

# Konvexní funkce

$\$ \$ \backslash \mathrm{def} \backslash \mathrm{scal} \# 1 \# 2 \{ \backslash \mathrm{langle} \# 1, \# 2 \backslash \mathrm{rangle} \} \backslash \mathrm{def} \backslash \mathrm{norm} \# 1 \{ \backslash \mathrm{left} \backslash \mathrm{Vert} \# 1 \backslash \mathrm{right} \backslash \mathrm{rVert} \}$   
 $\backslash \mathrm{def} \backslash \mathrm{dist} \{ \backslash \mathrm{rho} \} \backslash \mathrm{def} \backslash \mathrm{and} \{ \& \} \backslash \mathrm{def} \backslash \mathrm{brackets} \# 1 \{ \backslash \mathrm{left} \{ \# 1 \backslash \mathrm{right} \} \} \backslash \mathrm{def} \backslash \mathrm{parc} \# 1 \# 2 \{ \backslash \mathrm{frac} \{ \backslash \mathrm{partial} \# 1 \} \{ \backslash \mathrm{partial} \# 2 \} \} \backslash \mathrm{def} \backslash \mathrm{mtr} \# 1 \{ \backslash \mathrm{begin} \{ \mathrm{pmatrix} \} \# 1 \backslash \mathrm{end} \{ \mathrm{pmatrix} \} \}$   
 $\backslash \mathrm{def} \backslash \mathrm{bm} \# 1 \{ \backslash \mathrm{boldsymbol} \{ \# 1 \} \} \backslash \mathrm{def} \backslash \mathrm{mcal} \# 1 \{ \backslash \mathrm{mathcal} \{ \# 1 \} \}$   
 $\backslash \mathrm{def} \backslash \mathrm{vv} \# 1 \{ \backslash \mathrm{mathbf} \{ \# 1 \} \} \backslash \mathrm{def} \backslash \mathrm{vvp} \# 1 \{ \backslash \mathrm{pmb} \{ \# 1 \} \} \backslash \mathrm{def} \backslash \mathrm{ve} \{ \backslash \mathrm{varepsilon} \} \backslash \mathrm{def} \backslash \{ \backslash \mathrm{lambda} \}$   
 $\backslash \mathrm{def} \backslash \mathrm{th} \{ \backslash \mathrm{vartheta} \} \backslash \mathrm{def} \backslash \mathrm{a} \{ \backslash \mathrm{alpha} \} \backslash \mathrm{def} \backslash \mathrm{tagged} \# 1 \{ \{ \backslash \mathrm{text} \{ \# 1 \} \} \} \backslash \mathrm{def} \backslash \mathrm{tagEq} \# 1 \{ \backslash \mathrm{href} \{ \backslash \mathrm{#eq} \# 1 \} \{ \backslash \mathrm{text} \{ \# 1 \} \} \} \backslash \mathrm{def} \backslash \mathrm{tagDe} \# 1 \{ \backslash \mathrm{href} \{ \backslash \mathrm{#de} \# 1 \} \{ \backslash \mathrm{text} \{ \# 1 \} \} \} \backslash \mathrm{def} \backslash \mathrm{T} \# 1 \{ \backslash \mathrm{htmlId} \{ \mathrm{eq} \# 1 \} \{ \# 1 \} \} \backslash \mathrm{def} \backslash \mathrm{D} \# 1 \{ \backslash \mathrm{htmlId} \{ \mathrm{de} \# 1 \} \{ \backslash \mathrm{vv} \{ \# 1 \} \} \} \backslash \mathrm{def} \backslash \mathrm{conv} \# 1 \{ \backslash \mathrm{mathrm} \{ \mathrm{conv} \} \backslash, \# 1 \}$   
 $\backslash \mathrm{def} \backslash \mathrm{cone} \# 1 \{ \backslash \mathrm{mathrm} \{ \mathrm{cone} \} \backslash, \# 1 \} \backslash \mathrm{def} \backslash \mathrm{aff} \# 1 \{ \backslash \mathrm{mathrm} \{ \mathrm{aff} \} \backslash, \# 1 \} \backslash \mathrm{def} \backslash \mathrm{lin} \# 1 \{ \backslash \mathrm{mathrm} \{ \mathrm{Lin} \} \backslash, \# 1 \} \backslash \mathrm{def} \backslash \mathrm{span} \# 1 \{ \backslash \mathrm{mathrm} \{ \mathrm{span} \} \backslash, \# 1 \} \backslash \mathrm{def} \backslash \mathrm{O} \{ \backslash \mathrm{mathcal} \mathrm{O} \} \backslash \mathrm{def} \backslash \mathrm{ri} \# 1 \{ \backslash \mathrm{mathrm} \{ \mathrm{ri} \} \backslash, \# 1 \}$   
 $\backslash \mathrm{def} \backslash \mathrm{rd} \# 1 \{ \backslash \mathrm{mathrm} \{ \mathrm{r} \} \backslash \mathrm{partial} \backslash, \# 1 \} \backslash \mathrm{def} \backslash \mathrm{interior} \# 1 \{ \backslash \mathrm{mathrm} \{ \mathrm{int} \} \backslash, \# 1 \} \backslash \mathrm{def} \backslash \mathrm{proj} \{ \backslash \mathrm{Pi} \}$   
 $\backslash \mathrm{def} \backslash \mathrm{epi} \# 1 \{ \backslash \mathrm{mathrm} \{ \mathrm{epi} \} \backslash, \# 1 \} \backslash \mathrm{def} \backslash \mathrm{grad} \# 1 \{ \backslash \mathrm{mathrm} \{ \mathrm{grad} \} \backslash, \# 1 \} \backslash \mathrm{def} \backslash \mathrm{hess} \# 1 \{ \backslash \mathrm{nabla} ^ 2 \backslash, \# 1 \} \$ \$$

## Definice $\S D\{2.2.1\}$ (Konvexní funkce)

Nechť  $X \subseteq \mathbb{R}^n$  je **konvexní** množina. Funkce  $f: X \rightarrow \mathbb{R}$  se nazývá

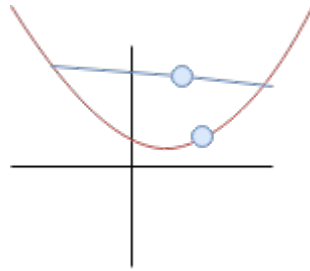
- **konvexní** na  $X$ , jestliže pro všechna  $x_1, x_2 \in X$  a každé  $\lambda \in [0, 1]$  platí
$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \tag{T\{2.2.1\}}$$
- **ostře konvexní** na  $X$ , jestliže nerovnost  $\tag{Eq\{2.2.1\}}$  je **ostrá** pro všechna  $x_1, x_2 \in X$ ,  $x_1 \neq x_2$  a každé  $\lambda \in (0, 1)$ .
- **silně konvexní** na  $X$  s **konstantou silné konvexnosti**  $\theta > 0$ , jestliže pro všechna  $x_1, x_2 \in X$  a každé  $\lambda \in [0, 1]$  platí
$$\underbrace{f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)}_{\tag{Eq\{2.2.1\}}} - \lambda(1 - \lambda)\mathrm{norm}\{x_1 - x_2\}^2 \tag{T\{2.2.2\}}$$

“ V praxi je **silná** konvexnost "silnější" než **ostrá** konvexnost a ta je silnější než "obyčejná" konvexnost

## Věta $\S D\{2.2.2\}$ (Konvexnost nadgrafu)

Nechť  $X \subseteq \mathbb{R}^n$  je konvexní množina a nechť  $f: X \rightarrow \mathbb{R}$ . Funkce  $f$  je **konvexní** na  $X$  právě tehdy, když její **nadgraf (epigraf)**  $\mathrm{epi} f := \{ [x, \beta] \mid x \in X, \beta \geq f(x) \}$  je **konvexní** množina.

Pro **ostře** konvexní funkci musí "tyto dva body" vždy ležet nad sebou (myšleny jejich souřadnice na ose  $y$ ). Navíc pro **silnou** konvexnost mezi nimi musí vždy být alespoň daná mezera.



“ Tyto body **nemusí** ležet nad sebou (na svislé přímce). Navíc ještě

$f$  **konvexní**  $\iff -f$  **konkávní**

## Kombinace konvexních funkcí

### Věta $\{2.2.3\}$ (Nezáporná lineární kombinace konvexních funkcí)

Nechť  $X \subseteq \mathbb{R}^n$  je konvexní množina, funkce  $f_1, \dots, f_m: X \rightarrow \mathbb{R}$  jsou konvexní na  $X$  a  $\alpha_1, \dots, \alpha_m \geq 0$  jsou daná čísla. Potom  $F(x) = \alpha_1 f_1(x) + \dots + \alpha_m f_m(x)$  je **konvexní**.

### Věta $\{2.2.4\}$ (Sublevel set)

Nechť  $X \subseteq \mathbb{R}^n$  je konvexní množina a  $f: X \rightarrow \mathbb{R}$  je *konvexní* funkce na  $X$ . Pak pro libovolné  $K \in \mathbb{R}$  je odpovídající **dolní vrstevnicová množina (sublevel set)**  $V_K := \{x \in X \mid f(x) \leq K\}$  také **konvexní**.

“ Platí pouze tato implikace:  $f$  **konvexní**  $\implies$  sublevel set **konvexní**  
Například  $x^3$  má konvexní sublevel set, ale sama konvexní **není**.

Přesněji říkáme, že pokud má funkce  $f$  **konvexní sublevel set**, pak  $f$  je **kvazikonvexní**.

### Věta $\{2.2.5\}$ (Jensen)

Nechť  $X \subseteq \mathbb{R}^n$  je konvexní množina a funkce  $f: X \rightarrow \mathbb{R}$  je konvexní na  $X$ . Pak pro libovolné  $m \in \mathbb{N}$ ,  $x_1, \dots, x_m \in X$  a čísla  $\lambda_1, \dots, \lambda_m \geq 0$  splňující  $\sum_{i=1}^m \lambda_i = 1$  platí  $f\left(\sum_{i=1}^m \lambda_i x_i\right) \leq \sum_{i=1}^m \lambda_i f(x_i)$ .  
 $\{2.2.3\}$  Je-li navíc funkce  $f$  **ostře konvexní** a  $\lambda_1, \dots, \lambda_m \in (0,1)$ , pak rovnost v  $\{2.2.3\}$  nastane *právě tehdy*, když  $x_1 = \dots = x_m$ .

První část **Věty**  $\tagDe{2.2.5}$  lze jistě podle **Definice**  $\tagDe{2.2.1}$  nahradit ekvivalencí

“ Z Jensenovy nerovnosti  $\tagEq{2.2.3}$  lze odvodit například **AG nerovnost**  
$$\frac{x_1 + \dots + x_m}{m} \leq \sqrt[m]{x_1 \cdot \dots \cdot x_m}$$

## Lokalizace minima konvexní funkce

### Věta $\tagD{2.2.6}$

Nechť  $X \subseteq \mathbb{R}^n$  je konvexní množina a funkce  $f: X \rightarrow \mathbb{R}$  konvexní. Potom následující tvrzení jsou pravdivá

- Libovolné **lokální minimum** funkce  $f$  na  $X$  je současně **globálním minimem**.
- Množina bodů množiny  $X$ , v nichž funkce  $f$  nabývá svého **minima** na  $X$ , je **konvexní**. Je-li funkce dokonce **ostře konvexní**, pak je tato množina **nejvýše jednoprvková**.
- Je-li funkce  $f$  **diferencovatelná** na **otevřené** množině  $U \subseteq X$  a  $x^* \in U$  je jejím **stacionárním bodem**, tj.  $\nabla f(x^*) = 0$ , pak  $x^*$  je bodem **globálního minima** funkce  $f$  na množině  $X$ .

Z Věty  $\tagDe{2.2.6}$  mimo jiné plyne, že je-li  $f: X \rightarrow \mathbb{R}$  (*ostře*) konvexní a **spojitá** funkce na konvexní a **kompaktní** množině  $X \subseteq \mathbb{R}^n$ , pak  $f$  má na  $X$  (*právě jedno*) **globální minimum**.

### Věta $\tagD{2.2.7}$ (Základní věta konvexního programování)

Máme-li konvexní funkci  $f: X \rightarrow \mathbb{R}$  na polytopu  $X := \text{conv}\{x_1, \dots, x_m\} \subseteq \mathbb{R}^n$ , pak je maximum funkce  $f$  na  $X$  dosaženo v některém z bodů  $x_1, \dots, x_m$ .

**Obecněji:** je-li  $X$  konvexní a **kompaktní** množina, pak maximum nastává v **extrémním bodě** (tj. v takovém bodě, který **není netriviální** konvexní kombinací dvou bodů z  $X$ )

“ Z Věty  $\tagDe{2.2.7}$  plyne **základní věta lineárního programování**:  
Je-li funkce  $f$  **afinní** (taková funkce je konvexní i konkávní zároveň), pak **globální maximum** nastává v některém z bodů  $x_1, \dots, x_m$ , tj. v některém z "**vrcholů**" polytopu.

# Vlastnosti konvexních funkcí

## Věta 2.4.1 (Spojitost konvexní funkce)

Nechť  $X \subseteq \mathbb{R}^n$  je konvexní množina a funkce  $f: X \rightarrow \mathbb{R}$  je konvexní na  $X$ . Pak  $f$  je **spojitá** pro každé  $x \in \text{int } X$ .

Dále ještě známe několik podmínek zaručujících konvexnost funkce  $f: \mathbb{R} \rightarrow \mathbb{R}$ :

- má-li  $f$  vlastní derivaci v **otevřeném** intervalu  $I$ , pak  $f$  je (ostře) konvexní na  $I$  právě tehdy, když  $f'$  je **neklesající (rostoucí)** na  $I$ .
- má-li  $f$  vlastní derivaci v **otevřeném** intervalu  $I$ , pak  $f$  je (ostře) konvexní na  $I$  právě tehdy, když pro každé  $x, x^* \in I$  platí  $f(x) \geq f(x^*) + f'(x^*)(x - x^*)$ , tj. graf funkce  $f$  leží **nad tečnou** sestrojenou v **libovolném** bodě.
- má-li  $f$  vlastní **druhou** derivaci v **otevřeném** intervalu  $I$ , pak  $f$  je konvexní na  $I$  právě tehdy, když funkce  $f''(x) \geq 0$  (je-li  $f''(x) > 0$  na  $I$ , pak **ostře konvexní**).

Tyto tvrzení si nyní rozšíříme pro  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  pro **silnou** konvexnost s konstantnou  $\theta$  (volbou  $\theta = 0$  dostáváme **ostrou** konvexnost)

## Věta 2.4.2

Nechť  $X \subseteq \mathbb{R}^n$  je konvexní množina a funkce  $f$  diferencovatelná na otevřené množině  $\text{int } X$ . Pak  $f$  je **silně** konvexní na  $X$  s konstantou **silné** konvexnosti  $\theta \geq 0$  právě tehdy, když pro každé  $x, x^* \in X$  platí  $f(x) \geq f(x^*) + \langle \text{grad } f(x^*), x - x^* \rangle + \frac{\theta}{2} \|x - x^*\|^2$  (tag 2.4.1)

Ve výrazu  $\langle \text{grad } f(x^*), x - x^* \rangle$  hraje úlohu **tečné nadroviny** v bodě  $x^*$  s **normálovým vektorem**  $\text{grad } f(x^*)$  (tečným jak na nadrovinu, tak na funkci  $f$  v bodě  $x^*$ )

Ještě jinými slovy je z Věty 2.4.2 plyne, že nadrovina daná  $\langle \text{grad } f(x) - \text{grad } f(x^*), x - x^* \rangle \geq \theta \|x - x^*\|^2$  **opěrnou nadrovinou** pro  $f$

## Důsledek 2.4.5

Nechť  $X \subseteq \mathbb{R}^n$  je konvexní množina splňující  $\text{int } X \neq \emptyset$ . Nechť funkce  $f: X \rightarrow \mathbb{R}$  je **dvakrát** spojitě diferencovatelná na **otevřené** množině  $\text{int } X$  s maticí druhých derivací  $\text{Hess } f(x)$  (*Hessova matice*). Pak  $f$  je **silně** konvexní na  $X$  s konstantou silné konvexnosti  $\theta \geq 0$  právě tehdy, když pro každé  $x \in X$  a  $h \in \mathbb{R}^n$

platí  $\|\text{hess } f(x)\|_h \geq 2 \|\text{h}\|^2 \tag{T{2.4.3}}$ , jinými slovy  $\text{hess } f(x) \geq 2I$  pro všechna  $x \in X$ .

---

Z Důsledku  $\tag{De{2.4.5}}$  plynou následující tři implikace

- $\text{hess } f(x) \geq 0$  pro všechna  $x \in X$   $\implies f$  je konvexní na  $X$
  - $\text{hess } f(x) > 0$  pro všechna  $x \in X$   $\implies f$  je **ostře** konvexní na  $X$
  - $\text{interior } X \neq \emptyset$  a  $f$  je konvexní na  $X$   $\implies \text{hess } f(x) \geq 0$  pro všechna  $x \in X$
- 

Revision #23

Created 30 December 2022 12:20:12 by Sceptri

Updated 7 June 2023 11:06:01 by Sceptri