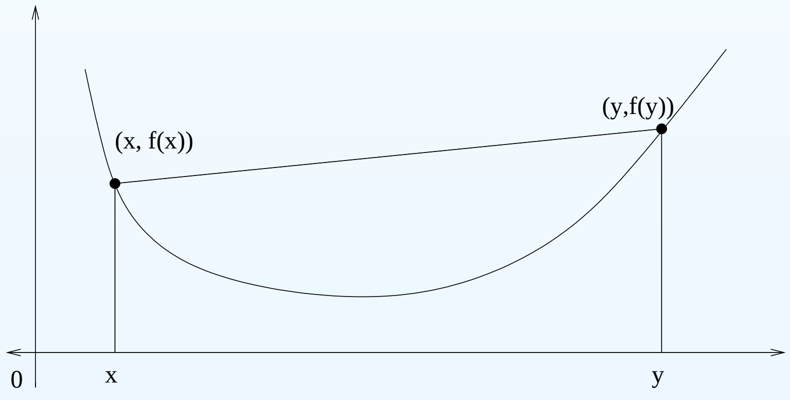
Konvexná optimalizácia

Konvexná analýza - funkcie

Funkcia $f : K \subseteq \mathbb{R}^n \to \mathbb{R}$ sa nazýva konvexná, ak množina K je konvexná a $\forall x, y \in K$, $\forall \lambda \in [0, 1]$ platí

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$



- $<, \geq, >$ - rýdzo konvexná, konkávna, rýdzo konkávna
- $f : K \subseteq \mathbb{R}^n \to \mathbb{R}$ je konvexná $\Leftrightarrow \forall x, y \in K$, $\forall \lambda \in [0, 1]$ je

$$g(\lambda) = f(\lambda x + (1 - \lambda)y)$$

konvexná.

Podúrovňové množiny

Ak $f : \mathbb{R}^n \to \mathbb{R}$ je konvexná

- tak $\forall \alpha \in \mathbb{R}$ je $S_\alpha = \{x \mid f(x) \leq \alpha\}$ **konvexná a uzavretá**.
- a existuje α_0 tak, že $S_{\alpha_0} = \{x \mid f(x) \leq \alpha_0\} \neq \emptyset$ a je ohraničená, tak S_α je ohraničená pre všetky $\alpha > \alpha_0$.

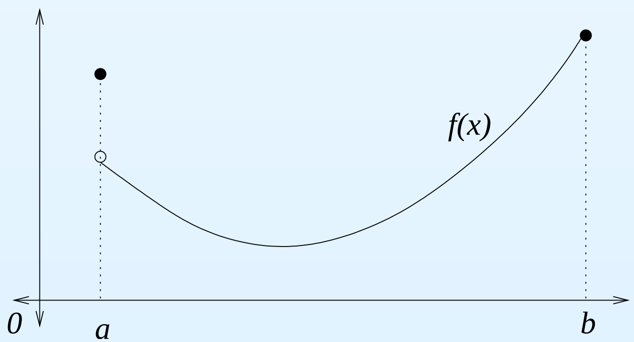
Minimá konvexnej funkcie

- Každé lokálne minimum konvexnej funkcie je jej globálnym minimom.
- Množina všetkých miním konvexnej funkcie tvorí konvexnú množinu.
- Rýdzo konvexná funkcia má nanajvýš jedno minimum.

Spojitosť

- Nech $K \subseteq \mathbb{R}^n$ je otvorená konvexná množina a $f : K \to \mathbb{R}$ je konvexná. Potom f je Lipschitzovská na ľubovolnej kompaktnej podmnožine $U \subseteq K$, t. j. $\forall x, y \in U$ existuje konštanta L tak, že
$$|f(x) - f(y)| \leq L\|x - y\|.$$
- Konvexná funkcia $f : K \subset \mathbb{R}^n \to \mathbb{R}$ definovaná na **otvorenej** konvexnej množine K je na nej spojitá.
- Konvexná funkcia $f : K \subset \mathbb{R}^n \to \mathbb{R}$ definovaná na **uzavretej** konvexnej množine K je na nej zhora polospojitá.

Konvexná funkcia definovaná na uzavretej množine na nej nemusí byť spojitá.



Podmienky I. rádu

Funkcia $f : K \subset \mathbb{R}^n \to \mathbb{R}$ definovaná na konvexnej množine K je konvexná **práve vtedy, ked'**

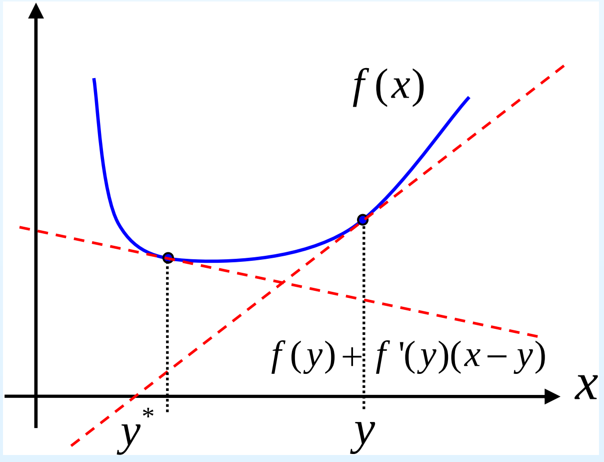
- (Taylorová aproximácia f leží pod grafom funkcie)

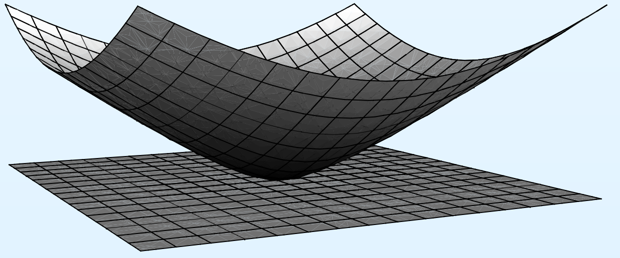
$$\forall x, y,\ x \neq y : f(x) \geq f(y) + \nabla f(y)^T (x - y)$$

- (Monotónnost' gradientu)

$$\forall x, y,\ x \neq y : (\nabla f(x) - \nabla f(y))^T (x - y) \geq 0$$

- $>, \leq, <$ - rýdzo konvexná, konkávna, rýdzo konkávna





Podmienky II. rádu

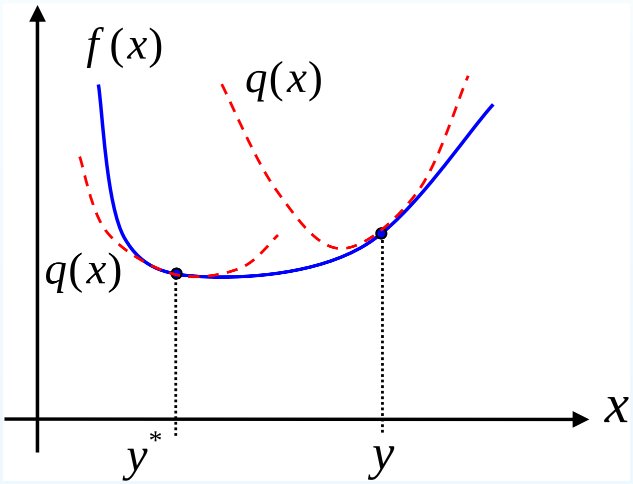
Funkcia $f : K \subset \mathbb{R}^n \rightarrow \mathbb{R}$ definovaná na konvexnej množine K je konvexná **práve vtedy, ked'**

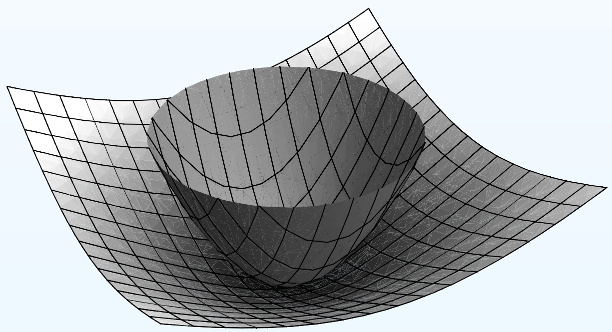
- (Kladná semidefinitnosť Hessovej matice)

$$\forall x : \nabla^2 f(x) \succeq 0$$

- $\preceq$ - konkávna,
- $\succ, \prec$ - rýdzo konvexná, rýdzo konkávna - platí **len implikácia - nie ekvivalencia** - $f(x) = x^4$

Podmienky II. rádu





Konvexná funkcia $f(x)$ a jej Taylorova aproximácia II. rádu v bode y - kvadratická konvexná funkcia

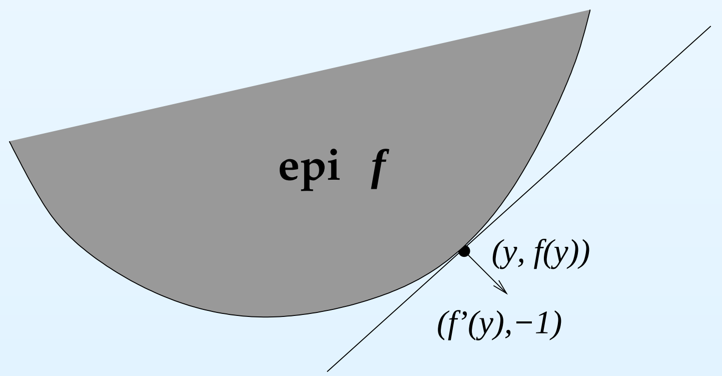
$$q(x) = f(y) + \nabla f(y)^T (x - y) + \frac{1}{2}(x - y)^T \nabla^2 f(y)(x - y)$$

Epigraf funkcie $f : K \subset \mathbb{R}^n \to \mathbb{R}$ je

$$\mathbf{epi} f = \{(x,t) \mid x \in K, f(x) \le t\}$$

- Funkcia je konvexná $\Leftrightarrow$ jej epigraf je konvexná množina.
- **Interpretácia podmienok I. rádu**: nadrovina s normálovým vektorom $(\nabla f(y), -1)$ je opornou nadrovinou $\mathbf{epi} f$ v hraničnom bode $(y, f(y))$

$$(x,t) \in \mathbf{epi} f \;\Rightarrow\; \begin{pmatrix} \nabla f(y) \\ -1 \end{pmatrix}^T \begin{pmatrix} x \\ t \end{pmatrix} \le \begin{pmatrix} \nabla f(y) \\ -1 \end{pmatrix}^T \begin{pmatrix} y \\ f(y) \end{pmatrix}$$



Operácie zachovávajúce konvexnosť

- **nezáporná lineárna kombinácia:** ak $f_1, \ldots, f_m$ sú konvexné, $w_1, \ldots, w_m \geq 0$ - tak
$$g(x) = w_1 f_1(x) + \cdots + w_m f_m(x)$$
je konvexná
- **afinna transformácia premenných:** ak f je konvexná, tak
$$g(x) = f(Ax + b)$$
je konvexná
- **maximum:** ak $f_1, \ldots, f_m$ sú konvexné, tak
$$g(x) = \max\{f_1(x), \ldots, f_m(x)\}$$
je konvexná

Operácie zachovávajúce konvexnosť

- **suprémum:**
  ak $\forall y \in \mathcal{C}$ je $f(x, y)$ konvexná v x a $\sup_{y\in\mathcal{C}} f(x, y) < \infty$, tak
  $$g(x) = \sup_{y\in\mathcal{C}} f(x, y)$$
  je konvexná

- **Príklad:** vzdialenosť bodu x do "najvzdialenejšieho" bodu množiny $\mathcal{C}$:
  $$g(x) = \sup_{y\in\mathcal{C}} \|x - y\|$$

Operácie zachovávajúce konvexnost'

- **infímum:**
  ak $f(x,y)$ je konvexná v (x,y), množina $\mathcal{C}$ je **konvexná** a $\inf_{y\in\mathcal{C}} f(x,y) > -\infty$, tak
  $$g(x) = \inf_{y\in\mathcal{C}} f(x,y)$$
  je konvexná

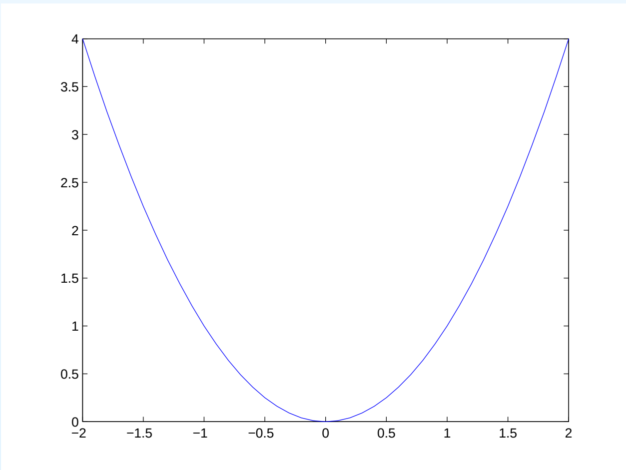
- **Príklad:** vzdialenost' bodu x od konvexnej množiny C:
  $$\mathrm{dist}(\mathrm{x},\mathrm{C}) = \inf_{\mathrm{y}\in\mathrm{C}} \|\mathrm{x}-\mathrm{y}\|$$

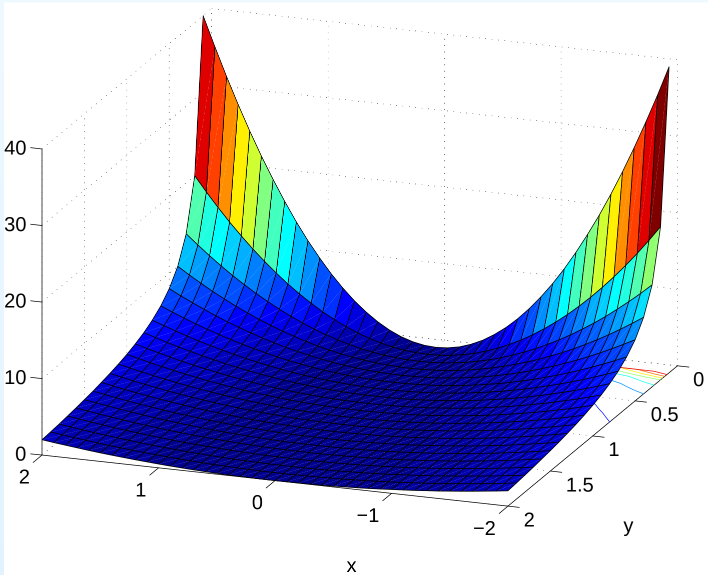
Operácie zachovávajúce konvexnosť

- Ak $f : \mathbb{R}^n \to \mathbb{R}$ je konvexná, tak **perspektíva**

$$g : \mathbb{R}^n \times \mathbb{R}_{++} \to \mathbb{R}, \; g(x,t) = tf\left(\frac{x}{t}\right)$$

  je konvexná.

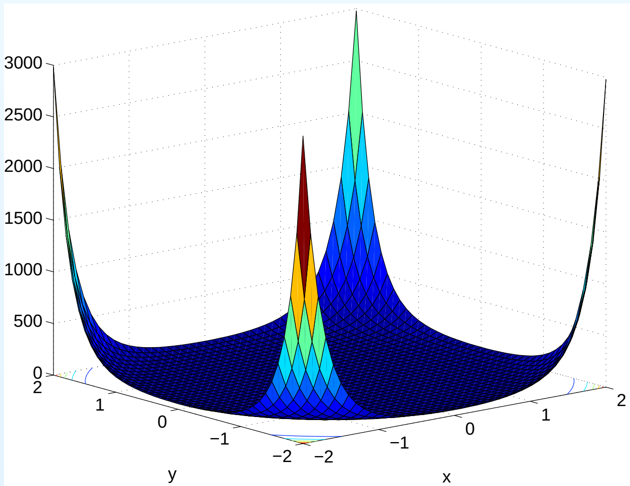




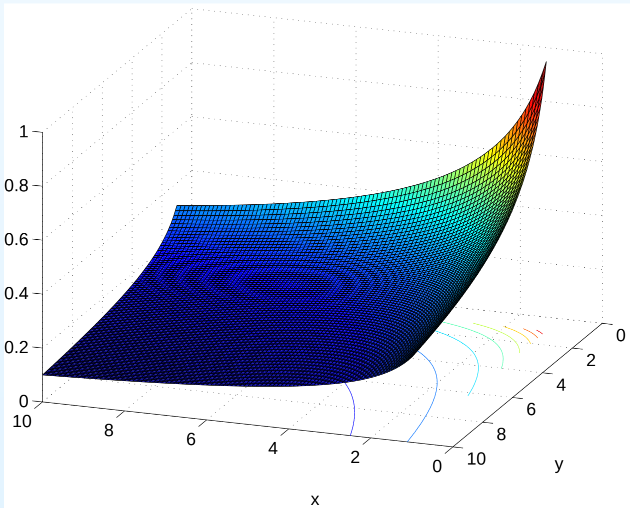
Kompozícia funkcií

$$h(t):\mathbb{R}\to\mathbb{R},\ f(x):\mathbb{R}^n\to\mathbb{R},\ H(x)=h(f(x)):\mathbb{R}^n\to\mathbb{R}$$

f	h	h		H
$\smile$	$\smile$	$\nearrow$	$\Rightarrow$	$\smile$
$\frown$	$\smile$	$\searrow$	$\Rightarrow$	$\smile$
$\frown$	$\frown$	$\nearrow$	$\Rightarrow$	$\frown$
$\smile$	$\frown$	$\searrow$	$\Rightarrow$	$\frown$



$$f(x_1,x_2)=\exp(x_1^2+x_2^2)$$



$$f(x_1,x_2)=1/\sqrt{x_1x_2}$$

! opačná implikácia v tabul'ke neplatí

$H(x_1, x_2) = \sqrt{x_1 x_2}$ je konkávna na $\mathbb{R}^2_{++}$, $h(t) = \sqrt{t}$ je konkávna, rastúca na $\mathbb{R}_{++}$, avšak $f(x_1, x_2) = x_1 x_2$ nie je konkávna (ani konvexná)

Vektorová kompozícia

$$h(y) : \mathbb{R}^m \to \mathbb{R},\ f_i(x) : \mathbb{R}^n \to \mathbb{R},\ i = 1, \ldots, m,$$

$$H : \mathbb{R}^n \to \mathbb{R},\ H(x) = h(f_1(x), \ldots, f_m(x))$$

$f_i, \forall i$	h	h		H
$\smile$	$\smile$	$\nearrow\nearrow\nearrow$	$\Rightarrow$	$\smile$
$\frown$	$\smile$	$\searrow\searrow\searrow$	$\Rightarrow$	$\smile$
$\frown$	$\frown$	$\nearrow\nearrow\nearrow$	$\Rightarrow$	$\frown$
$\smile$	$\frown$	$\searrow\searrow\searrow$	$\Rightarrow$	$\frown$

Kvázikonvexné funkcie

Funkcia $f : K \subseteq \mathbb{R}^n \to \mathbb{R}$ sa nazýva kvázikonvexná, ak množina K je konvexná a $\forall x, y \in K$, $\forall \lambda \in [0, 1]$ platí

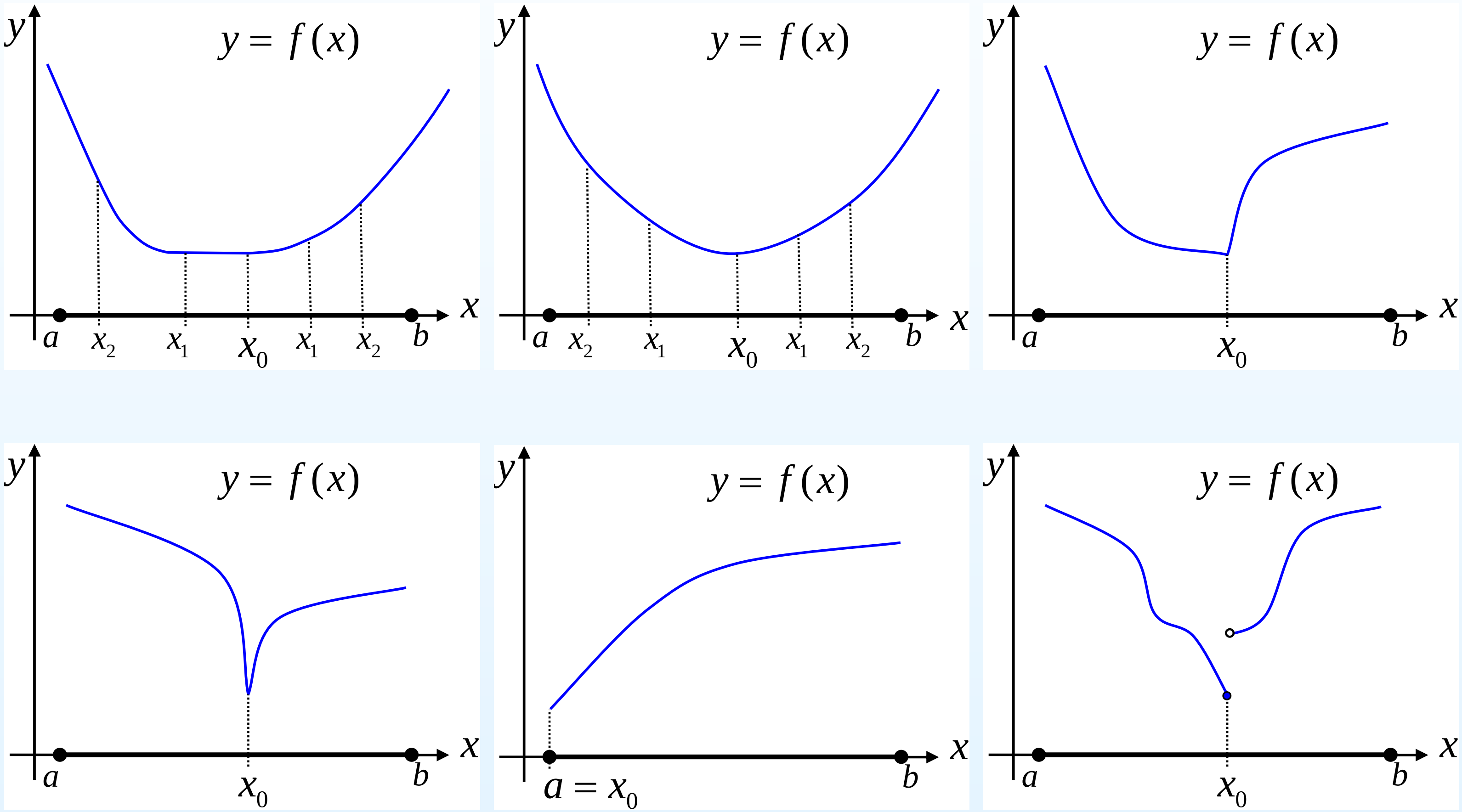
$$f(\lambda x + (1-\lambda)y) \leq \max\{f(x), f(y)\}.$$

- $<, \geq, >$ - rýdzo kvázikonvexná, kvázikonkávna, rýdzo kvázikonkávna
- kvázilineárne funkcie
- **Ekvivalentná definícia:** Funkcia $f : K \subseteq \mathbb{R}^n \to \mathbb{R}$ sa nazýva kvázikonvexná, ak množina K je konvexná a $\forall \alpha \in \mathbb{R}^n$ je $S_\alpha = \{x \mid f(x) \leq \alpha\}$ konvexná
- $f : K \subseteq \mathbb{R}^n \to \mathbb{R}$ je kvázikonvexná $\Leftrightarrow$ $\forall x, y \in K$, $\forall \lambda \in [0, 1]$ je

  $$g(\lambda) = f(\lambda x + (1-\lambda)y)$$

  kvázikonvexná.

Kvázikonvexné funkcie



Operácie zachovávajúce kvázikonvexnosť

- **vážené maximum:** ak $f_1, \ldots, f_m$ sú konvexné, $w_1, \ldots, w_m \geq 0$ - tak
$$g(x) = \max\{w_1 f_1(x), \ldots, w_m f_m(x)\}$$
je konvexná
- **suprémum:**
ak $\forall y \in \mathcal{C}$ je $f(x, y)$ kvázikonvexná v x a $\sup_{y \in \mathcal{C}} f(x, y) < \infty$, tak
$$g(x) = \sup_{y \in \mathcal{C}} f(x, y)$$
je kvázikonvexná
- **infímum:**
ak $f(x, y)$ je kvázikonvexná v (x, y), množina $\mathcal{C}$ je **konvexná** a $\inf_{y \in \mathcal{C}} f(x, y) > -\infty$, tak
$$g(x) = \inf_{y \in \mathcal{C}} f(x, y)$$
je kvázikonvexná

Podmienky I. rádu

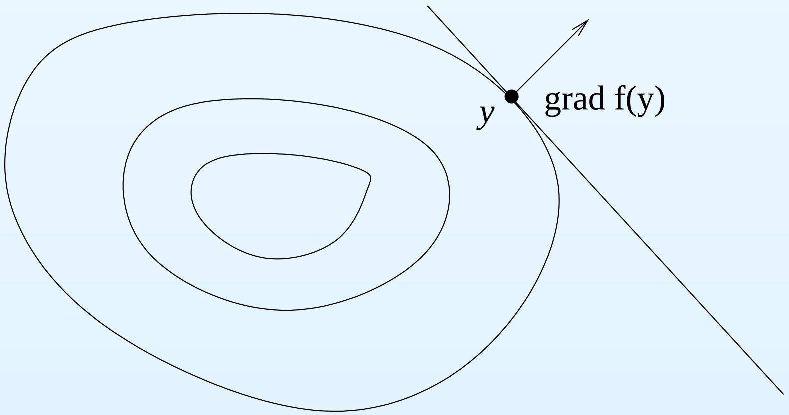
- Funkcia $f : K \subset \mathbb{R}^n \to \mathbb{R}$ definovaná na konvexnej množine K je kvázikonvexná **práve vtedy, ked'**

$$\forall x, y,\ x \neq y : f(x) \leq f(y) \ \Rightarrow\ \nabla f(y)^T (x - y) \leq 0$$

- Geometrická interpretácia: ak $\nabla f(y) \neq 0$, tak $\nabla f(y)$ je normálový vektor opornej nadroviny podúrovňovej množiny

$$S = \{x \mid f(x) \leq f(y)\}$$

v bode y.



Podmienky II. rádu

- Ak f je kvázikonvexná, tak $\forall x, y$ platí

$$y^T \nabla f(x) = 0 \;\Rightarrow\; y^T \nabla^2 f(x) y \geq 0$$

- Ak f spĺňa $\forall x, y \neq 0$

$$y^T \nabla f(x) = 0 \;\Rightarrow\; y^T \nabla^2 f(x) y > 0,$$

  tak f je kvázikonvexná.

- $\forall x : \nabla f(x) = 0$ je $\nabla^2 f(x)$ kladne (semi)definitná
- Ak $\nabla f(x) \neq 0$, tak $\nabla^2 f(x)$ je kladne (semi)definitná na podpriestore $\nabla f(x)^\perp$, resp. matica

$$H(x) = \begin{pmatrix} \nabla^2 f(x) & \nabla f(x) \\ \nabla f(x)^T & 0 \end{pmatrix}$$

  má práve jedno záporné vlastné číslo.

Silnokonvexné funkcie

Funkcia $f : \mathbb{R}^n \to \mathbb{R}$ sa nazýva silnokonvexná, ak existuje $\beta > 0$ tak, že $\forall x \neq y$ a $\forall \lambda \in (0, 1)$ platí

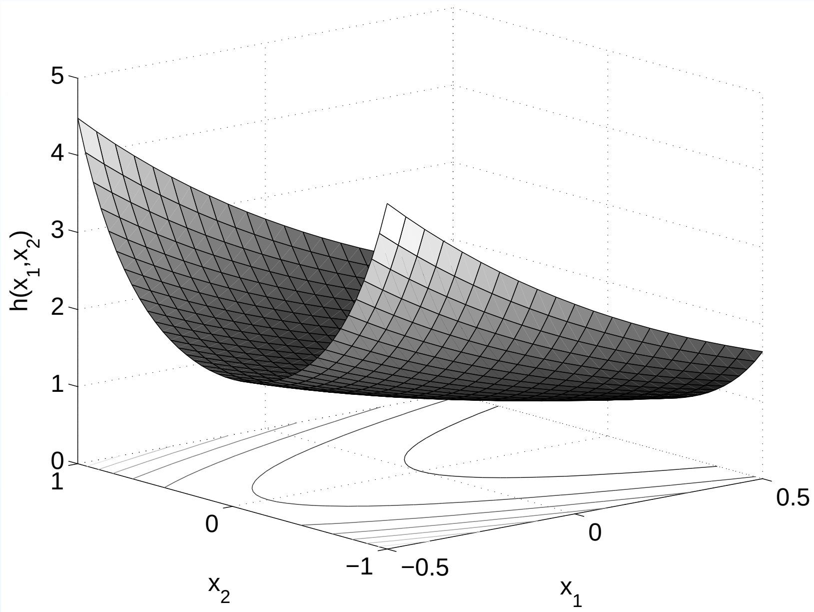
$$f(\lambda x + (1-\lambda)y) + \beta\lambda(1-\lambda)\|x-y\|^2 \leq \lambda f(x) + (1-\lambda)f(y).$$

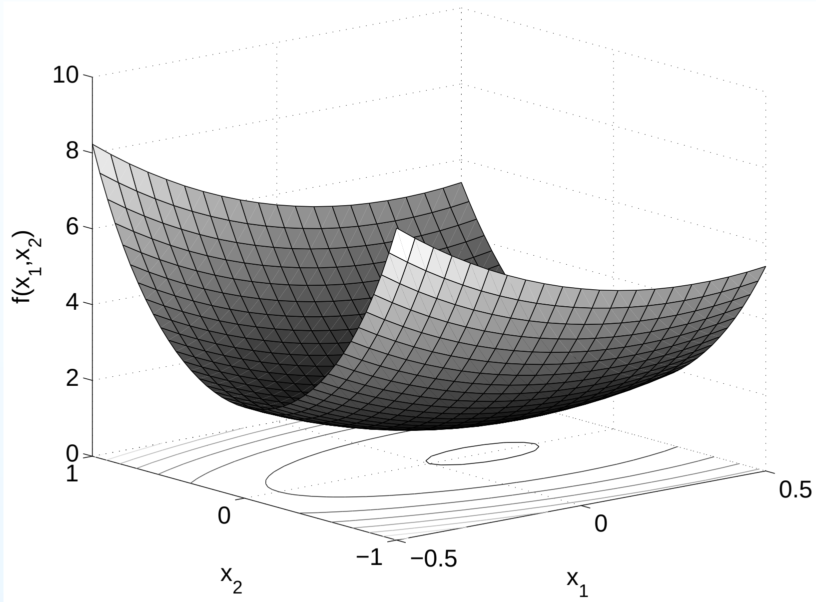
- $q(x) = \beta x^T x = \beta\|x\|_2^2,\ \beta > 0$ je "najslabšia" silnokonvexná funkcia
- $f(x)$ je silnokonvexná $\Leftrightarrow$ existuje $\beta > 0$ tak, že funkcia $h(x) = f(x) - q(x)$ je konvexná.
- Ak $f(x)$ silnokonvexná, tak $\forall \alpha$ je podúrovňová množina

  $$S_\alpha = \{x \mid f(x) \leq \alpha\}$$

  konvexná a kompaktná.

Prednáška tretia: Konvexná analýza - funkcie





Konvexná funkcia

$$h(x_1, x_2) = e^{-x_1 + x_2^2}$$

a silnokonvexná funkcia

$$f(x) = h(x) + \beta x^T x$$

Podmienky I. a II. rádu pre silnokonvexné funkcie

Funkcia $f(x)$ je silnokonvexná $\Leftrightarrow$ ak existuje $\beta > 0$ tak, že

- (kvadratická dolná hranica)

$$f(x) \geq f(y) + \nabla f(y)^T(x-y) + \beta\|x-y\|^2$$

- (silná monotónnosť - koercívnosť gradientu)

$$(\nabla f(x) - \nabla f(y))^T(x-y) \geq 2\beta\|x-y\|^2$$

- (spektrum Hessovej matice je vpravo od β)

$$\nabla^2 f(x) \succeq \beta I$$

Zovšeobecnená konvexnosť

Kužeľ $\mathcal{K}$ sa nazýva **vlastný** ak spĺňa nasledovné vlastnosti:

- $\mathcal{K}$ je konvexný;
- $\mathcal{K}$ je uzavretý;
- $\mathcal{K}$ je "plný" - t. j. $\mathrm{int}(\mathcal{K}) \neq \emptyset$;
- $\mathcal{K}$ je špicatý, t. j. $x \in \mathcal{K} \wedge -x \in \mathcal{K} \Rightarrow x = 0$

Príklad: $\mathbb{R}^n_+$, $\mathcal{S}^n_+$, $\mathcal{C}_2$

Čiastočné usporiadanie asociované s kužeľom $\mathcal{K}$:

$$x \preceq_{\mathcal{K}} y \;\Leftrightarrow\; y - x \in \mathcal{K}, \quad x \prec_{\mathcal{K}} y \;\Leftrightarrow\; y - x \in \mathrm{int}(\mathcal{K})$$

Vlastnosti zovšeobecnených nerovností

vlastnosť	$\preceq_K$	$\prec_K$
invariantnosť	$x \preceq_K y,\ u \preceq_K v \Rightarrow$ $x + u \preceq_K y + v$ $x \preceq_K y, \alpha \geq 0 \Rightarrow \alpha x \preceq_K \alpha y$	$x \prec_K y,\ u \preceq_K v \Rightarrow$ $x + u \prec_K y + v$ $x \prec_K y, \alpha > 0 \Rightarrow \alpha x \prec_K \alpha y$
reflexívnosť	$x \preceq_K x$	! $x \not\prec_K x$
tranzitívnosť	$x \preceq_K y,\ y \preceq_K z \Rightarrow$ $x \preceq_K y$	$x \prec_K y,\ y \prec_K z \Rightarrow$ $x \prec_K z$
antisymetria	$x \preceq_K y,\ y \preceq_K x \Rightarrow x = y$	—
	—	$x \prec_K y,\ \exists u, v$ dosť malé: $x + u \prec_K y + v$
	$x_i \preceq_K y_i, \forall i,\ x_i \to x, y_i \to y$ $\Rightarrow\ x \preceq_K y$	—

- $\mathcal{K} = \mathbb{R}^n_+$
  zovšeovecnená nerovnosť' je ekvivalentná porovnávaniu vektorov po zložkách: $x \preceq_{\mathcal{K}} y \;\Leftrightarrow\; x_i \leq y_i \; \forall i = 1, 2, \ldots, n$

- $\mathcal{K} = \mathcal{S}^n_+$
  zovšeovecnená nerovnosť' zodpovedá tzv. **Löwnerovmu usporiadaniu symetrických matíc** $\preceq$:

  - $A = Q\Lambda Q^T \preceq \alpha I$ - spektrum matice A je zhora ohraničené konštantou α
  - pre kladne semidefinitné matice nerovnosť' $0 \preceq A \preceq B$ implikuje napríklad $h(A) \leq h(B)$, resp. $\det(A) \leq \det(B)$

Zovšeobecnenie pojmu konvexnosti:

Nech $\mathcal{K} \subseteq \mathbb{R}^m$ je vlastný kužel' a $\preceq_{\mathcal{K}}$ je príslušná zovšeobecnená nerovnost'. Potom funkcia $f : \mathbb{R}^n \to \mathbb{R}^m$ sa nazýva $\mathcal{K}$-**konvexná** ak $\forall x, y \in \mathbb{R}^n$ a $\forall \lambda \in [0, 1]$ platí

$$f(\lambda x + (1-\lambda)y) \preceq_{\mathcal{K}} \lambda f(x) + (1-\lambda)f(y).$$

Príklad: Maticová konvexnost' Funkcia $f : \mathbb{R}^n \to \mathcal{S}^m$ sa nazýva maticovo konvexná ak

$$f(\lambda x + (1-\lambda)y) \preceq \lambda f(x) + (1-\lambda)f(y).$$

pre $\forall x, y \in \mathbb{R}^n$ a $\forall \lambda \in [0, 1]$.
Ekvivalentná definícia: funkcia $z^T f(x) z$ je konvexná $\forall z \in \mathbb{R}^m$. Napr. funkcia $f : \mathbb{R}^{n\times m} \to \mathcal{S}^n$, $f(X) = XX^T$ je maticovo konvexná, lebo pre pevné z je funkcia $z^T XX^T z = \|X^T z\|^2$ konvexná kvadratická funkcia.