Lagrangeovy multiplikátory

Hledáme lokální extrémy funkce $f: \mathbb{R}^n \to \mathbb{R}$ na množině $X = \{x \in \mathbb{R}^n \mid g(x) = 0\}$

$$\left.\begin{array}{l} \min/\max \ f(x_1, x_2, \dots, x_n) \\ \text{z.p.} \ \ g(x_1, x_2, \dots, x_n) = 0 \quad i = 1, \dots, m \end{array}\right\} \begin{array}{l} \text{extrémy funkce } f \\ \text{vázané rovnostmi } g(x) = 0 \end{array}$$

- regulární bod zobrazení $g: \mathbb{R}^n \to \mathbb{R}^m$
  - i) zobrazení g je v bodě x spojitě diferencovatelné
  - ii) Jacobiho matice $g'(x)$ má lineárně nezávislé řádky

- podmínka prvního řádu

  $f: \mathbb{R}^n \to \mathbb{R}$ je diferencovatelná v bodě $x \in \mathbb{R}^n$, x je regulární bod zobrazení $g: \mathbb{R}^n \to \mathbb{R}^m$.
  Jestliže x je lok. extrém f vázaný $g(x) = 0$, pak $\nabla f(x) \perp \operatorname{null} g'(x)$

  - Podmínku $\nabla f(x) \perp \operatorname{null} g'(x)$ lze zapsat jako $\nabla f(x) \in \operatorname{span}\{\nabla g_1(x), \dots, \nabla g_m(x)\}$,
    tzn. $\nabla f(x) = \lambda_1 \nabla g_1(x) + \lambda_2 \nabla g_2(x) + \dots + \lambda_m \nabla g_m(x)$

$$\boxed{\begin{aligned} f'(x) + \lambda^T g'(x) &= 0 \\ g(x) &= 0 \end{aligned}}$$

($\lambda_i \in \mathbb{R}$, proto nezáleží na tom, zda použijeme $+\lambda_i$ nebo $-\lambda_i$)

  - kompaktněji: $\mathcal{L}(x, \lambda) = f(x) + \lambda^T g(x)$

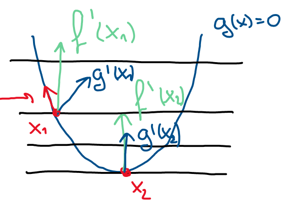
$$\boxed{\begin{aligned} \mathcal{L}_x(x, \lambda) &= 0 \\ \mathcal{L}_\lambda(x, \lambda) &= 0 \end{aligned}}$$

intuice: $f(x,y) = y$; $g(x,y) = y - x^2$ $\qquad\qquad$ $g(x,y) = y^3 - x^2$

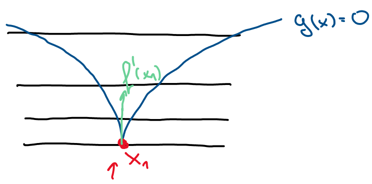
$\min \ y$ $\qquad$ $\min \ y$

z.p. $y - x^2 = 0$ $\qquad$ z.p. $y^3 - x^2 = 0$

f' a g' nejsou v x_1 rovnoběžné, pohybem po null $g'(x)$ lze změnit hodnotu f



f' a g' jsou v x_2 rovnoběžné, f je na null $g'(x_2)$ konst.



x_1 není regulární bod $g(x)$, věta nelze použít

$\left(g'(x,y) = [-2x \ \ 3y^2]\big|_{(0,0)} = [0 \ \ 0] \right.$
$\left. \hookrightarrow \text{nemá lin. nez. řádky v } (0,0) \right)$

- podmínka druhého řádu

  Hessova matice $\dfrac{\partial^2 \mathcal{L}(x, \lambda)}{\partial x^2} = f''(x) + \sum_{i=1}^{m} \lambda_i g_i''(x)$

  $f: \mathbb{R}^n \to \mathbb{R}$ a $g: \mathbb{R}^n \to \mathbb{R}^m$ dvakrát dif. v bodě $x \in \mathbb{R}^n$, x je regulární bod g
  Pak x je lok. min/max f vázané $g(x) = 0$ $\Rightarrow$ $\exists \lambda: \mathcal{L}'(x, \lambda) = 0$ a Hessova matice je pos./neg. semidef na null $g'(x)$

  $\exists \lambda: \mathcal{L}'(x, \lambda) = 0$ a Hessova matice je pos./neg. def. na null $g'(x)$ $\Rightarrow$ x je ostré lok. min/max f vázané $g(x) = 0$

11.12) Rozdělení pravděpodobnosti diskrétní náhodné proměnné je funkce $p: \{1,\dots,n\} \to \mathbb{R}_+$ splňující $\sum_{x=1}^{n} p(x) = 1$

a) Najděte rozdělení s maximální entropií $-\sum_{x=1}^{n} p(x) \log p(x)$

$$\left.\begin{aligned} &\max_{p\in\mathbb{R}^n} \sum_{x=1}^{n} p_x \log p_x \\ &\text{z.p.} \ \sum_{x=1}^{n} p_x = 1 \end{aligned}\right\} \mathcal{L}(p,\lambda) = \sum_{x=1}^{n} p_x \log p_x + \lambda\left(\sum_{x=1}^{n} p_x - 1\right)$$

$$\left.\begin{aligned} \frac{\partial \mathcal{L}}{\partial p_x} &= \log p_x + 1 + \lambda \\ \frac{\partial \mathcal{L}}{\partial \lambda} &= \sum_{x=1}^{n} p_x - 1 \end{aligned}\right\} \overset{!}{=} 0 \qquad \begin{aligned} &\log p_x + 1 + \lambda = 0 \quad \forall x \in \{1,\dots,n\} \\ &\log p_x = -(1+\lambda) \Rightarrow \log p_x = \log p_y \ \forall x,y \Rightarrow \underline{p_x = p_y \ \forall x,y} \\ &p_x = c \ \forall x: \\ &\sum_{x=1}^{n} p_x - 1 = 0 \Rightarrow n \cdot c = 1 \Rightarrow c = \frac{1}{n} \end{aligned}$$

Pro rozdělení s maximální entropií platí $p_x = \frac{1}{n} \ \forall x \in \{1,\dots,n\} \Rightarrow$ <u>rovnoměrné rozdělení</u>

b) Dokažte Gibbsovu nerovnost: pro každá dvě rozdělení platí

$$\sum_{x=1}^{n} p_x \log q_x \le \sum_{x=1}^{n} p_x \log p_x \quad \text{, rovnost pouze pro } p=q$$

Zkusíme maximalizovat $f(q) = \sum_{x=1}^{n} p_x \log q_x$ za podmínky $\sum_{x=1}^{n} q_x = 1$

($q_x \ge 0$ splněno implicitně, pro $\lim_{q_x \to 0^+} f(q) = -\infty$)

$$\left.\begin{aligned} &\max_{q\in\mathbb{R}^n} f(q) \\ &\text{z.p.} \ \sum_{x=1}^{n} q_x = 1 \end{aligned}\right\} \mathcal{L}(q,\lambda) = \sum_{x=1}^{n} p_x \log q_x + \lambda\left(\sum_{x=1}^{n} q_x - 1\right)$$

$$\left.\begin{aligned} \frac{\partial \mathcal{L}}{\partial q_x} &= \frac{p_x}{q_x} + \lambda \\ \frac{\partial \mathcal{L}}{\partial \lambda} &= \sum_{x=1}^{n} q_x - 1 \end{aligned}\right\} \overset{!}{=} 0 \qquad \begin{aligned} \lambda q_1 &= -p_1 \\ \lambda q_2 &= -p_2 \\ &\vdots \end{aligned}$$

$$\oplus$$

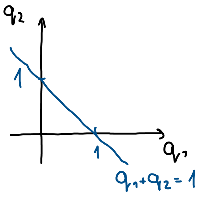
$$\lambda \sum q_x = -\sum p_x$$

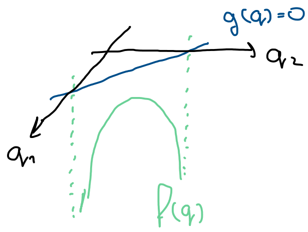
$$\lambda = -1 \Rightarrow q_x = p_x \ \forall x \in \{1,\dots,n\}$$

Bod podezřelý z extrému (vázaný rovností $\sum q_x = 1$): $q_x = p_x \ \forall x \in \{1,\dots,n\}$.

Protože je to jediný podezřelý bod, všechny body $\{q \in \mathbb{R}^n \mid g(q) = 0\}$ jsou regulární a funkce $f(q)$ klesá na hranici svého def. oboru do $-\infty$, bude to lokální maximum.

Pro $q \in \mathbb{R}^2$





11.17) Minimalizujte $x^T A x$ za podmínky $b^T x = 1$, kde A je symetrická pos. def.

$$\left.\begin{aligned} &\min x^T A x \\ &\text{z.p.} \ b^T x = 1 \end{aligned}\right\} \mathcal{L}(x,\lambda) = x^T A x + \lambda(b^T x - 1)$$

$$\left.\begin{aligned} \frac{\partial \mathcal{L}}{\partial x} &= x^T(A^T + A) + \lambda b^T = 2x^T A + \lambda b^T \\ \frac{\partial \mathcal{L}}{\partial \lambda} &= b^T x - 1 \end{aligned}\right\} \overset{!}{=} 0 \qquad \begin{aligned} 2x^T A + \lambda b^T &= 0 \\ b^T x &= 1 \end{aligned}$$

$$2x^T A = -\lambda b^T \ / \cdot A^{-1}$$

$$x^T = -\tfrac{1}{2}\lambda b^T A^{-1}$$

$$x = -\tfrac{1}{2}\lambda A^{-1} b$$

dosazením do $b^T x = 1$

$$b^T x = -\tfrac{1}{2}\lambda b A^{-1} b = 1$$

$$\lambda = -\frac{2}{b^T A^{-1} b}$$

$$x^T = \frac{b^T A^{-1}}{b^T A^{-1} b} \Rightarrow x = \frac{A^{-1} b}{b^T A^{-1} b}$$

11.18) Minimalizujte x^TCx za podmínky $Ax=b$, kde A má LN řádky a C je symetrická pos. def.

$$\left.\begin{array}{l} \min\ x^TCx \\ \text{z.p. } Ax=b \end{array}\right\} \mathcal{L}(x,\lambda) = x^TCx + \underset{\text{vektor!}}{\lambda^T}(Ax-b)$$

$$\left.\begin{array}{l} \frac{\partial \mathcal{L}}{\partial x} = x^T(C+C^T) + \lambda^T A = 2x^TC + \lambda^T A \\ \frac{\partial \mathcal{L}}{\partial \lambda} = x^TA^T - b^T \end{array}\right\} \overset{!}{=} 0$$

$$2x^TC + \lambda^T A = 0$$
$$Ax = b$$

$$2x^TC = -\lambda^T A \quad / \cdot C^{-1}$$
$$x^T = -\frac{1}{2}\lambda^T A C^{-1} \quad /^T$$
$$x = -\frac{1}{2}C^{-1}A^T\lambda$$

$$A(-\frac{1}{2}C^{-1}A^T\lambda) = b$$
$$-\frac{1}{2}AC^{-1}A^T\lambda = b$$
$$\lambda = -2(AC^{-1}A^T)^{-1}b$$

$$x = C^{-1}A^T(AC^{-1}A^T)^{-1}b$$