Optimalizace

Konvexní funkce

Tomáš Kroupa Tomáš Werner

2024

Fakulta elektrotechnická
ČVUT v Praze

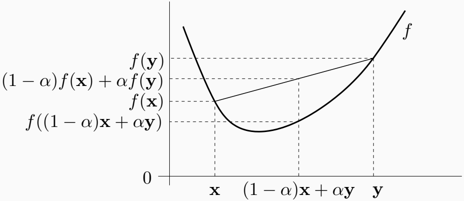
Konvexní funkce

Definice

Funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ je konvexní na konvexní množině $X \subseteq \mathbb{R}^n$, jestliže pro všechna $\mathbf{x}, \mathbf{y} \in X$ a každé $\alpha \in [0,1]$ platí nerovnost

$$f((1-\alpha)\mathbf{x} + \alpha\mathbf{y}) \leq (1-\alpha)f(\mathbf{x}) + \alpha f(\mathbf{y}).$$

Funkce f je konkávní na X, je-li $-f$ je konvexní na X.

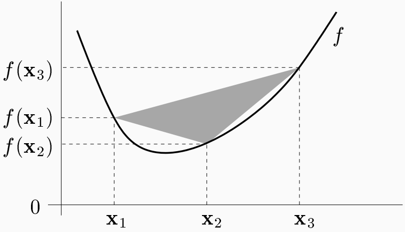


Jensenova nerovnost

Tvrzení

Nechť f je konvexní funkce. Pro všechna $\mathbf{x}_1, \ldots, \mathbf{x}_k \in X$ a všechna $\alpha_1, \ldots, \alpha_k \geq 0$ splňující $\alpha_1 + \cdots + \alpha_k = 1$ platí

$$f(\alpha_1 \mathbf{x}_1 + \cdots + \alpha_k \mathbf{x}_k) \leq \alpha_1 f(\mathbf{x}_1) + \cdots + \alpha_k f(\mathbf{x}_k)$$

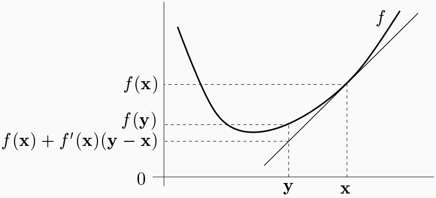


Konvexní diferencovatelné funkce

Podmínka prvního řádu

Nechť f je diferencovatelná. Funkce f je konvexní, právě když pro všechna $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ platí nerovnost

$$f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y} - \mathbf{x}) \leq f(\mathbf{y}).$$



Podmínka druhého řádu

Nechť f je dvakrát diferencovatelná. Funkce f je konvexní, právě když je pro každé $\mathbf{x} \in \mathbb{R}^n$ Hessova matice $f''(\mathbf{x})$ pozitivně semidefinitní.

Příklady konvexních funkcí (1)

Definice

Funkce

$$\mathbf{x} \in \mathbb{R}^n \mapsto \|\mathbf{x}\| \in [0, \infty)$$

je norma, pokud platí:

- $\|\mathbf{x}\| = 0$ právě tehdy, když $\mathbf{x} = \mathbf{0}$
- $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$ pro každé $\alpha \in \mathbb{R}$ a $\mathbf{x} \in \mathbb{R}^n$
- $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ pro každé $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

p-normy

$$\|\mathbf{x}\|_p = \left(|x_1|^p + \cdots + |x_n|^p\right)^{\frac{1}{p}}, \qquad p \in [1, \infty]$$

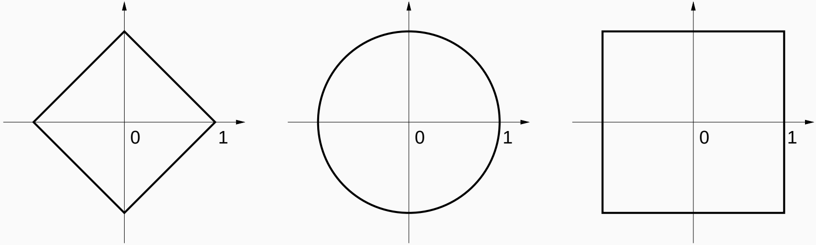
Tři důležité normy

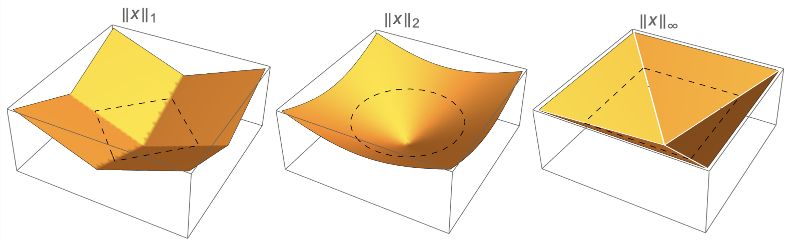
Manhattanská $\|\mathbf{x}\|_1 = |x_1| + \cdots + |x_n|$

Eukleidovská $\|\mathbf{x}\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$

Čebyševova $\|\mathbf{x}\|_\infty = \max\{|x_1|, \ldots, |x_n|\} = \lim\limits_{p \to \infty} \|\mathbf{x}\|_p$

Jednotkové kružnice p-norem pro $p = 1, 2, \infty$





$$\|\mathbf{x}\|_1 \geq \|\mathbf{x}\|_2 \geq \|\mathbf{x}\|_\infty$$

Příklady konvexních funkcí (2)

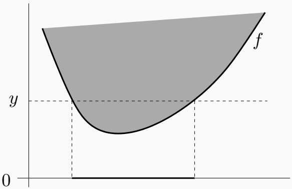
- $f(x) = e^{ax}$ pro $a \in \mathbb{R}$
- $f(x) = x \log_2 x$, kde $x > 0$ a $f(0) = 0$ pro $x = 0$
- $f(\mathbf{x}) = \ln(e^{x_1} + \cdots + e^{x_n})$
- $f(\mathbf{x}) = \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$, kde $\mathbf{A}$ je pozitivně semidefinitní
- $f(\mathbf{x}) = \mathbf{a}^\mathsf{T}\mathbf{x} + b$

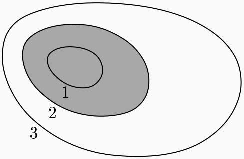
Nad grafem funkce a pod vrstevnicí

- Epigraf funkce f je množina $\{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid f(\mathbf{x}) \leq y\}$.
- Subkontura výšky y je množina $\{\mathbf{x} \in \mathbb{R}^{n} \mid f(\mathbf{x}) \leq y\}$.

Věta

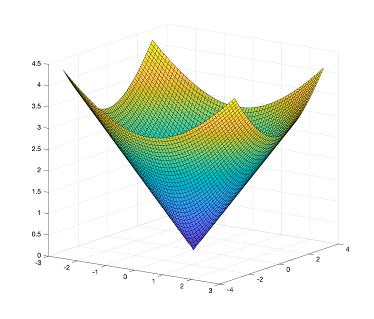
- f je konvexní, právě když její epigraf je konvexní množina.
- Každá subkontura konvexní funkce je konvexní množina.

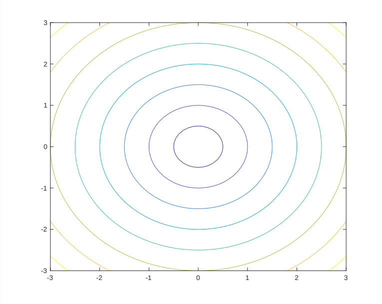




Příklad: Eukleidovská norma

- Epigraf normy $\|.\|_2$ na $\mathbb{R}^n$ je $\{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid \|\mathbf{x}\|_2 \leq y\}$
- Subkontury jsou kruhy o poloměru y





Epigrafový tvar úlohy minimalizace

Každá úloha tvaru

Úloha 1

Minimalizuj $f(\mathbf{x})$ za podmínky $\mathbf{x} \in \mathbb{R}^n$

má stejnou optimální hodnotu jako úloha s lineární účelovou funkcí

Úloha 2

Minimalizuj y za podmínek $f(\mathbf{x}) \leq y$, $\mathbf{x} \in \mathbb{R}^n, y \in \mathbb{R}$

Operace zachovávající konvexitu (1)

Nezáporné lineární kombinace

Jsou-li $g_1, \dots, g_k \colon \mathbb{R}^n \to \mathbb{R}$ konvexní funkce a $\alpha_1, \dots, \alpha_k \geq 0$, pak je funkce $f = \alpha_1 g_1 + \dots + \alpha_k g_k$ konvexní.

Příklad (Shannonova entropie)

Funkce $H(\mathbf{p}) = -\sum_{i=1}^{n} p_i \log_2 p_i$ je konkávní na množině diskrétních pravděpodobnostních rozdělení

$$\{\mathbf{p} \in \mathbb{R}^n \mid p_1 + \dots + p_n = 1,\ p_1, \dots, p_n \geq 0\}.$$

Operace zachovávající konvexitu (2)

Skládání funkcí

Následující funkce jsou konvexní:

- $h = g \circ f$, kde $f\colon \mathbb{R} \to \mathbb{R}$ je konvexní a $g\colon \mathbb{R} \to \mathbb{R}$ je konvexní a neklesající.
- $h(\mathbf{x}) = g(\mathbf{A}\mathbf{x} + \mathbf{b})$, kde $g\colon \mathbb{R}^m \to \mathbb{R}$ je konvexní.

Příklad

Funkce $f\colon \mathbb{R}^{2n} \to \mathbb{R}$ daná pomocí $f(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$, kde $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ a $\|.\|$ je norma, je konvexní.

Operace zachovávající konvexitu (3)

Věta

Nechť $g_i \colon \mathbb{R}^n \to \mathbb{R}$ jsou konvexní funkce pro všechna $i \in I$. Pak

$$f(\mathbf{x}) := \max_{i \in I} g_i(\mathbf{x})$$

je konvexní funkce, existuje-li pro každé $\mathbf{x} \in \mathbb{R}^n$ maximum výše.

Příklady

- $f(\mathbf{x}) = \max_{i=1}^{k} (\mathbf{a}_i^\mathsf{T} \mathbf{x} + b_i)$
- $f(\mathbf{x}) = \max_{\mathbf{y} \in C} \|\mathbf{x} - \mathbf{y}\|$ pro nějakou množinu $C \subseteq \mathbb{R}^n$
- $f(\mathbf{c}) = \max \{\mathbf{c}^\mathsf{T} \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$