na vrhu

Optimizacijske metode
5. Reševanje

Vladimir Batagelj

Univerza v Ljubljani
FMF, matematika

Kazalo

Minimum funkcij na $\mathbb{R}^n$

Naj bo $P(x), x \in \mathbb{R}^n$ v okolici točke x^* vsaj dvakrat zvezno odvedljiva. Razvijmo jo v Taylorjevo vrsto okoli x^*

$$P(x^* + \varepsilon y) = P(x^*) + \varepsilon y^T \nabla P(x^*) + \frac{1}{2}\varepsilon^2 y^T H(P, x^*) y + \varepsilon^2 \mathcal{O}(\varepsilon y)$$

pri čemer je $\varepsilon > 0$ in $y \in \mathbb{R}^n$.

Kdaj ima funkcija $P(x)$ v točki x^* minimum? V neki dovolj majhni okolici točke x^* funkcija $P(x)$ nima vrednosti manjše od $P(x^*)$.

Recimo, da bi obstajal tak y, da je

$$y^T \nabla P(x^*) < 0$$

Tedaj, ker ε^2 gre hitreje proti 0 kot ε, je mogoče najti $\delta > 0$ tako, da za vsak $\varepsilon, 0 < \varepsilon \leq \delta$ velja

$$P(x^* + \varepsilon y) < P(x^*)$$

To pa pomeni, da točka x^* ni minimum – protislovje.

. . . Minimum funkcij na $\mathbb{R}^n$

Torej je potreben pogoj za to, da je v točki x^* minimum

$$\forall y : y^T \nabla P(x^*) \geq 0$$

Toda, ker mora ta neenakost veljati tudi za vektorja y in $-y$, je to mogoče le, če je $\nabla P(x^*) = 0$. Podoben sklep velja tudi za maksimum.

Točkam, v katerih je $\nabla P(x) = 0$ pravimo *stacionarne* točke.

Naj bo x^* stacionarna točka. Tedaj je

$$P(x^* + \varepsilon y) = P(x^*) + \frac{1}{2}\varepsilon^2 y^T H(P, x^*) y + \varepsilon^2 \mathcal{O}(\varepsilon y)$$

Če bi bil $y^T H(P, x^*) y < 0$ za kak $y \in \mathbb{R}^n$, bi zopet veljalo za dovolj majhne $\varepsilon > 0$

$$P(x^* + \varepsilon y) < P(x^*)$$

kar pomeni, da točka x^* ni minimum – pritislovje.

. . . Minimum funkcij na $\mathbb{R}^n$

Torej mora v minimumu veljati

$$\forall y \in \mathbb{R}^n : y^T H(P, x^*) y \geq 0$$

Če v tem pogoju za vse $y \neq 0$ velja strogi neenačaj, je točka x^* *strogi* minimum; sicer je potrebno stvar še preučiti. Povedano strnemo v naslednji izrek:

IZREK 1 *Potreben pogoj za to, da ima zvezna in dvakrat zvezno odvedljiva funkcija P naloge $(\mathbb{R}^n, P, \mathrm{Min})$ (lokalni) minimum v točki x^*, je*

$$\nabla P(x^*) = 0$$

in, da je Hessova matrika $H(P, x^*)$ pozitivno semi-definitna. Če je pozitivno definitna, je v točki x^* strogi (lokalni) minimum.

Pri dvakrat zvezno odvedljivih konveksnih funkcijah je drugi pogoj avtomatično izpolnjen. Zato zadostuje že prvi.

Regresijska premica

Imamo n podatkov, meritev

$$(x_i, y_i), \qquad i = 1, \ldots, n$$

pri čemer imata vsaj dve točki različno absciso. Kako določiti premico $y = ax + b$, ki jih najbolje povzema? Običajno uporabljamo *metodo najmanjših kvadratov*. Premico določimo tako, da je vsota kvadratov napak – odstopanj posameznih točk od premice najmanjša. Tako dobimo optimizacijsko nalogo $(\mathbb{R}^2, P, \mathrm{Min})$, kjer je

$$P(a, b) = \sum_{i=1}^{n} ((ax_i + b) - y_i)^2$$

Naloga zadošča pogojem izreka.

...Regresijska premica

Iz pogoja $\nabla P = 0$ dobimo enačbi

$$\begin{aligned}\frac{\partial P}{\partial a} &= \sum_{i=1}^{n} 2((ax_i + b) - y_i)x_i = 0 \\ \frac{\partial P}{\partial b} &= \sum_{i=1}^{n} 2((ax_i + b) - y_i) = 0\end{aligned}$$

z rešitvijo

$$a = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}, \qquad b = \frac{1}{n}(\sum y - a\sum x)$$

oziroma, če vpeljemo oznako $\overline{z} = \frac{1}{n}\sum z$

$$a = \frac{\overline{xy} - \overline{x}\,\overline{y}}{\overline{x^2} - \overline{x}^2}, \qquad b = \overline{y} - a\overline{x}$$

. . . Regresijska premica

Poglejmo še Hessovo matriko

$$H = \begin{bmatrix} \frac{\partial^2 P}{\partial a^2} & \frac{\partial^2 P}{\partial a \partial b} \\ \frac{\partial^2 P}{\partial b \partial a} & \frac{\partial^2 P}{\partial b^2} \end{bmatrix} = 2 \begin{bmatrix} \sum x^2 & \sum x \\ \sum x & n \end{bmatrix}$$

Ker je $\Delta_1 = 2\sum x^2 > 0$ in $\Delta_2 = 4(n\sum x^2 - (\sum x)^2) = 2\sum\sum(x_i - x_j)^2 > 0$, je matrika H pozitivno definitna in zato funkcija P strogo konveksna. Torej je *regresijska premica* enolično določena.

Kriterijska funkcija, ki meri napako po metodi najmanjših kvadratov je precej občutljiva na večje napake in tujke (outliers). Zato v takih primerih nekateri raje uporabljajo kriterijsko funkcijo, ki temelji na absolutnih napakah

$$P(a, b) = \sum_{i=1}^{n} |(ax_i + b) - y_i|$$

Seveda pa te naloge ne moremo več ugnati z gornjimi sredstvi.

Sedla

Točka $(x^*, u^*) \in \Phi \times \Psi$ *je minmax-sedlo* funkcije $G : \Phi \times \Psi \to \overline{\mathbb{R}}$ natanko takrat, ko velja:

$$\forall (x,u) \in \Phi \times \Psi : G(x, u^*) \geq G(x^*, u^*) \geq G(x^*, u)$$

in je *maxmin-sedlo* funkcije G natanko takrat, ko velja:

$$\forall (x,u) \in \Phi \times \Psi : G(x, u^*) \leq G(x^*, u^*) \leq G(x^*, u)$$

Če enega od pogojev za sedlo pomnožimo z -1 se neenakosti obrnejo – dobimo drugi pogoj. Torej velja:

IZREK 2 *Funkciji $G(x,u)$ in $-G(x,u)$ imata ista sedla, a nasprotnih vrst.*

...Sedla

Količino $G(x^*, u^*)$ imenujemo *vrednost* sedla (x^*, u^*). Da je slednje smiselno, nam zagotavlja naslednja lastnost:

IZREK 3 *Vsa sedla iste vrste imajo enako vrednost.*

Dokaz: Zaradi simetrije bomo dokazovali samo za minmax-sedla. Bodita (x^*, u^*) in $(\bar{x}, \bar{u})$ minmax-sedli. Potem po definiciji velja:

$$G(x^*, \bar{u}) \geq G(\bar{x}, \bar{u}) \geq G(\bar{x}, u^*) \geq G(x^*, u^*) \geq G(x^*, \bar{u})$$

kar je mogoče le, če povsod velja enačaj. Torej je res $G(x^*, u^*) = G(\bar{x}, \bar{u})$.

$\square$

. . . Sedla

Iz gornjega dokaza je le še majhen korak do lastnosti:

IZREK 4 *Množica sedel iste vrste funkcije* $G : \Phi \times \Psi \to \overline{\mathbb{R}}$ *je oblike* $\Phi_s \times \Psi_s$*, kjer sta* $\Phi_s \subseteq \Phi$ *in* $\Psi_s \subseteq \Psi$ *.*

Dokaz: Bodita (x^*, u^*) in $(\bar{x}, \bar{u})$ minmax-sedli. Tedaj sta minmax sedli tudi $(x^*, \bar{u})$ in $(\bar{x}, u^*)$. Pokažimo, na primer, da je $(x^*, \bar{u})$ minmax-sedlo. Napišimo pogoja: za vsak $(x, u) \in \Phi \times \Psi$ velja:

$$G(x, u^*) \geq G(x^*, u^*) \geq G(x^*, u) \qquad \text{in} \qquad G(x, \bar{u}) \geq G(\bar{x}, \bar{u}) \geq G(\bar{x}, u)$$

Iz prejšnjega dokaza vemo $G(\bar{x}, u^*) = G(x^*, u^*) = G(\bar{x}, \bar{u}) = G(x^*, \bar{u})$. Torej je res tudi

$$G(x, \bar{u}) \geq G(\bar{x}, \bar{u}) = G(x^*, \bar{u}) = G(x^*, u^*) \geq G(x^*, u)$$

kar je bilo treba pokazati. □

Izrek o mini-maxu

Sedaj pa si oglejmo še najpomembnejši rezultat tega razdelka, ki daje potreben in zadosten pogoj za obstoj sedla.

IZREK 5 *Funkcija* $G : \Phi \times \Psi \to \overline{\mathbb{R}}$ *ima minmax-sedlo natanko takrat, ko obstaja točka* $(x^*, u^*) \in \Phi \times \Psi$, *tako da velja:*

$$\min_{x\in\Phi} \sup_{u\in\Psi} G(x,u) = \max_{u\in\Psi} \inf_{x\in\Phi} G(x,u) = G(x^*,u^*)$$

in ima maxmin-sedlo natanko takrat, ko obstaja točka $(x^*, u^*) \in \Phi \times \Psi$, *tako da velja:*

$$\max_{x\in\Phi} \inf_{u\in\Psi} G(x,u) = \min_{u\in\Psi} \sup_{x\in\Phi} G(x,u) = G(x^*,u^*)$$

Dokaz izreka najdete v zapiskih.

Primeri sedel

Brez večjih težav lahko sestavimo primere funkcije G (običajno so podane z matriko), ki so ali brez sedla, ali imajo samo sedla enega tipa, ali pa imajo sedla obeh tipov. Tako ima na primer funkcija G podana z matriko:

		u				$\min_u$	$\max_u$
		1	2	3	4		
	1	1	0	5	2	0	5
x	2	2	2	4	3	2	4
	3	0	1	7	2	0	7
$\min_x$		0	0	4	2		4 = 4
$\max_x$		2	2	7	3	2 = 2	

dve maxmin-sedli $(2, 1)$ in $(2, 2)$ ter eno minmax-sedlo $(2, 3)$.

…Primeri sedel

Za nadaljne izreke o obstoju sedel je potrebno privzeti dodatne lastnosti za funkcijo G in množici Φ in Ψ; npr. odvedljivost, konveksnost/konkavnost, ...

Iz analize vemo: Naj bo $G : \mathbb{R} \times \mathbb{R} \to \overline{\mathbb{R}}$ v okolici točke (x^*, u^*) dvakrat zvezno odvedljiva funkcija. Za to, da je (x^*, u^*) *analitično sedlo*, zadostuje, da velja $\nabla G(x^*, u^*) = 0$ in je $H(G, (x^*, u^*))$ negativno definitna.

Vendar pa ni vsako analitično sedlo tudi sedlo v smislu tega razdelka, kar je očitno iz izreka 3.

Prirejene in dualne naloge

Naj bo dana funkcija $G : \Phi \times \Psi \to \overline{\mathbb{R}}$, ki določa funkciji

$$P(x) = \sup_{u \in \Psi} G(x,u) \qquad \text{in} \qquad Q(u) = \inf_{x \in \Phi} G(x,u)$$

potem pravimo, da sta si nalogi (Φ, P, Min) in (Ψ, Q, Max) *prirejeni* glede na G.

IZREK 6 *Za s funkcijo $G : \Phi \times \Psi \to \overline{\mathbb{R}}$ določeni funkciji P in Q velja:*

$$\forall (x,u) \in \Phi \times \Psi : P(x) \geq Q(u)$$

Dokaz:

$$P(x) = \sup_{v \in \Psi} G(x,v) \geq G(x,u) \geq \inf_{y \in \Phi} G(y,u) = Q(u)$$

$\square$

Dualne naloge

V primeru, ko obstaja točka $(x^*, u^*) \in \Phi \times \Psi$, za katero velja enakost $P(x^*) = Q(u^*)$ pravimo, da sta si prirejeni nalogi *dualni* ali *pridruženi*; točki (x^*, u^*) pa *dualna točka*.

Opomba: Gornja definicija dualnosti je nekoliko ožja od tiste, ki jo običajno srečamo v literaturi. Z njo se odpovemo primerom, ko je neka od množic Ψ in Φ prazna.

Dualne naloge – lastnosti

IZREK 7 *Bodita glede na G prirejeni nalogi (Φ, P, Min) in (Ψ, Q, Max) dualni in naj bo (x^*, u^*) dualna točka, potem velja:*

a. $x^* \in \mathrm{Min}(\Phi, P)$ *in* $u^* \in \mathrm{Max}(\Psi, Q)$

b. $\min(\Phi, P) = \max(\Psi, Q) = G(x^*, u^*) = P(x^*) = Q(u^*)$

Dokaz:

a. Iz izreka 6 izhaja, da velja za vsak $x \in \Phi : P(x) \geq Q(u^*)$. Po predpostavki $Q(u^*) = P(x^*)$ izhaja naprej $P(x) \geq P(x^*)$; kar pomeni, da je res $x^* \in \mathrm{Min}(\Phi, P)$. Podobno pokažemo tudi, da je $u^* \in \mathrm{Max}(\Psi, Q)$.

b. Ker sta po delu **a** $x^* \in \mathrm{Min}(\Phi, P)$ in $u^* \in \mathrm{Max}(\Psi, Q)$, je po definiciji in predpostavki $\min(\Phi, P) = P(x^*) = Q(u^*) = \max(\Psi, Q)$.

Pokazati moramo še, da lahko v ta spisek enakosti dodamo tudi $G(x^*, u^*)$. Izhajamo iz neenakosti $G(x, u^*) \geq Q(u^*) = P(x^*) \geq G(x^*, u)$. Postavimo vanjo $x = x^*$ in $u = u^*$ pa dobimo

$$G(x^*, u^*) \geq Q(u^*) = P(x^*) \geq G(x^*, u^*)$$

kar je mogoče le, če povsod veljajo enačaji. S tem je trditev dokazana. □

Dualnost in sedla

Zvezo med dualnostjo in sedli nam daje naslednji izrek:

IZREK 8 *Glede na funkcijo* $G : \Phi \times \Psi \to \overline{\mathbb{R}}$ *prirejeni nalogi* (Φ, P, Min) *in* (Ψ, Q, Max) *sta si dualni natanko takrat, ko ima funkcija* G *minmax-sedlo.*

Dokaz: V eno smer (**dualnost** $\Rightarrow$ **sedlo**) je trditev že skoraj dokazana. Naj bo (x^*, u^*) dualna točka. Upoštevajmo v neenakosti iz prejšnjega dokaza končno ugotovitev $P(x^*) = Q(u^*) = G(x^*, u^*)$ pa imamo:

$$G(x, u^*) \geq G(x^*, u^*) \geq G(x^*, u)$$

Torej je (x^*, u^*) res minmax-sedlo.

V nasprotno smer (**sedlo** $\Rightarrow$ **dualnost**) pa sklepamo takole: Naj bo (x^*, u^*) minmax-sedlo. Potem je po definiciji $\forall u \in \Psi : G(x^*, u) \leq G(x^*, u^*)$. Torej tudi:

$$P(x^*) = \sup_{u \in \Psi} G(x^*, u) \leq G(x^*, u^*) \quad \text{in podobno} \quad Q(u^*) = \inf_{x \in \Phi} G(x, u^*) \geq G(x^*, u^*)$$

oziroma $P(x^*) \leq Q(u^*)$. Od tu po izreku 6 sledi enakost $P(x^*) = Q(u^*)$. $\square$

Enakovrednost dualnih nalog

Vzemimo sedaj poljubna $x^* \in \mathrm{Min}(\Phi, P)$ in $u^* \in \mathrm{Max}(\Psi, Q)$. Če sta si nalogi (Φ, P, Min) in (Ψ, Q, Max) dualni, potem je točka (x^*, u^*) po izreku 7.b dualna točka in po izreku 8 minmax sedlo. Po drugi strani iz izreka 8 in izreka 7.a izhaja, da je vsako minmax-sedlo $(x^*, u^*) \in \mathrm{Min}(\Phi, P) \times \mathrm{Max}(\Psi, Q)$. Kar pomeni, da množica minmax sedel funkcije G sovpada z množico $\mathrm{Min}(\Phi, P) \times \mathrm{Max}(\Psi, Q)$.

Vpeljimo še relaciji

$$\tau(x) = \{u : Q(u) = P(x)\} \qquad \text{in} \qquad \theta(u) = \{x : P(x) = Q(u)\}$$

pa vidimo iz pravkar povedanega, da mora veljati $\forall x \in \mathrm{Min}(\Phi, P) : \tau(x) = \mathrm{Max}(\Psi, Q)$ oziroma

$$\tau(\mathrm{Min}(\Phi, P)) = \bigcup_{x \in \mathrm{Min}(\Phi, P)} \tau(x) = \mathrm{Max}(\Psi, Q)$$

in podobno za θ. Torej sta τ in θ translatorja. Povzemimo:

. . . Enakovrednost dualnih nalog

IZREK 9 *Glede na G dualni nalogi (Φ, P, Min) in (Ψ, Q, Max) sta enakovredni in velja:*

$$\{(x, u) \in \Phi \times \Psi : P(x) = Q(u)\} = \mathrm{Min}(\Phi, P) \times \mathrm{Max}(\Psi, Q)$$

Lagrangeova prirejenost

Najpogosteje pridemo do prirejene naloge z Langrangeovo funkcijo. Poglejmo, kako to naredimo:

Naj bo optimizacijsko območje Φ naloge (Φ, P, Min) podano takole

$$\Phi = \{x \in \Omega : P_i(x) \leq 0, i \in I,\ P_j(x) = 0, j \in J\}, \qquad I \cup J = 1..n$$

in naj bo $\Phi \neq \emptyset$. V primeru, ko je $\Phi \neq \emptyset$, pravimo, da so omejitve in naloga *neprotislovne*. Če obstaja točka $x \in \Omega$, tako da je $P_i(x) < 0, i \in I$, je naloga *strogo neprotislovna*. Točki x rečemo tudi Slaterjeva točka.

Opomba: Pogoj oblike $P_j(x) = 0$ nadomestimo z dvema enakovrednima pogojema

$$P_j(x) \leq 0 \qquad \text{in} \qquad -P_j(x) \leq 0$$

Predpostavimo začasno, $J = \emptyset$.

Lagrangeova funkcija

Označimo $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ in $\mathbf{P}(x) = (P_1(x), P_2(x), \ldots, P_n(x))$ ter

$$(\mathbf{u}, \mathbf{v}) = \sum_{i \in I} u_i v_i$$

potem lahko definiramo *Lagrangeovo funkcijo* $L : \Omega \times (\mathbb{R}_0^+)^n \to \overline{\mathbb{R}}$ naloge (Φ, P, Min) takole:

$$L(x, \mathbf{u}) = P(x) + (\mathbf{u}, \mathbf{P}(x))$$

Od tu dobimo (glede na L) prirejeni nalogi $(\Omega, P^*, \mathrm{Min})$ in $((\mathbb{R}_0^+)^n, Q, \mathrm{Max})$, kjer je

$$P^*(x) = \sup_{\mathbf{u} \in (\mathbb{R}_0^+)^n} L(x, \mathbf{u}) \qquad \text{in} \qquad Q(\mathbf{u}) = \inf_{x \in \Omega} L(x, \mathbf{u})$$

Pokažimo najprej, da je naloga $(\Omega, P^*, \mathrm{Min})$ enakovredna osnovni nalogi!

Prva prirejena naloga je enakovredna osnovni

Nastopita dva primera:

a) $x \in \Phi$; tedaj so vsi $P_i(x) \leq 0$; kar pomeni, da je tudi $\mathbf{P}(x) \leq \mathbf{0}$ in dalje $(\mathbf{u}, \mathbf{P}(x)) \leq 0$. Toda za $\mathbf{u} = \mathbf{0}$ je $(\mathbf{u}, \mathbf{P}(x)) = 0$ in zato

$$\sup_{\mathbf{u} \in (\mathbb{R}_0^+)^n} (\mathbf{u}, \mathbf{P}(x)) = 0$$

Torej je v tem primeru $P^*(x) = P(x)$.

b) $x \notin \Phi$; tedaj je vsaj en $P_i(x) > 0$. Definirajo zaporedje $(\mathbf{u}_k)$ takole:

$$\mathbf{u}_{kj} = k\delta_{ij}$$

Tedaj za $k \to \infty$ tudi $(\mathbf{u}_k, \mathbf{P}(x)) \to \infty$. Torej je

$$P^*(x) = \sup_{\mathbf{u} \in (\mathbb{R}_0^+)^n} (\mathbf{u}, \mathbf{P}(x)) = \infty$$

. . . Prva prirejena naloga je enakovredna osnovni

Povzemimo :

$$P^*(x) = \begin{cases} P(x) & x \in \Phi \\ \infty & x \in \Omega \setminus \Phi \end{cases}$$

Ker je po predpostavki $\Phi \neq \emptyset$, izhaja od tu enakovrednost osnovne naloge (Φ, P, Min) in naloge $(\Omega, P^*, \mathrm{Min})$; če pa je naloga $((\mathbb{R}_0^+)^n, Q, \mathrm{Max})$ dualna nalogi $(\Omega, P^*, \mathrm{Min})$, zaradi tranzitivnosti enakovrednosti in enakovrednosti dualnih nalog izhaja, da je tudi naloga $((\mathbb{R}_0^+)^n, Q, \mathrm{Max})$ enakovredna osnovni nalogi.

Druga prirejena naloga

Seveda moramo funkcijo $Q(\mathbf{u})$ določiti v vsakem primeru posebej. Poglejmo še primer pogoja oblike $P_i(x) = 0$. Ta nam v Lagrangeovi funkciji prispeva člena

$$u^+ P_i(x) - u^- P_i(x) = (u^+ - u^-) P_i(x) = u P_i(x)$$

Iz te enakosti vidimo, da lahko oba člena nadomestimo z enim samim, pri tem pa moramo pogoja $u^+, u^- \geq 0$ nadomestiti (omiliti) z $u \in \mathbb{R}$.

Če so P, P_i konveksne in je naloga rešljiva, ima L sedlo.

Druga prirejena naloga je 'lepa'

IZREK 10 *Glede na L prirejena funkcija $Q(\mathbf{u})$ je konkavna.*

Dokaz: Vzemimo $\mathbf{v}, \mathbf{w} \in (\mathbb{R}_0^+)^n$ in naj bo

$$\mathbf{u} = \lambda\mathbf{v} + (1-\lambda)\mathbf{w}, \qquad \lambda \in [0,1]$$

Tedaj velja

$$\begin{aligned} Q(\mathbf{u}) &= \inf_{x\in\Omega} L(x,\mathbf{u}) = \inf_{x\in\Omega}(P(x) + (\mathbf{u}, \mathbf{P}(x))) \\ &= \inf_{x\in\Omega}((\lambda + (1-\lambda))P(x) + (\lambda\mathbf{v} + (1-\lambda)\mathbf{w}, \mathbf{P}(x))) \end{aligned}$$

upoštevajmo še $\inf(f+g) \geq \inf f + \inf g$ pa lahko nadaljujemo

$$\begin{aligned} &\geq \lambda \inf_{x\in\Omega}(P(x) + (\mathbf{v}, \mathbf{P}(x))) + (1-\lambda) \inf_{x\in\Omega}(P(x) + (\mathbf{w}, \mathbf{P}(x))) \\ &= \lambda Q(\mathbf{v}) + (1-\lambda)Q(\mathbf{w}) \end{aligned}$$

□

Druga prirejena naloga

Opomba: V gornjih razmišljanjih o prirejenih nalogah ni nobenih posebnih predpostavk o naravi P, P_i in Φ. Na primer, ko je $\Omega = \mathbb{Z}^m$, funkcija Q ni odvedljiva; vendar je zaradi konkavnosti vsak lokalni maksimum u^* tudi globalni.

V splošnem je prirejeno nalogo torej lažje reševati. Žal prirejeni nalogi nista vedno dualni (enakovredni). Kljub temu zaradi

$$\max((\mathbb{R}_0^+)^n, Q) = Q(\mathbf{u}^*) \leq \min(\Phi, P)$$

vrednost $Q(\mathbf{u}^*)$ ali njeno oceno lahko uporabljamo za spodnjo mejo P.

Primer uporabe Lagrangeove prirejenosti

Pokažimo, kako lahko uporabimo dualnost pri reševanju naloge (Φ, P, Min)

$$P(x, y) = x^2 + y^2$$

$$\Phi = \{(x, y) \in \mathbb{R}^2 : 2x + y \leq -4\}$$

Pripadajoča Lagrangeova funkcija je

$$L(x, y; u) = x^2 + y^2 + u \cdot (2x + y + 4), \qquad u \geq 0$$

Ker sta kriterijska funkcija in omejitev konveksni ima L sedlo. Poiščimo

$$Q(u) = \inf_{(x,y)\in\mathbb{R}^2} L(x, y; u)$$

Ker je L zvezno odvedljiva, lahko inf določimo po izreku iz pogojev

$$\frac{\partial L}{\partial x} = 2x + 2u = 0 \qquad \frac{\partial L}{\partial y} = 2y + u = 0$$

…Primer uporabe Lagrangeove prirejenosti

od koder dobimo

$$\hat{x} = -u \qquad \hat{y} = -\frac{u}{2}$$

Torej je

$$Q(u) = L(\hat{x}, \hat{y}, u) = -\frac{5u^2}{4} + 4u$$

Funkcija $Q(u)$ je parabola in doseže maksimum v temenu $u^* = \frac{8}{5}$. Torej je

$$x^* = -\frac{8}{5} \qquad y^* = -\frac{4}{5} \qquad Q(u^*) = P(x^*, y^*) = \frac{16}{5}$$

Karush-Kuhn-Tuckerjev izrek

Poglejmo ponovno postopek iz primera na nalogi (Φ, P, Min), kjer je

$$\Phi = \{\mathbf{x} \in \mathbb{R}^m : P_i(\mathbf{x}) \leq 0, i \in I,\ P_j(\mathbf{x}) = 0, j \in J\}$$

Pripadajoča Lagrangeova funkcija je

$$L(\mathbf{x}, \mathbf{u}) = P(\mathbf{x}) + (\mathbf{u}, \mathbf{P}(\mathbf{x}))$$

pri čemer, če zapišemo $\mathbf{u} = (\mathbf{u}_I, \mathbf{u}_J)$, je $\mathbf{u}_I \geq \mathbf{0}$. Določimo

$$Q(\mathbf{u}) = \inf_{\mathbf{x} \in \mathbb{R}^m} L(\mathbf{x}, \mathbf{u})$$

oziroma, še nekoliko več, določimo $\mathbf{x}^*(\mathbf{u}) \in \mathbb{R}^m$, za katerega je

$$Q(\mathbf{u}) = L(\mathbf{x}^*(\mathbf{u}), \mathbf{u})$$

Če je L zvezno odvedljiva po $\mathbf{x}$, mora po izreku 1 $\mathbf{x}^*(\mathbf{u})$ zadoščati pogoju

$$\nabla_x L = \mathbf{0}$$

... Karush-Kuhn-Tuckerjev izrek

in, če naj si z rešitvijo prirejene naloge pomagamo pri reševanju osnovne naloge, morata biti nalogi dualni

$$P(\mathbf{x}^*) = Q(\mathbf{u}^*) = L(\mathbf{x}^*, \mathbf{u}^*)$$

Torej mora veljati

$$P(\mathbf{x}^*) = L(\mathbf{x}^*, \mathbf{u}^*) = P(\mathbf{x}^*) + (\mathbf{u}^*, \mathbf{P}(\mathbf{x}^*))$$

oziroma

$$0 = (\mathbf{u}^*, \mathbf{P}(\mathbf{x}^*)) = \sum_{i \in I} u_i^* P_i(\mathbf{x}^*) + \sum_{j \in J} u_j^* P_j(\mathbf{x}^*)$$

Ker je $P_j(\mathbf{x}^*) = 0, j \in J$, je drugi člen enak 0 in preostane $\sum_{i \in I} u_i^* P_i(\mathbf{x}^*) = 0$. Toda, $u_i^* \geq 0, P_i(\mathbf{x}^*) \leq 0, i \in I$, in zato končno

$$u_i^* P_i(\mathbf{x}^*) = 0, \quad i \in I$$

. . . Karush-Kuhn-Tuckerjev izrek

Iz povedanega lahko povzamemo (potrebnostni) del Karush-Kuhn-Tucker-jevega izreka:

IZREK 11 *Naj bodo P in $P_k, k \in K = I \cup J$ zvezno odvedljive funkcije in naj ima Lagrangeova funkcija L sedlo. Potem je potreben pogoj za to, da ima naloga (Φ, P, Min) v točki $\mathbf{x}^* \in \Phi$ lokalni minimum, obstoj realnih števil $u_j, j \in J$ in $u_i \geq 0, i \in I$, za katere velja:*

$$\nabla_x L = \nabla_x P(\mathbf{x}^*) + \sum_{i \in I} u_i \nabla_x P_i(\mathbf{x}^*) + \sum_{j \in J} u_j \nabla_x P_j(\mathbf{x}^*) = \mathbf{0}$$

in

$$u_i P_i(\mathbf{x}^*) = 0, \qquad i \in I$$

Poleg pogojem za u_i mora rešitev $\mathbf{x}^*$, če naj bo dopustna, zadoščati še omejitvam

$$P_i(\mathbf{x}^*) \leq 0, \quad i \in I \qquad \text{in} \qquad P_j(\mathbf{x}^*) = 0, \quad j \in J$$

...Karush-Kuhn-Tuckerjev izrek

Z rešitvijo tako dobljenega sistema enačb dobimo kandidate za "lokalna" sedla $(\mathbf{x}^*, \mathbf{u}^*)$.

Za konveksne naloge velja tudi obrat.

IZREK 12 *Če je osnovna naloga* (Φ, P, Min) *konveksna, strogo neprotislovna in* $J = \emptyset$*, so Karush-Khun-Tuckerjevi pogoji tudi zadostni.*

... Karush-Kuhn-Tuckerjev izrek

Dokaz: Iz konveksnosti funkcij P in P_i izhaja, da v točki $\mathbf{x}^*$ velja

$$\begin{aligned} P(\mathbf{x}) &\geq P(\mathbf{x}^*) + \nabla P^T(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) \\ P_i(\mathbf{x}) &\geq P_i(\mathbf{x}^*) + \nabla P_i^T(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*), \quad i \in I \end{aligned}$$

Pomnožimo zadnje neenakosti z $u_i^* \geq 0$, seštejmo jih in prištejmo prvo pa dobimo

$$P(\mathbf{x}) + \sum_{i \in I} u_i^* P_i(\mathbf{x}) \geq P(\mathbf{x}^*) + \sum_{i \in I} u_i^* P_i(\mathbf{x}^*) +$$

$$(\nabla P^T(\mathbf{x}^*) + \sum_{i \in I} u_i^* \nabla P_i^T(\mathbf{x}^*))(\mathbf{x} - \mathbf{x}^*)$$

Predzadnji in zadnji člen sta po K-K-T izreku enaka 0. Torej je za $\mathbf{x} \in \Phi$, ker je $\sum u_i^* P_i(\mathbf{x}) \leq 0$

$$P(\mathbf{x}) \geq P(\mathbf{x}^*) - \sum_{i \in I} u_i^* P_i(\mathbf{x}) \geq P(\mathbf{x}^*)$$

kar pomeni $\mathbf{x}^* \in \mathrm{Min}(\Phi, P)$. $\square$

Primerne omejitve

Omejitev $P_i(\mathbf{x}) \leq 0$ je *zasičena* ali *dejavna* v točki $\mathbf{x}_0$, če je $P_i(\mathbf{x}_0) = 0$.

Pri odgovoru na vprašanje, kdaj so pogoji izreka tudi zadostni, potrebujemo pojem *primernosti* omejitev. Mi se bomo izognili definiciji tega pojma in navedli nekaj zadostnih pogojev za primernost, ki pokrivajo večino praktičnih primerov:

- linearnost funkcij P_i;
- konveksnost funkcij P_i in nepraznost notranjosti množice Φ;
- neodvisnost gradientov omejitev, ki so zasičene v dani točki.

Pokazati je mogoče, da so pogoji iz izreka zadostni, če je osnovna naloga konveksna in so omejitve primerne v točki $\mathbf{x}^*$.

Karush-Kuhn-Tuckerjev izrek – primer

Določimo minimum funkcije $P(x, y) = x$ pri omejitvah $(x-1)^2 + (y+2)^2 \leq 16$ in $x^2 + y^2 \geq 13$.

Najprej zapišimo omejitvi v "standardni" obliki

$$\begin{aligned} P_1(x, y) &= (x-1)^2 + (y+2)^2 - 16 \leq 0 \\ P_2(x, y) &= -x^2 - y^2 + 13 \leq 0 \end{aligned}$$

Poleg teh dveh neenačb dobimo iz K-K-T pogojev še zveze

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} + u \begin{bmatrix} 2(x-1) \\ 2(y+2) \end{bmatrix} + v \begin{bmatrix} -2x \\ -2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$u \cdot ((x-1)^2 + (y+2)^2 - 16) = 0$$

$$v \cdot (x^2 + y^2 - 13) = 0$$

pri čemer sta $u, v \geq 0$.

...Karush-Kuhn-Tuckerjev izrek – primer

Iz prve zveze dobimo enačbi

$$u \cdot (x-1) - vx = -\frac{1}{2} \qquad \text{in} \qquad u \cdot (y+2) = vy$$

Do točk, v katerih je morda minimum, pridemo z analizo rešitev tega sistema enačb.

1. $v = 0$; iz enačb $u \cdot (y+2) = 0$ in $u \cdot (x-1) = -\frac{1}{2}$ izhaja, da je $u \neq 0$ in $y = -2$. Iz enačbe $(x-1)^2 + (y+2)^2 = 16$ dobimo $(x-1)^2 = 16$. Koren $x = 5$ odpade, ker je pripadajoči $u = -\frac{1}{8} \leq 0$. Drugi koren da rešitev

$$x = -3,\ y = -2,\ u = \frac{1}{8},\ v = 0,\ P = -3$$

...Karush-Kuhn-Tuckerjev izrek – primer

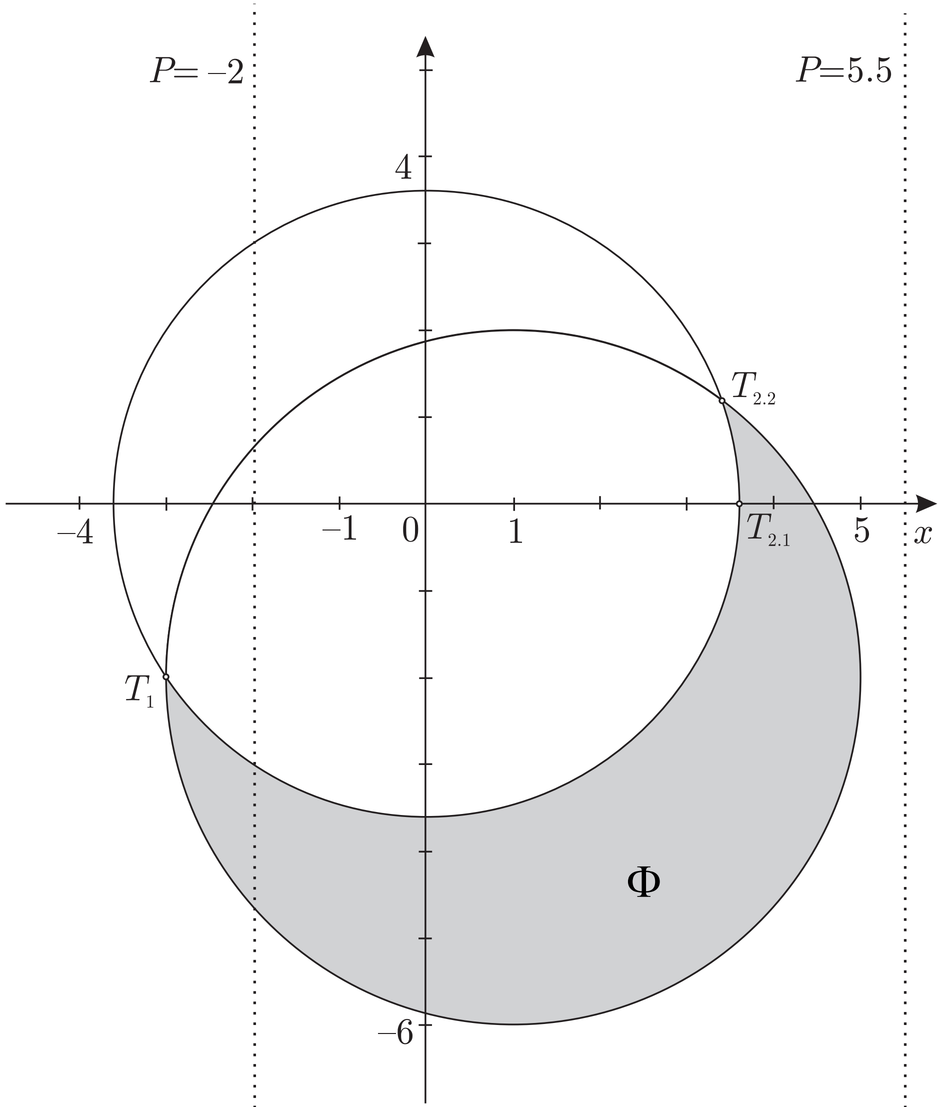
2. $v \neq 0$; tedaj je $x^2 + y^2 = 13$;
2.1. $u = 0$; tedaj je $y = 0$ in zato $x = \pm\sqrt{13}$. Negativni koren odpade, ker je pripadajoči v negativen. Ostane rešitev

$$x = \sqrt{13},\ y = 0,\ u = 0,\ v = \frac{1}{2\sqrt{13}},\ P = \sqrt{13}$$

2.2. $u \neq 0$; dobimo sistem kvadratnih enačb $x^2 + y^2 = 13$ in $(x-1)^2 + (y+2)^2 = 16$, iz katerega izhaja $x = 2y + 1$ in $5y^2 + 4y - 12 = 0$.
Koren $y = -2$ smo že dobili v primeru 1 (iz $vy = 0$ pa izhaja protislovje).
Koren $y = \frac{6}{5}$ pa nam da novo rešitev

$$x = \frac{17}{5},\ y = \frac{6}{5},\ u = \frac{3}{40},\ v = \frac{1}{5},\ P = \frac{17}{5}$$

. . . Karush-Kuhn-Tuckerjev izrek – primer



Nalogo $(\Phi, P, \min)$ si lahko ponazorimo tudi na sliki. Iz nje vidimo, da je rešitev **1** globalni minimum, rešitev **2.2** lokalni minimum in rešitev **2.1** le K-K-T točka.

Numerični postopki

Obstaja več postopkov za reševanje optimizacijskih nalog oblike

$$(\mathbb{R}^n, P, \mathrm{Min})$$

Ločijo se predvsem po tem, koliko upoštevajo lepe lastnosti kriterijske funkcije P. Osnovna delitev je tako na:

- postopki, ki uporabljajo samo funkcijske vrednosti: pokoordinatni spust, Hooke-Jeevesov postopek, Nelder-Meadov postopek;
- postopki, ki uporabljajo tudi prve odvode kriterijske funkcije: gradientni postopek;
- postopki, ki upoštevajo tudi druge odvode kriterijske funkcije: Davidon-Fletcher-Powellov postopek, Fletcher-Reevesov postopek.

. . . Numerični postopki

Za probleme z omejitvami obstajata dva pristopa:

- premi postopki – pazimo, da ne zapustimo množice dopustnih rešitev; linearizacija problema;
- prevedba na probleme brez omejitev: Lagrangeova dualnost, kazenske funkcije (v naslednjem razdelku);
- postopki za posebne vrste problemov: linearno programiranje, Marquardtov postopek za določanje parametrov po metodi najmanjših kvadratov.

Pokoordinatni spust

```
izberi začetno točko x ∈ ℝ^n in koračni vektor h ∈ ℝ^n;
loop
    y := x; change := FALSE;
    for i := 1 to n do begin
        z := y + h e_i;
        if P(z) < P(y) then
            begin y := z; change := TRUE end
        else begin
            z := y − h e_i;
            if P(z) < P(y) then
                begin y := z; change := TRUE end
        end;
    end;
    x := y;
    if not change then begin
        if error(h) < epsilon then exit;
        h := h/q;
    end;
end;
```

Običajne vrednosti parametrov so $q = 10$ in $epsilon = 0.0001$.

Hooke-Jeevesov postopek

Hooke-Jeevesov (1961) postopek je izboljšava pokoordinatnega spusta. Pri njem izkoristimo informacijo, ki smo jo nabrali v zadnjem koraku spusta, in se še poskusimo v dobljeni smeri premakniti kar se da daleč. Stavek $x := y$; nadomestimo s stavki

$d := y - x$;
repeat $x := y;\ y := y + d$ **until** $P(y) \geq P(x)$;

Pravimo tudi, da v zanki **for** raziščemo okolico dane točke x, v dodatku pa napredujemo.

Obstaja več izpopolnitev tega postopka, ki različno popravljajo korak v posamezni smeri.

Nelder-Meadov simpleksni postopek

Prvo zamisel postopka so podali Spendley, Hext in Himsworth (1962). Precej sta ga izpopolnila Nelder in Mead (1965). Zelo dobro se obnese, če razsežnost naloge ni prevelika (do 6).

določi začetne točke simpleksa S: $x_i \in \mathbb{R}^n$, $i = 1, \ldots, n+1$;
$y_i := P(x_i)$; $i = 1, \ldots, n+1$;
while velikost$(S) > \varepsilon$ **do begin**
  $x_M :=$ najvišja točka v S; $y_M := P(x_M)$;
  $x_d :=$ druga najvišja točka v S; $y_d := P(x_d)$;
  $x_m :=$ najnižja točka v S; $y_m := P(x_m)$;
  $x_0 := \frac{1}{n} \sum_{i \neq M} x_i$ težišče $S \setminus \{x_M\}$;
  ZRCALJENJE za $\alpha > 0$: $x_z - x_0 = \alpha(x_0 - x_M)$ oziroma
    $x_z := (1+\alpha)x_0 - \alpha x_M$; $y_z = P(x_z)$;
  if $y_z < y_d$ **then**
    if $y_z < y_m$ **then begin**
      RAZTEG za $\gamma > 1$: $x_r - x_0 = \gamma(x_z - x_0)$ oziroma
        $x_r := \gamma x_z + (1-\gamma)x_0$; $y_r = P(x_r)$;
      if $y_r < y_m$ **then begin** $x_M := x_r$; $y_M := y_r$ **end**
      else begin $x_M := x_z$; $y_M := y_z$ **end**
    end else begin $x_M := x_z$; $y_M := y_z$ **end**

...Nelder-Meadov simpleksni postopek

else begin
if $y_z > y_M$ **then begin**
$x := x_z; x_z := x_M; x_M := x;$
$y := y_z; y_z := y_M; y_M := y;$
end;
SKRČITEV za $0 < \beta < 1$: $x_s - x_0 = \beta(x_M - x_0)$ oziroma
$x_s := \beta x_M + (1-\beta)x_0; y_s = P(x_s);$
if $y_s \geq y_M$ **then begin**
POMANJŠAMO simpleks S:
$x_i := \frac{1}{2}(x_m + x_i); y_i := P(x_i)$; za $i \neq m$
end else begin $x_M := x_s; y_M := y_s$ **end**
end
end;
x^* := najnižja točka v S; $y^* := P(x^*)$;

velikost$(S) = \sqrt{\frac{1}{n+1}\sum_{i=1}^{n+1}(y_i - \overline{y})^2}$, $\overline{y} = \frac{1}{n+1}\sum_{i=1}^{n+1} y_i$

začetne točke: dana x_1, $x_{i+1} = x_1 + ke_i$, $i = 1, \ldots, n$.

običajne vrednosti parametrov so $\alpha = 1$, $\beta = 0.5$ in $\gamma = 2$.

Gradientni postopek

Naj bo kriterijska funkcija P odvedljiva. Kako izbrati v pomiku $x \to x+hs$, $h \geq 0$ enotski smerni vektor s tako, da bo padec vrednosti P največji?

Vpeljimo funkcijo

$$Q(s) = P(x+hs) - P(x) = \sum_{i=1}^{n} \frac{\partial P}{\partial x_i}(x)\, hs_i$$

Tako dobimo optimizacijsko nalogo (Φ, Q, Min), kjer množico dopustnih rešitev sestavljajo vsi enotski vektorji $\Phi = \{s \in \mathbb{R}^n : \sum_{i=1}^{n} s_i^2 = 1\}$. Pripadajoča Lagrangeova funkcija je

$$\mathcal{L}(s,\lambda) = Q(s) + \lambda(\sum_{i=1}^{n} s_i^2 - 1) = \sum_{i=1}^{n} \frac{\partial P}{\partial x_i}(x)\, hs_i + \lambda(\sum_{i=1}^{n} s_i^2 - 1)$$

kjer je $\lambda \in \mathbb{R}$.

…Gradientni postopek

Od tu dobimo sistem enačb

$$\frac{\partial \mathcal{L}}{\partial s_i} = h\frac{\partial P}{\partial x_i}(x) + 2\lambda s_i = 0, \qquad i = 1, \ldots, n$$

iz katerih izhaja $s = -\frac{h}{2\lambda}\nabla P(x) = -\alpha\nabla P(x)$, kjer je $\alpha \geq 0$; ali

$$Q(s) = -\alpha\sum_{i=1}^{n}(\frac{\partial P}{\partial x_i}(x))^2 \leq 0$$

Torej je smer največjega pada vrednosti kriterijske funkcije smer nasprotna njenemu gradientu. Ta ugotovitev je osnova *gradientnega postopka*

izberi začetno točko x;
repeat
    $s := -\nabla P(x)$;
    določi minimum y funkcije $P(x)$ na premici $x + \lambda s$;
    $error := |x - y|;\ x := y$;
until $error < epsilon$;

Kazenske metode

Vzemimo optimizacijsko nalogo (Φ, P, Min), kjer je Φ oblike

$$\Phi = \{x \in \Omega : P_i(x) \leq 0,\ i \in I\}$$

in naj bo funkcija $k : \mathbb{R} \to \overline{\mathbb{R}}$ določena s predpisom

$$k(y) = \begin{cases} 0 & y \leq 0 \\ \infty & y > 0 \end{cases}$$

Postavimo za $x \in \Omega$

$$K(x) = \sum_{i \in I} k(P_i(x)) \quad \text{in} \quad Q(x) = P(x) + K(x)$$

Tedaj je $K(x) = \begin{cases} 0 & x \in \Phi \\ \infty & x \notin \Phi \end{cases}$ in dalje $Q(x) = \begin{cases} P(x) & x \in \Phi \\ \infty & x \notin \Phi \end{cases}$.
Torej, če je $\Phi \neq \emptyset$, je

$$(\Phi, P, \mathrm{Min}) \approx (\Omega, Q, \mathrm{Min})$$

… Kazenske metode

Druga naloga je brez omejitev! No, kriterijska funkcija Q praviloma ne ohranja lepih lastnosti (npr. zveznosti), ki jih morda imajo funkcije P in $P_i, i \in I$.

Ali je mogoče zamisel popraviti tako, da ne izgubimo (vseh) lepih lastnosti funkcij P in $P_i, i \in I$? V nekaterih primerih je odgovor pritrdilen. To lahko storimo celo na več načinov. Prve korake je naredil Courant (okrog 1943), podrobneje pa sta leta 1968 pristop razdelala Fiacco in McCormick.

Nalogi (Φ, P, Min) priredimo družino nalog $(\Omega, Q(., r), \mathrm{Min})$ odvisnih od parametra $r \in \mathbb{R}^+$, za katero velja

$$(\Omega, Q(., r), \mathrm{Min}) \quad \xrightarrow[r\to\infty]{} \quad (\Omega, Q, \mathrm{Min})$$

pri čemer imajo funkcije $Q(., r), r \in \mathbb{R}$ zahtevane lepe lastnosti.

Metoda zunanje točke

Pri *metodi zunanje točke* izberemo za funkcijo k, na primer,

$$k(y) = \begin{cases} 0 & y \leq 0 \\ y & y > 0 \end{cases} \quad \text{ali} \quad k(y) = \begin{cases} 0 & y \leq 0 \\ y^2 & y > 0 \end{cases}$$

in nato postavimo

$$K(x,r) = r \sum_{i \in I} k(P_i(x)) \quad \text{ter} \quad Q(x,r) = P(x) + K(x,r)$$

Za $f : \Omega \to \bar{\mathbb{R}}$ označimo $f^+(x) = \max(0, f(x))$. Za funkcijo f^+ je mogoče pokazati

IZREK 13 *Naj bo $f : \mathbb{R}^n \to \mathbb{R}$ zvezno odvedljiva na $\mathbb{R}^n$. Tedaj je zvezno odvedljiva tudi funkcija $g(x) = (f^+(x))^2$ in velja*

$$\frac{\partial g(x)}{\partial x_i} = 2f^+(x)\frac{\partial f(x)}{\partial x_i}$$

za vsak $x \in \mathbb{R}^n$.

Courant-Beltramijeva kazenska funkcija

Od tu neposredno izhaja: Če so kriterijska funkcija in omejitve naloge (Φ, P, Min), $\Phi \subseteq \mathbb{R}^n$ zvezno odvedljive, je zvezno odvedljiva tudi zunanja kriterijska funkcija

$$Q(x, r) = P(x) + r \sum_{i \in I} (P_i^+(x))^2$$

Kazenski člen imenujemo tudi *Courant-Beltramijeva kazenska funkcija*.

Zvezna odvedljivost funkcije $Q(x, r)$ postane pomembna pri numeričnem reševanju nalog $(\Omega, Q(., r), \mathrm{Min})$, saj lahko uporabimo učinkovitejše postopke, ki temelje na odvodih.

Izrek o zunanjih kazenskih funkcijah

O zvezi med nalogami $(\Omega, Q(., r), \mathrm{Min}),\ r \in \mathbb{R}^+$ in nalogo (Φ, P, Min) je mogoče pokazati naslednji izrek:

IZREK 14 *Naj zunanja kazenska funkcije $K(x, r)$ zadošča za vsak $r \in \mathbb{R}^+$ in za vsak $x \in \Omega$ pogojem*

a. $K(x, r) \geq 0$

b. $K(x, r) = 0 \Leftrightarrow x \in \Phi$

c. $K(x, r)$ *je zvezna*

d. *kriterijska funkcija $P(x)$ je zvezna in je izpolnjen vsaj en od pogojev:*

  d1. $\|x\| \to \infty \Rightarrow P(x) \to \infty$

  d2. *Φ je omejena množica in $\|x\| \to \infty \Rightarrow K(x, r) \to \infty$*

Tedaj ima zaporedje $x^*(r), r \to \infty$ rešitev nalog $(\Omega, Q(., r), \mathrm{Min})$ vsaj eno stekališče in $K(x^*(r), r) \to 0$. Vsako stekališče zaporedja $x^*(r)$ pripada množici $\mathrm{Min}(\Phi, P)$.

Postopek za zunanje kazenske funkcije

Iz izreka izhaja naslednji postopek določanja rešitev naloge (Φ, P, Min):

> izberi začetni r (npr. $r := 1$);
> **loop**
>   določi $x^*(r) \in \mathrm{Min}(\Omega, Q(., r))$ z izbranim
>   postopkom reševanja nalog brez omejitev;
>   **if** $K(x^*(r), r)$ je dovolj majhna **exit**;
>   povečaj r (npr. $r := 10r$)
> **endloop**

Metoda zunanje točke – primer

Prikažimo metodo zunanje točke na nalogi (Φ, P, Min) za

$$\Phi = \{(x, y) \in \mathbb{R}^2 : 3x - 2y \geq 4\}$$

$$P(x, y) = x^2 + y^2$$

Nalogi pripada zunanja kriterijska funkcija

$$Q(x, y; r) = \begin{cases} x^2 + y^2 & 3x - 2y \geq 4 \\ x^2 + y^2 + r(4 + 2y - 3x)^2 & 3x - 2y < 4 \end{cases}$$

Določimo rešitev naloge $(\mathbb{R}^2, Q(., r), \mathrm{Min})$. V tem primeru lahko to naredimo analitično

$$\begin{aligned} \frac{\partial Q}{\partial x} = 0 \quad &= \quad 2x - 6r(4 + 2y - 3x) \\ \frac{\partial Q}{\partial y} = 0 \quad &= \quad 2y + 4r(4 + 2y - 3x) \end{aligned}$$

. . . Metoda zunanje točke – primer

oziroma, ko preuredimo

$$(1+9r)x - 6ry = 12r \quad \text{in} \quad -6rx + (1+4r)y = -8r$$

od koder dobimo

$$x^*(r) = \frac{12r}{1+13r} = \frac{12}{13+\frac{1}{r}} \quad \text{in} \quad y^*(r) = \frac{-8r}{1+13r} = \frac{-8}{13+\frac{1}{r}}$$

Točke $(x^*(r), y^*(r)) \notin \Phi$, saj je

$$3x^*(r) - 2y^*(r) = \frac{52r}{1+13r} = \frac{4}{1+\frac{1}{13r}} < 4$$

Stekališče zaporedja $(x^*(r), y^*(r))$ pa je $(\frac{12}{13}, -\frac{8}{13}) \in \Phi$.

…Metoda zunanje točke – primer

r	$x^*(r)$	$y^*(r)$
1	0.8571	-0.5714
2	0.8884	-0.5926
4	0.9057	-0.6038
8	0.9143	-0.6095
16	0.9187	-0.6124
32	0.9209	-0.6139
64	0.9220	-0.6146
∞	0.9231	-0.6154

Metoda notranje točke

Glavna slabost metode zunanje točke je, da se optimalnim točkam približujemo od zunaj – zunaj množice dopustnih rešitev Φ. Fiacco in McCormick (1968) sta predlagala nov pristop, ki odpravlja tudi to slabost za naloge (Φ, P, Min), za katere je notranjost množice Φ

$$\mathrm{Int}\Phi = \{x \in \Omega : P_i(x) < 0, i \in I\}$$

neprazna. Točke, ki niso notranje, sestavljajo rob množice Φ

$$\partial\Phi = \Phi \setminus \mathrm{Int}\Phi$$

Vpeljala sta *zaprečno funkcijo*

$$B(x) = \begin{cases} -\sum_{i \in I} \frac{1}{P_i(x)} & x \in \mathrm{Int}\Phi \\ \infty & \text{sicer} \end{cases}$$

in kazensko kriterijsko funkcijo

$$Q(x,t) = P(x) + tB(x), \qquad t \geq 0$$

...Metoda notranje točke

Za $x \in \mathrm{Int}\Phi$ je $B(x) \geq 0$. Funkcija $Q(x,t)$ na $\mathrm{Int}\Phi$ ohranja konveksnost in zveznost.

Kasneje so si izmislili še druge zaprečne funkcije, kot je na primer

$$B(x) = \begin{cases} -\sum_{i \in I} \ln(-P_i(x)) & x \in \mathrm{Int}\Phi \\ \infty & \text{sicer} \end{cases}$$

Zopet je mogoče pokazati:

... Metoda notranje točke

IZREK 15 *Naj bo* Φ *zaprta množica,* $\mathrm{Int}\Phi \neq \emptyset$ *in vsak* $x \in \Phi$ *je stekališče notranjih točk. Naj notranja kazenska funkcije* $B(x)$ *zadošča pogojem*

a. $B(x) \geq 0, \ x \in \mathrm{Int}\Phi$

b. $B(x)$ *je zvezna na* $\mathrm{Int}\Phi$

c. $x \to \partial\Phi \Rightarrow B(x) \to \infty$

d. kriterijska funkcija $P(x)$ *je zvezna in je izpolnjen vsaj en od pogojev*

  d1. $\|x\| \to \infty \Rightarrow P(x) \to \infty$

  d2. Φ *je omejena množica*

Tedaj ima zaporedje $x^*(t), t \to 0$ *rešitev nalog* $(\Omega, Q(., t), \mathrm{Min})$ *vsaj eno stekališče in* $tB(x^*(t)) \to 0$*. Vsako stekališče zaporedja* $x^*(t)$ *pripada množici* $\mathrm{Min}(\Phi, P)$.

Metoda notranje točke – primer

Metodo notranje točke si oglejmo na preprosti nalogi (Φ, P, Min)

$$\Phi = \{x \in \mathbb{R} : x \geq 2\} \quad \text{in} \quad P(x) = x$$

Nalogi pripada notranja kriterijska funkcija

$$Q(x; t) = x + \frac{t}{x-2}$$

Rešitev naloge $(\mathbb{R}, Q(., t), \mathrm{Min})$ lahko določimo analitično

$$\frac{\partial Q}{\partial x} = 0 = 1 - \frac{t}{(x-2)^2}$$

od koder dobimo $x(t) = 2 \pm \sqrt{t}$ oziroma $x^*(t) = 2 + \sqrt{t} \in \Phi$. Ker je $\frac{\partial^2 Q}{\partial x^2}(x^*(t)) > 0$, je v tej točki minimum. Stekališče zaporedja $(x^*(t)),\ t \to 0$ je točka $2 \in \Phi$.

... Metoda notranje točke – primer

t	$x^*(t)$	$Q(x^*(t), t)$
1	3	4
0.25	2.5	3
0.01	2.1	2.2
0.0001	2.01	2.02
0	2	2