Matematika 1	**Izročki: 12. in 13. teden**

Optimizacija. 3. del.
Prirejena funkcija. Karush-Kuhn-Tuckerjevi pogoji.

Polona Oblak

Naj bodo $f, g_i, h_j \colon \mathbb{R}^n \to \mathbb{R}$ dane funkcije več spremenljivk. Radi bi našli rešitev naslednjega problema

$$(P\star) \qquad \begin{aligned} &\text{minimizirajmo}_{\vec{x}}\ f(\vec{x}) \\ &\text{pri pogojih } g_i(\vec{x}) \leq 0 \text{ za } i = 1, 2, \ldots, m \\ &\qquad\qquad\quad h_j(\vec{x}) = 0 \text{ za } j = 1, 2, \ldots, r. \end{aligned}$$

Definirajmo še množice $\mathcal{D}_{g_i} = \{x \in \mathbb{R}^n;\ g_i(x) \leq 0\}$ za $i = 1, 2, \ldots, m$, $\mathcal{D}_{h_j} = \{x \in \mathbb{R}^n;\ h_j(x) = 0\}$ za $j = 1, 2, \ldots, r$, in

$$\mathcal{D} = \mathcal{D}_f \cap \left(\cap_{i=1}^m \mathcal{D}_{g_i}\right) \cap \left(\cap_{j=1}^r \mathcal{D}_{h_j}\right).$$

Sedaj lahko problem $(P\star)$ zapišemo ekvivalentno kot

$$(P\star) \qquad \min_{\vec{x}\in\mathcal{D}} f(\vec{x}).$$

Definirajmo Lagrangevo funkcijo

$$\begin{aligned} L(\vec{x}, \vec{\lambda}, \vec{\mu}) &= f(\vec{x}) - \vec{\lambda}^{\mathsf{T}} \vec{G}(\vec{x}) - \vec{\mu}^{\mathsf{T}} \vec{H}(\vec{x}) = \\ &= f(\vec{x}) - \sum_{i=1}^m \lambda_i g_i(\vec{x}) - \sum_{j=1}^r \mu_j h_j(\vec{x}), \end{aligned}$$

kjer je

$$\vec{G}(\vec{x}) = \begin{bmatrix} g_1(\vec{x}) \\ \vdots \\ g_m(\vec{x}) \end{bmatrix}, \vec{H}(\vec{x}) = \begin{bmatrix} h_1(\vec{x}) \\ \vdots \\ h_r(\vec{x}) \end{bmatrix}, \vec{\lambda} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{bmatrix}, \text{ in } \vec{\mu} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_r \end{bmatrix}.$$

Funkcijo

$$K(\vec{\lambda}, \vec{\mu}) = \inf_{\vec{x}\in\mathcal{D}} L(\vec{x}, \vec{\lambda}, \vec{\mu}) = \inf_{\vec{x}\in\mathcal{D}} \left\{ f(\vec{x}) - \vec{\lambda}^{\mathsf{T}} \vec{G}(\vec{x}) - \vec{\mu}^{\mathsf{T}} \vec{H}(\vec{x}) \right\}$$

imenujemo *prirejena funkcija* problema $(P\star)$. Pri tem spremenljivke $\vec{\lambda}$ in $\vec{\mu}$ imenujemo *prirejene spremenljivke*. Opazimo:

(1) $K(\vec{\lambda}, \vec{\mu})$ je konkavna funkcija (neodvisno od lastnosti funkcij f, g_i, h_j originalnega problema).

(2) Če je $\lambda_i \leq 0$ za $i = 1, 2, \ldots, m$, potem velja $K(\vec{\lambda}, \vec{\mu}) \leq \min_{\vec{x}\in\mathcal{D}} f(x)$ za vse $\vec{\lambda} \leq \vec{0}$ ter vse $\vec{\mu}$.

Problem

$$\begin{aligned} &\text{maksimizirajmo}_{\vec{\lambda},\vec{\mu}}\ K(\vec{\lambda},\vec{\mu}) \\ (D\star) \qquad &\text{pri pogoju}\ \ \lambda_i \leq 0 \text{ za } i = 1,2,\ldots,m \end{aligned}$$

imenujemo *prirejeni problem* problema $(P\star)$.

Označimo z $\vec{x}^*$ vektor iz $\mathcal{D}$, ki reši problem $(P\star)$ in z $\vec{\lambda}^*$ in $\vec{\mu}^*$ prirejeni spremenljivki, ki rešita prirejeni problem $(D\star)$. Naj bo torej $p^* = f(\vec{x}^*)$ rešitev problema $(P\star)$ in $d^* = K(\vec{\lambda}^*, \vec{\mu}^*)$ rešitev problema $(D\star)$. Potem iz (2) sledi

$$d^* \leq p^*.$$

Če

- je $(P\star)$ linearni program (t,.j. f je linearna in h_j so afine funkcije), ali če
- so f, g_i konveksne funkcije in h_j afine,

potem velja $d^* = p^*$.

V primeru, ko je $d^* = p^*$, sledi, da morajo $\vec{x}^*$, $\vec{\lambda}^*$ in $\vec{\mu}^*$ zadoščati *Karush–Kuhn–Tuckerjevim pogojem*:

$$\text{(KKT)} \qquad \begin{aligned} \frac{\partial L(\vec{x},\vec{\lambda}^*,\vec{\mu}^*)}{\partial \vec{x}}(\vec{x}^*) &= 0, \\ g_i(\vec{x}^*) &\leq 0 \ \text{ za } i = 1,2,\ldots,m, \\ h_j(\vec{x}^*) &= 0 \ \text{ za } j = 1,2,\ldots,r, \\ \lambda_i^* &\leq 0 \ \text{ za } i = 1,2,\ldots,m, \\ \lambda_i^* g_i(\vec{x}^*) &= 0 \ \text{ za } i = 1,2,\ldots,m. \end{aligned}$$

1. Nadaljnje branje

(1) Wilhelm Forst, Dieter Hoffmann: Optimization - Theory in Practice, Springer, 2010.

2. Domača naloga

ϟ (1) Območje A v $\mathbb{R}^n$ je konveksno, če je za poljubni točki $x, y \in A$ tudi točka $tx + (1-t)y \in A$. Pokažite, da je presek konveksnih množic konveksna množica.

ϟ (2) Naj bodo $g_i \colon \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$, $i = 1,2,\ldots,m$, konveksne funkcije na konveksni množici $\mathcal{D}$. Pokažite, da je tedaj območje

$$\mathcal{D}_0 = \{x \in \mathcal{D};\ g_i(x) \leq 0 \text{ za } i = 1,2,\ldots,m\}$$

konveksna množica.

↯ (3) Rešite naslednji problem:

$$\begin{aligned} \text{minimizirajte}_{x,y} \quad & (x-2)^2 + 2(y-1)^2 \\ \text{pri pogojih} \quad & x + 4y \leq 3, \\ & y - x \leq 0. \end{aligned}$$

↯ (4) Zapišite Karush–Kuhn–Tuckerjeve pogoje naslednjih problemov:

(a)

$$\begin{aligned} \text{minimizirajte}_{x,y,z} \quad & x + y^2 + z^2 \\ \text{pri pogojih} \quad & x^2 + 2y^2 \leq 4 \\ & x + y + z = 1. \end{aligned}$$

(b)

$$\begin{aligned} \text{minimizirajte}_{x,y,z} \quad & xyz \\ \text{pri pogojih} \quad & x^2 + 2y^2 \leq 4 \\ & x^2 + y^2 + z^2 = 3 \end{aligned}$$

(c)

$$\begin{aligned} \text{minimizirajte}_{x,y,z} \quad & 2x + 2y^2 + xz \\ \text{pri pogojih} \quad & x^2 + y^2 = 1 \\ & x + z = 0 \\ & xy \geq 0 \end{aligned}$$

(Naloge, označene s ↯ preverjajo razumevanje osnovnih pojmov in so primeri nalog s teoretičnih izpitov.)