$ sadulpunkt ja x^* ülesande (2) lahend. Kui teoreemi eeldustel $x^* \in \mathbb{R}_+^n$ on selline, et (8) ei ole täidetud ühegi $y^* \in \mathbb{R}_+^m$ korral, aga regulaarsuse tingimus on täidetud, siis x^* ei ole (2) lahend.

Koonuse kriteerium

Vt ülesande (2) korral tingimusi (8). Olgu

$x \in \Omega = \{x \in \mathbb{R}^n \mid F(x) \leq 0, x \geq 0\}$, $y \in \mathbb{R}_+^m$, siis

$$\mathcal{L}_x(x, y) = f'(x) + \sum_{i=1}^m y_i f'_i(x) = z = \sum_{j=1}^n z_j e^j \geq 0 \Leftrightarrow z_j \geq 0, j = 1, \dots, n.$$

Tingimus $\mathcal{L}_y(x, y) = F(x) \leq 0$ on täidetud $x \in \Omega$ tõttu. Olgu aktiivsetele kitsendustele vastavad indeksite hulgad $M = \{i \in \{1, \dots, m\} \mid f_i(x) = 0\}$, $N = \{j \in \{1, \dots, n\} \mid x_j = 0\}$. Siis

$$y \cdot \mathcal{L}_y(x, y) = 0 \Leftrightarrow \underbrace{y}_{\geq 0} \cdot \underbrace{F(x)}_{\leq 0} = 0 \Leftrightarrow y_i f_i(x) = 0, i = 1, \dots, m \Leftrightarrow$$

$$\Leftrightarrow y_i = 0, i \notin M$$

$$x \cdot \mathcal{L}_x(x, y) = 0 \Leftrightarrow \underbrace{x}_{\geq 0} \cdot \underbrace{z}_{\geq 0} = 0 \Leftrightarrow x_j z_j = 0, j = 1, \dots, n \Leftrightarrow z_j = 0, j \notin N$$

$$(8) \quad \begin{aligned} \mathcal{L}_x(x, y) &\geq 0, & \Leftrightarrow & z \geq 0, z = \mathcal{L}_x(x, y) \\ \mathcal{L}_y(x, y) &\leq 0, & \Leftarrow & x \in \Omega \\ x \cdot \mathcal{L}_x(x, y) &= 0, & \Leftrightarrow & z_j = 0, j \notin N \\ y \cdot \mathcal{L}_y(x, y) &= 0 & \Leftrightarrow & y_i = 0, i \notin M. \end{aligned}$$

Lause

Tingimused (8) on rahuldatud $x \in \Omega$ ja $y \geq 0$ korral parajasti siis, kui

$$z \geq 0, \quad z_j = 0, j \notin N, \quad y_i = 0, i \notin M. \quad (11)$$

Koonuse kriteerium

Kirjutame võrduse $\mathcal{L}_x(x, y) = z$ kujul

$$-f'(x) = \sum_{i=1}^m y_i f'_i(x) + \sum_{j=1}^n z_j (-e^j)$$

ning tingimuste (8) täidetuse korral kujul

$$-f'(x) = \sum_{i \in M} y_i f'_i(x) + \sum_{j \in N} z_j (-e^j).$$

Teoreem (koonuse kriteerium)

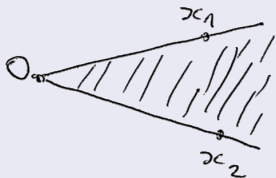
Regulaarse kumerate diferentseeruvate funktsioonidega ülesande (2) lubatav lahend x on optimaalne parajasti siis, kui mingite $\lambda \geq 0$, $\mu \geq 0$ korral

$$-f'(x) = \sum_{i \in M} \lambda_i f'_i(x) + \sum_{j \in N} \mu_j (-e^j).$$

Koonuse kriteerium

Definitsioon

Vektorruumis E nimetatakse **koonuseks** hulka $C \subset E$, mille puhul iga $x_1, x_2 \in C$ ja iga $\lambda_1, \lambda_2 \geq 0$ korral $\lambda_1 x_1 + \lambda_2 x_2 \in C$.



Hulka

$$C(x_1, \dots, x_m) = \left\{ \sum_{i=1}^m \mu_i x_i \mid \mu_i \geq 0, i = 1, \dots, m \right\}$$

nimetatakse vektorite $x_1, \dots, x_m \in E$ poolt moodustatud (**polüeedriliseks**) **koonuseks**.

Näiteks $\mathbb{R}_+^n = C(e^1, \dots, e^n)$.

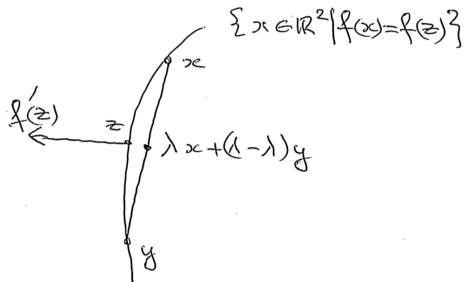
Kui $M = \emptyset$ ja $N = \emptyset$, st $f_i(x) < 0$, $i = 1, \dots, m$, ja $x_j > 0$, $j = 1, \dots, n$, siis x on Ω sisepunkt (piisab eeldada f_i pidevust, mis järeldeb diferentseeruvusest või ka kumerusest). Kui $x \in \Omega^0$, siis $M = \emptyset$ (Ül 41) ja $N = \emptyset$.

Lause

Kui kumeras ülesandes (2) $x^ \in \Omega$ korral $f_i(x^*) < 0$, $i = 1, \dots, m$, ja $x_j^* > 0$, $j = 1, \dots, n$, siis x^* on optimaalne lahend parajasti siis, kui $f'(x) = 0$.*

Koonuse kriteeriumi kasutamine

Kus asuvad tingimust $f(x) \leq 0$ rahuldavad punktid?

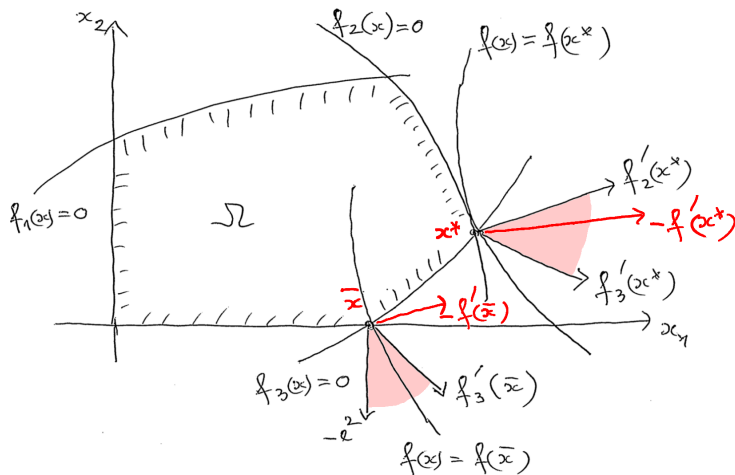


Siin

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) = f(z).$$

Tuletis $f'(z)$ näitab funktsiooni kiireima kasvu suunda.

Koonuse kriteeriumi kasutamine



Punktis x^* $M = \{2, 3\}$, $N = \emptyset$, $-f'(x^*) = \lambda_2 f'_2(x^*) + \lambda_3 f'_3(x^*)$, x^* opt.

Punktis $\bar{x}$ $M = \{3\}$, $N = \{2\}$, $-f'(\bar{x}) \notin C(f'_3(\bar{x}), -e^2)$, $\bar{x}$ ei ole opt.