OVM - nahradna hodina 18.4.

20 April 2024 09:54

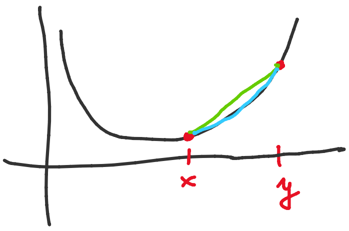
Logist. reg.: $y_i \sim Alt(\pi_i)$, $y_i \in \{1, 0\}$

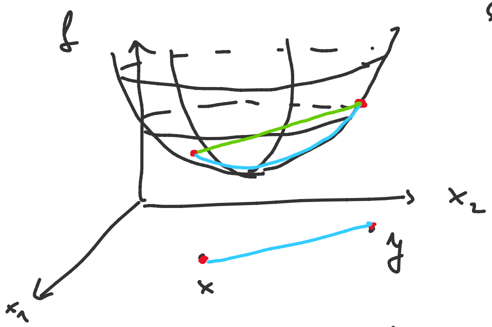
π_i : pp. úspechu

- met. doterat: ťažké dôkazy konvergencie
  : fungujú pre <u>konvexné fcie</u>: $f: \mathbb{R}^n \to \mathbb{R}$ je konv. ak:

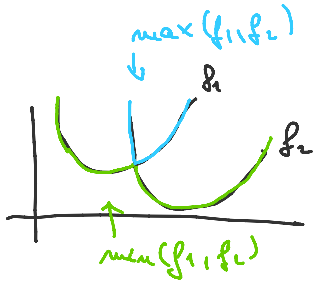
$$\forall x,y \in \mathbb{R}^n\ \forall \alpha \in [0,1]: \ f(\underbrace{\alpha x + (1-\alpha) y}_{\text{konv. komb. } x,y}) \le \underbrace{\alpha f(x) + (1-\alpha) f(y)}_{\text{konv. komb. } f(x),\, f(y)}$$

1-rozmer





- ak $\exists \nabla^2 f(x)\ \forall x$: $\quad f$ - konv. $\iff \underbrace{\nabla^2 f(x) \succeq 0}_{\nabla^2 f(x) \text{ je k.s.d.}} \quad \forall x \in \mathbb{R}^n$
- f - konv. $\Rightarrow$ každé lok. min. na $\mathbb{R}^n$ je aj glob. min. na $\mathbb{R}^n$
  $\Rightarrow$ ak $\nabla f(\tilde{x}) = 0$, potom $\tilde{x}$ - glob. min.
- súčet konv. fcií je konv. fcia
- maximum konv. fcií je konv. fcia
  $\to$ minimum konk fcií je konk. fcia



<u>Ohraničená optimalizácia</u>

$$\left\{ \begin{array}{l} \min\limits_{x \in \mathbb{R}^n} f(x) \\ x \in M \end{array} \right\} \qquad f: \mathbb{R}^n \to \mathbb{R}, \quad M \subseteq \mathbb{R}^n$$

$x \in M$: prípustné rieš.

$x^* \in M$ a $f(x^*) \le f(x)\ \forall x \in M \Rightarrow x^*$: opt. rieš.

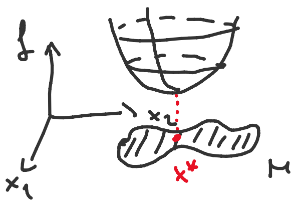
M: množ. príp. riešení

praktický zápis:
$$\left\{ \begin{array}{ll} \min\limits_{x} f(x) & \\ f_i(x) \le 0 & \forall i = 1, \dots, m \\ g_j(x) = 0 & \forall j = 1, \dots, p \end{array} \right\}$$

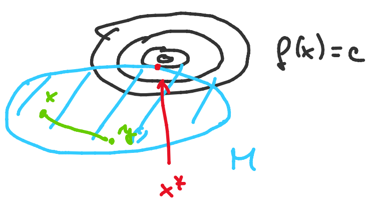
Pr.:
$$\left\{ \begin{array}{l} \min\ 2x_1^2 + x_1 x_2 \\ 1 \le x_1 \le 2 \\ x_1^2 + x_2^2 = 1 \end{array} \right\} \to \quad 1 \le x_1: \ \underbrace{1 - x_1 \le 0}_{f_1(x) \le 0} \qquad x_1 \le 2: \ \underbrace{x_1 - 2 \le 0}_{f_2(x) \le 0}$$

$x_1^2 + x_2^2 - 1 = 0$: $g_1(x) = 0$

$\Updownarrow$
$$\left\{ \begin{array}{ll} \min\ 2x_1^2 + x_1 x_2 & \\ 1 - x_1 \le 0 & \leftarrow f_1 \\ x_1 - 2 \le 0 & \leftarrow f_2 \\ x_1^2 + x_2^2 - 1 = 0 & \leftarrow g_1 \end{array} \right\}$$



vrstevnice:



$\Rightarrow$ komplikované <
- úč. fcia - môže byť nepekná
- M - môže byť nepekná

pekná úč. fcia: konvexná fcia

pekná M: <u>konvexná množ.</u>: $\forall x,y \in M\ \forall \alpha \in [0,1]: \ \alpha x + (1-\alpha) y \in M$

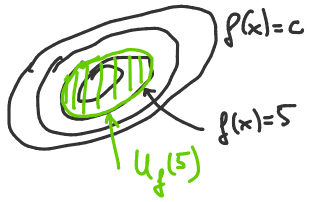
dobre riešiteľné: konvexná optimalizácia (konvexné programovanie) : KP

KP: $\left\{ \begin{array}{l} \min\limits_x f(x) \\ x \in M \end{array} \right\}$ f - konv. fcia, M - konv. množ. resp. $\left\{ \begin{array}{l} \max\limits_x f(x) \\ x \in M \end{array} \right\}$ f - konk. fcia, M - konv. množ.

chceli by sme: $(*) \left\{ \begin{array}{l} \min\limits_x f(x) \\ f_i(x) \le 0 \ \forall i \\ g_j(x) = 0 \ \forall j \end{array} \right\}$ f - konv., f_i ?, g_j ?

platí: ak f - konv. $\Rightarrow$ každá jej podúrovňová množ. je konv.

$\hookrightarrow U_f(c) = \{ x \in \mathbb{R}^n \mid f(x) \le c \}$



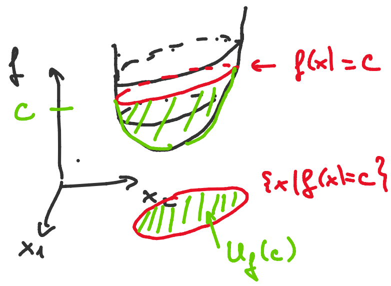
Dôkaz: f - konv.

$x, y \in U_f(c), \ \alpha \in [0,1] \overset{CHD}{\Rightarrow} \alpha x + (1-\alpha) y \in U_f(c)$

$\downarrow$

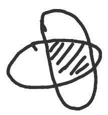
$f(x) \le c, \ f(y) \le c \overset{CHD}{\Rightarrow} f(\alpha x + (1-\alpha) y) \le c$

$f(\alpha x + (1-\alpha) y) \underset{f\text{-konv.}}{\le} \alpha \underbrace{f(x)}_{\le c} + (1-\alpha) \underbrace{f(y)}_{\le c} \le \alpha c + (1-\alpha) c = c$ ✓ ■



$(*)$: $M = \{ x \mid f_i(x) \le 0 \ \forall i, \ g_j(x) = 0 \ \forall j \} = \left(\bigcap\limits_{i=1}^{m} \underbrace{\{ x \mid f_i(x) \le 0 \}}_{U_{f_i}(0)} \right) \cap \left(\bigcap\limits_{j=1}^{p} \{ x \mid g_j(x) = 0 \} \right)$

platí: prienik konv. množ. je konv. množina



$\Downarrow$

zvolíme f_i - konv. $\forall i \Rightarrow$ bude konv. množ.

$g_j(x) = 0 \Leftrightarrow g_j(x) \le 0 \qquad g_j(x) \le 0 \leftarrow \{ x \mid g_j(x) \le 0 \}$ bude konv. ak g_j - konv.

$\qquad g_j(x) \ge 0 \Leftrightarrow -g_j(x) \le 0 \leftarrow \{ x \mid -g_j(x) \le 0 \}$ bude konv. ak $-g_j$ - konv.

$(*)$: $\left\{ \begin{array}{l} \min f(x) \\ f_i(x) \le 0 \ \forall i \\ g_j(x) \le 0 \ \forall j \\ -g_j(x) \le 0 \ \forall j \end{array} \right\}$ $\begin{array}{l} \leftarrow f_i \text{ - konv.} \\ \leftarrow g_j \text{ - konv.} \\ \leftarrow -g_j \text{ - konv.} \Leftrightarrow g_j \text{ - konk.} \end{array}$ $\Big\}$ g_j - lineárna $\forall j$ (afinná) : $g_j(x) = a_j^T x - b_j \ \forall j = 1, \dots, p$

$a_j \in \mathbb{R}^n, \ b_j \in \mathbb{R}$

$$\begin{array}{c} a_1^T x - b_1 = 0 \\ \vdots \\ a_p^T x - b_p = 0 \end{array} \Leftrightarrow \begin{array}{c} a_1^T x = b_1 \\ \vdots \\ a_p^T x = b_p \end{array} \Leftrightarrow Ax = b, \qquad A = \begin{bmatrix} a_1^T \\ \vdots \\ a_p^T \end{bmatrix} \in \mathbb{R}^{p \times n}, \ b \in \mathbb{R}^p$$

$\Rightarrow$ záver:

$\left\{ \begin{array}{l} \min\limits_x f(x) \\ f_i(x) \le 0 \ \forall i = 1, \dots, m \\ Ax = b \end{array} \right\}$ $f, f_1, \dots, f_m$ - konv., $A \in \mathbb{R}^{p \times n}$, $b \in \mathbb{R}^n$ $\searrow$ je úlohou KP

– štandardná úloha KP

- toto nie je nutná podm. na to, aby úloha bola KP
- ale: každá úloha KP sa dá zapísať v tomto tvare

Pr.: $\left\{ \min\limits_x x_1^2 + x_2^2 \right.$ $\Big\}$ $\leftarrow$ je konv. ✓

Príp. rieš.:

$x_1 \dots \le 0 \Leftrightarrow x_1 \le 0$

Pr.: $\begin{cases} \min\limits_x \; x_1^2+x_2^2 \\ \dfrac{x_1}{1+x_2^2} \le 0 \\ (x_1+x_2)^2 = 0 \end{cases}$ $\begin{array}{l} \leftarrow \text{je konv. } \checkmark \\ \leftarrow \text{nie je konv.} \\ \leftarrow \text{nie je lin. } \times \end{array}$

Príp. rieš.:

$\dfrac{x_1}{\underbrace{1+x_2^2}_{>0}} \le 0 \iff x_1 \le 0$ ↑ konv. fcia ✓

$(x_1+x_2)^2 = 0 \iff x_1+x_2 = 0$ ↑ lin. fcia ✓

$\Updownarrow$

$\begin{cases} \min\limits_x \; x_1^2+x_2^2 \\ x_1 \le 0 \\ x_1+x_2 = 0 \end{cases}$ $\begin{array}{l} \leftarrow \text{je konv. } \checkmark \\ \leftarrow f_1 - \text{je konv. } \checkmark \\ \leftarrow g_1 - \text{je lin. } \checkmark \end{array}$

• v tvare štandardnej úlohy KP $\Rightarrow$ pôvodná úloha je konvexná

KP: pekné vlastnosti – "zaručene" fungujúce algoritmy
– každé lok. min. v M je aj globálnym v M

Dualita: pre všeob. úlohu ohr. opt.

$\begin{cases} \min f(x) \\ f_i(x) \le 0 \quad \forall i = 1 \dots m \\ g_j(x) = 0 \quad \forall j = 1 \dots p \end{cases}$ (1) $\begin{array}{l} \leftarrow \lambda \in \mathbb{R}^m \\ \leftarrow \nu \in \mathbb{R}^p \end{array}$ } Lagrangeove multiplikátory

Lagrangeova fcia: $\mathcal{L}(x,\lambda,\nu) = f(x) + \underbrace{\sum_{i=1}^m \lambda_i f_i(x) + \sum_{j=1}^p \nu_j g_j(x)}_{\text{penalty za porušenie ohr.}}$ $\quad \lambda_i \ge 0 \; \forall i$, $\nu_j \in \mathbb{R} \; \forall j$

namiesto (1) riešime $\{ \min\limits_x \mathcal{L}(x,\lambda,\nu) \}$ (2) – voľná opt.

• ozn. optimálnu hodnotu (2): $h(\lambda,\nu) = \min\limits_x \mathcal{L}(x,\lambda,\nu)$
$= \min\limits_x \Big(\underbrace{f(x) + \sum_i \lambda_i f_i(x) + \sum_j \nu_j g_j(x)}_{\text{lin. fcia v } \lambda,\nu \;\Rightarrow\; \text{aj konkávne v } \lambda,\nu} \Big)$

vieme: min. konk. fcií je konkávna
$\Rightarrow h(\lambda,\nu)$ – konk. fcia (vždy, bez ohľadu na tvar f, f_i, g_j)

Tvrdenie:
$h(\lambda,\nu) \le f(x^*) \quad \forall \lambda \ge 0_m, \nu \in \mathbb{R}^p$, pre x^* – opt. rieš. (1)

Dôkaz: $h(\lambda,\nu) = \min\limits_x \big(f(x) + \sum_i \lambda_i f_i(x) + \sum_j \nu_j g_j(x) \big) \le f(x^*) + \underbrace{\sum_i \underbrace{\lambda_i}_{\ge 0} \underbrace{f_i(x^*)}_{\le 0}}_{\le 0} + \cancel{\sum_j \nu_j \underbrace{g_j(x^*)}_{=0}} \le f(x^*)$ ■

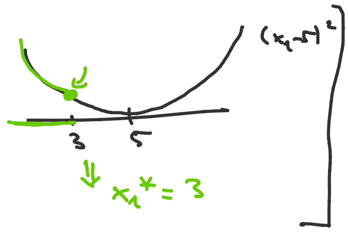
$\lambda_i \ge 0 \; \forall i$
$\nu_j \in \mathbb{R}$

min. je $\le$ ak fcná hodnota v ľubovoľnom x (aj v x^*)

x^* – opt. pre (1) $\Rightarrow$ aj príp. pre (1) $\Rightarrow$ $f_i(x^*) \le 0 \; \forall i$, $g_j(x^*) = 0 \; \forall j$

Pr.: $\begin{cases} \min\limits_x \; (x_1-5)^2 + x_2^2 \\ x_1 \le 3 \end{cases} \leftarrow \lambda$
↑ $x_1 - 3 \le 0$

Riešenie: $x_2^* = 0 \Rightarrow$ ostáva $(x_1-5)^2 \to \min\limits_{x_1 \le 3}$
$\Downarrow$
$x_1^* = 3$
$\Rightarrow x^* = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$
$f(x^*) = 4 + 0 = 4$



$\mathcal{L}(x,\lambda) = (x_1-5)^2 + x_2^2 + \lambda(x_1-3) \to \min\limits_x$

$\dfrac{\partial \mathcal{L}}{\partial x_1} = 2(x_1-5) + 0 + \lambda = 0 \quad /:2$
$x_1 - 5 = -\frac{\lambda}{2}$

$\dfrac{\partial \mathcal{L}}{\partial x_2} = 0 + 2x_2 + 0 = 0$
$x_2 = 0$

$$\frac{\partial \mathcal{L}}{\partial x_1} = 2(x_1-5)+0+\lambda = 0 \quad /:2$$

$$x_1-5 = -\frac{\lambda}{2}$$

$$x_1 = 5-\frac{\lambda}{2}$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = 0+2x_2+0=0$$

$$x_2 = 0$$

$$\Rightarrow h(\lambda) = \mathcal{L}\left(\begin{pmatrix}5-\frac{\lambda}{2}\\0\end{pmatrix},\lambda\right) = \left(5-\frac{\lambda}{2}-5\right)^2+0+\lambda\left(\overbrace{5-\frac{\lambda}{2}-3}^{2-\frac{\lambda}{2}}\right) = \frac{\lambda^2}{4}+2\lambda-\frac{\lambda^2}{2} = -\frac{\lambda^2}{4}+2\lambda$$

Twierdzenie : $h(\lambda) \le \underbrace{f(x^*)}_{4} \quad \forall \lambda \ge 0$

np. $\lambda=0$: $h(\lambda) = 0+0=0 \le 4$ ✓

$\lambda = 1$: $h(\lambda) = -\frac{1}{4}+2 = \frac{7}{4} \le 4$ ✓

$\lambda = 4$: $h(\lambda) = -\frac{4^{\cancel{2}}}{\cancel{4}}+2\cdot 4 = 4 \le 4$ ✓

$\lambda = 3$ ... ≤ 4 ...