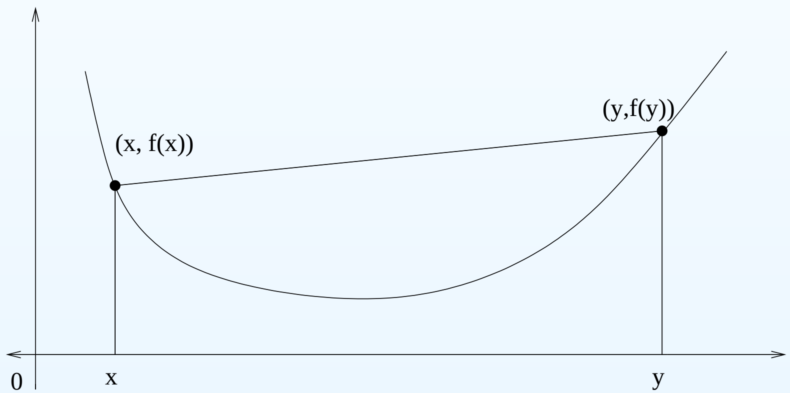
Convex Optimization

Convex Analysis - Functions

A function $f : K \subseteq \mathbb{R}^n \to \mathbb{R}$ is **convex**, if K is a convex set and $\forall x, y \in K, x \neq y$, $\forall \lambda \in (0,1)$ we have

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



- $<, \geq, >$ - strictly convex, concave, strictly concave
- $f : K \subseteq \mathbb{R}^n \to \mathbb{R}$ is convex iff $\forall x, y \in K, x \neq y$, $\forall \lambda \in (0,1)$
$$g(\lambda) = f(\lambda x + (1-\lambda)y)$$
is convex.

Sub-level sets

If $f : \mathbb{R}^n \to \mathbb{R}$ is convex

- then $\forall \alpha \in \mathbb{R}$ is $S_\alpha = \{x \mid f(x) \leq \alpha\}$ **convex and closed**.
- and if there exists α_0 such that $S_{\alpha_0} = \{x \mid f(x) \leq \alpha_0\}$ is nonempty and bounded, then S_α is bounded for any $\alpha > \alpha_0$.

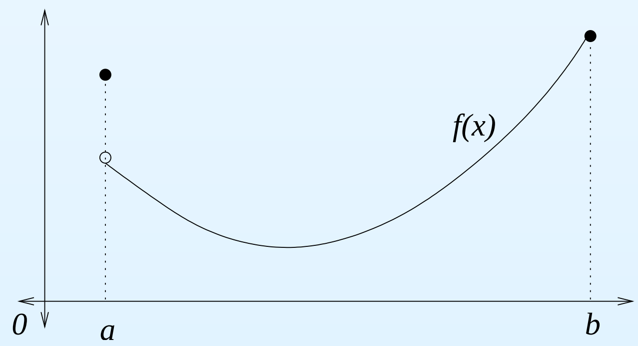
Minima of a convex function

- Every local minimum of a convex function is a global minimum.
- The set of all minima of a convex function is a convex set.
- Strictly convex function has at most one minimum.

Continuity

- Assume $K \subseteq \mathbb{R}^n$ is an open and convex set and $f : K \to \mathbb{R}$ is convex. Then f is Lipschitz continuous on any compact subset of $U \subseteq K$, that is there exists a constant L such that $\forall x, y \in U$
$$|f(x) - f(y)| \leq L\|x - y\|.$$
- Convex function $f : K \subset \mathbb{R}^n \to \mathbb{R}$ defined on an **open** convex set K is continuous on K.
- Convex function $f : K \subset \mathbb{R}^n \to \mathbb{R}$ defined on a **closed** convex set K is upper semi-continuous on K.

Convex function defined on a **closed** convex set K is not necessarily continuous.



First order conditions

The function $f : K \subset \mathbb{R}^n \to \mathbb{R}$ defined on an convex set K is convex **iff**

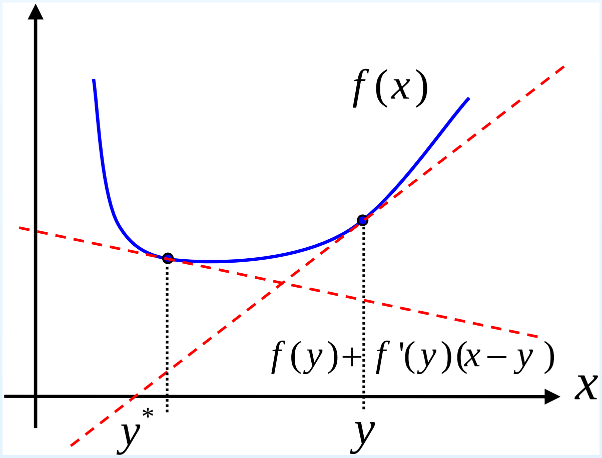
- (First order Taylor approximation lays below the graph of f)

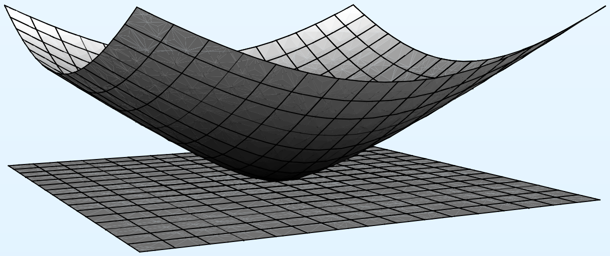
$$\forall x, y,\ x \neq y : f(x) \geq f(y) + \nabla f(y)^T (x - y)$$

- (Monotone gradient)

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

- $>, \leq, <$ - strictly convex/ concave/ strictly concave





Second order conditions

Th function $f : K \subset \mathbb{R}^n \to \mathbb{R}$ defined on an convex set K is convex **iff**

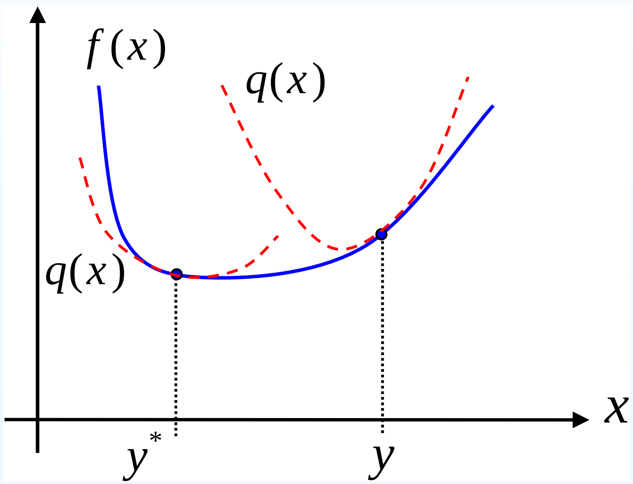
- (Positive semidefinite Hessian matrix)

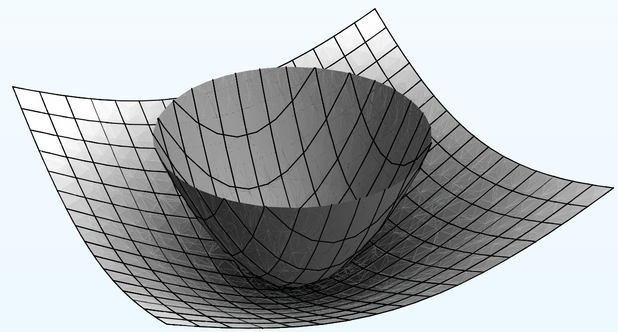
$$\forall x : \nabla^2 f(x) \succeq 0$$

- $\preceq$ - concave,

- $\succ, \prec$ - strictly convex/ strictly concave - **only one implication holds** - $f(x) = x^4$

Second order conditions





Convex function $f(x)$ and its second order Taylor approximation at the point y - convex quadratic function

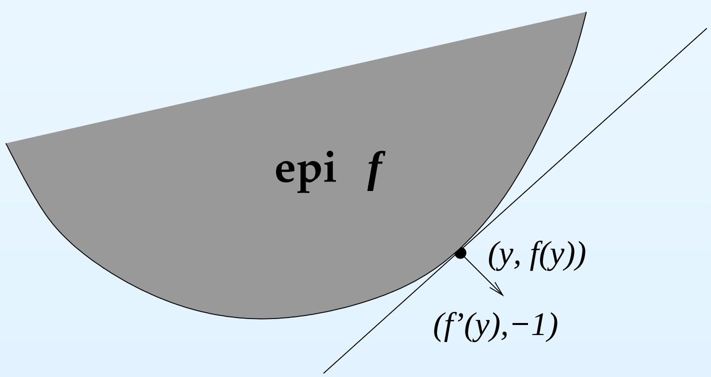
$$q(x) = f(y) + \nabla f(y)^T(x-y) + \frac{1}{2}(x-y)^T\nabla^2 f(y)(x-y)$$

Epigraph of a function $f : K \subset \mathbb{R}^n \to \mathbb{R}$ is the set

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

- The function is convex $\Leftrightarrow$ its epigraph is a convex set.
- **First order condition interpretation**: the hyperplane defined by the normal vector $(\nabla f(y), -1)$ is the supporting hyperplane $\mathbf{epi} f$ at the boundary point $(y, f(y))$

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



Operations preserving convexity

- **nonnegative linear combination:** if $f_1, \dots, f_m$ are convex $w_1, \dots, w_m \geq 0$ - then
$$g(x) = w_1 f_1(x) + \cdots + w_m f_m(x)$$
is convex.
- **affine transformation of variables:** if f is convex, then
$$g(x) = f(Ax + b)$$
is convex
- **point-wise maximum:** if $f_1, \dots, f_m$ are convex, then
$$g(x) = \max\{f_1(x), \dots, f_m(x)\}$$
is convex

Operations preserving convexity

- **supremum:**
  if $\forall y \in \mathcal{C}$ is $f(x,y)$ convex in x and $\sup_{y\in\mathcal{C}} f(x,y) < \infty$, then

$$g(x) = \sup_{y\in\mathcal{C}} f(x,y)$$

  is convex.

- **Example:** distance (of the point x) to the farthest point of the set $\mathcal{C}$:

$$g(x) = \sup_{y\in\mathcal{C}} \|x - y\|$$

Operations preserving convexity

- **infimum:**
  if $f(x, y)$ is convex in (x, y), the set $\mathcal{C} \neq \emptyset$ is **convex** and $\inf_{y \in \mathcal{C}} f(x, y) > -\infty$, then
  $$g(x) = \inf_{y \in \mathcal{C}} f(x, y)$$
  is convex

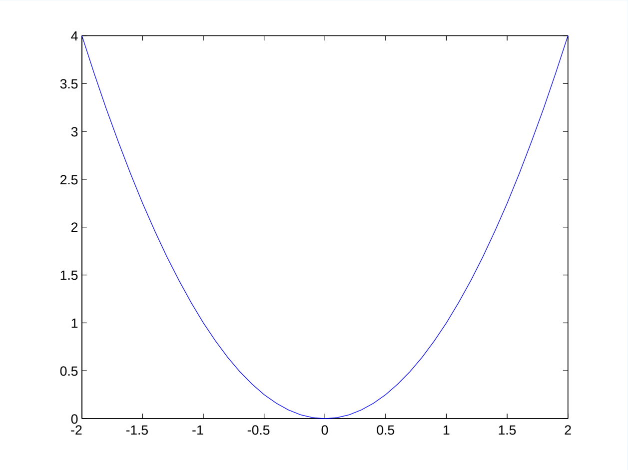
- **Example:** distance of the point x from the convex set C:
  $$\mathrm{dist}(\mathrm{x}, \mathrm{C}) = \inf_{\mathrm{y} \in \mathrm{C}} \|\mathrm{x} - \mathrm{y}\|$$

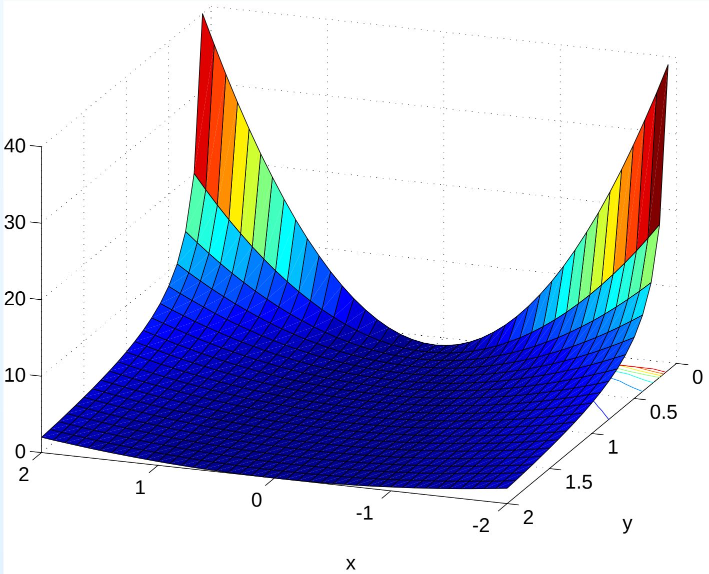
Operations preserving convexity

- If $f : \mathbb{R}^n \to \mathbb{R}$ is convex, then the **perspective function**

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

  is convex

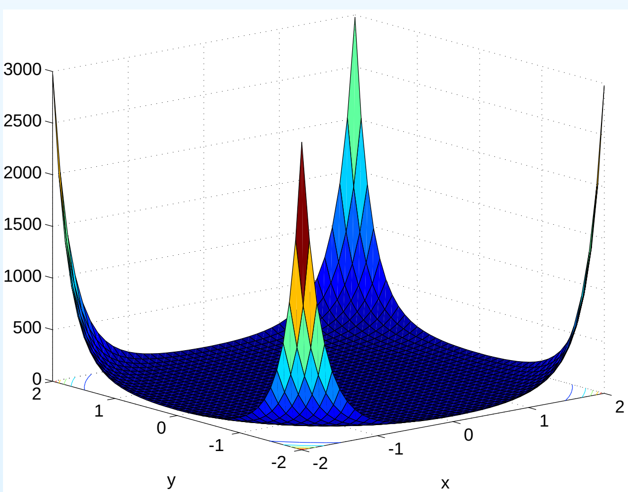




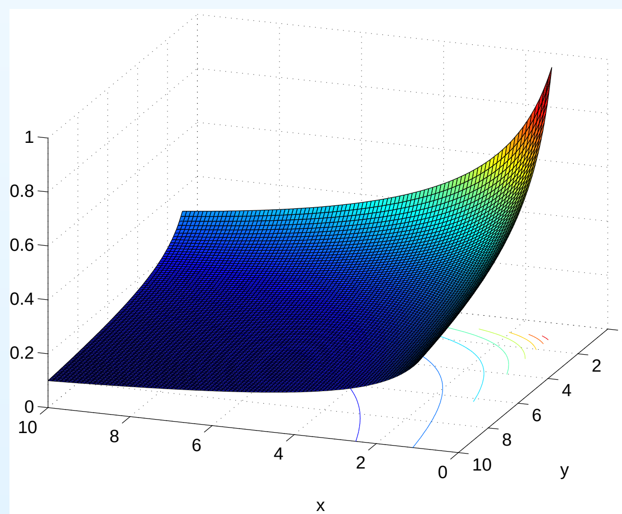
Composition

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

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



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



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

! the reverse implication does not hold!

$H(x_1, x_2) = \sqrt{x_1 x_2}$ is concave on $\mathbb{R}^2_{++}$, $h(t) = \sqrt{t}$ is concave and increasing on $\mathbb{R}_{++}$, but $f(x_1, x_2) = x_1 x_2$ is not concave (and not convex)

Vector composition

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

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

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

Quasi-convex functions

The function $f : K \subseteq \mathbb{R}^n \to \mathbb{R}$ is called quasi-convex, if the set K is convex and $\forall x, y \in K$, $\forall \lambda \in [0, 1]$ it holds

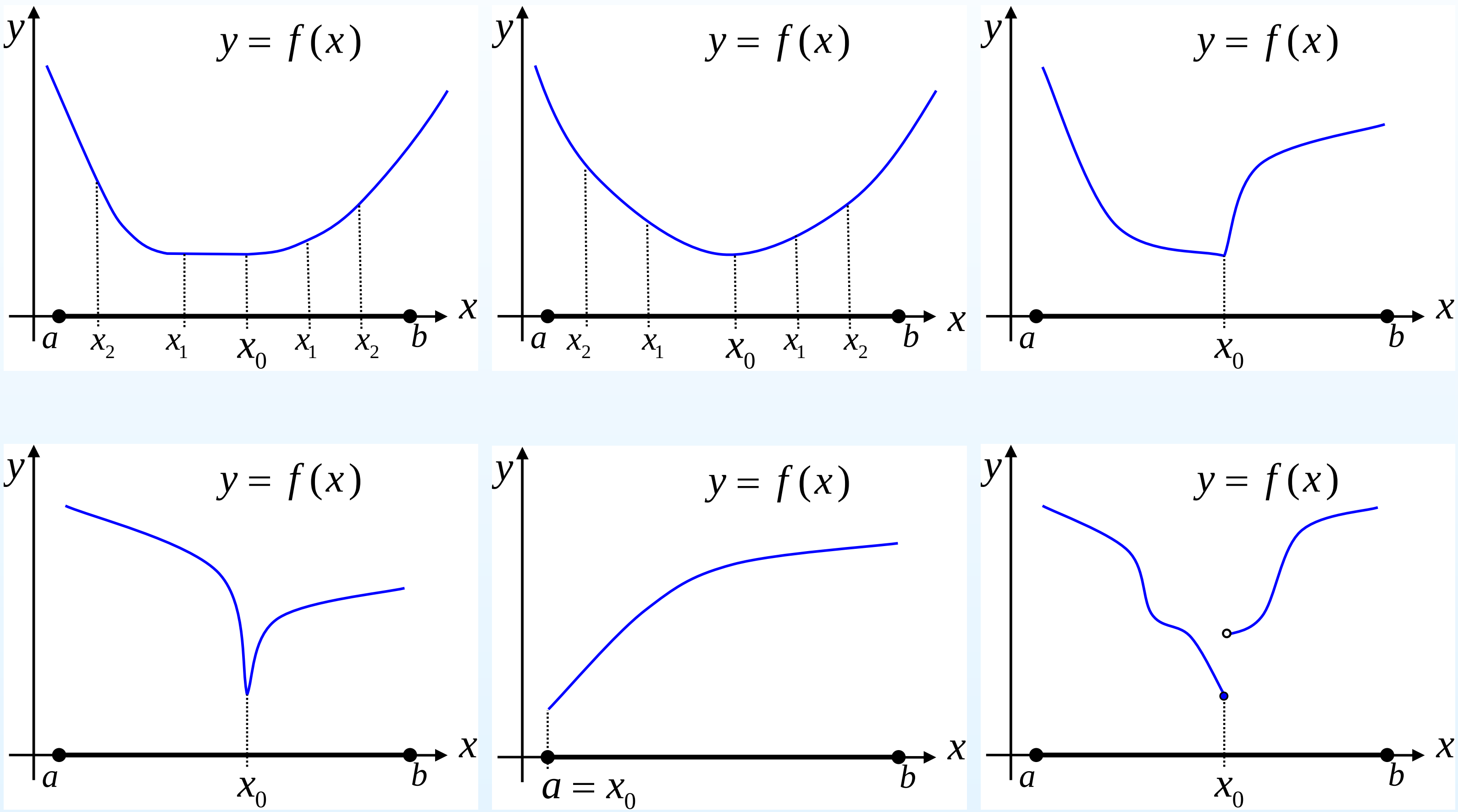
$$f(\lambda x + (1 - \lambda)y) \leq \max\{f(x), f(y)\}.$$

- $<, \geq, >$ - strictly quasi-convex, quasi-concave, strictly quasi-convex
- quasilinear functions
- **Equivalent definition:** The function $f : K \subseteq \mathbb{R}^n \to \mathbb{R}$ is called quasi-convex, if the set K is convex and $\forall \alpha \in \mathbb{R}^n$ are the sub-level sets $S_\alpha = \{x \mid f(x) \leq \alpha\}$ convex.
- $f : K \subseteq \mathbb{R}^n \to \mathbb{R}$ is quasi-convex $\Leftrightarrow \forall x, y \in K$, $\forall \lambda \in [0, 1]$ is the function

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

  quasi-convex.

Quasi-convex functions



Operations preserving quasi-convexity

- **weighted maximum:** if $f_1, \ldots, f_m$ are quasi-convex, $w_1, \ldots, w_m \geq 0$ - then
$$g(x) = \max\{w_1 f_1(x), \ldots, w_m f_m(x)\}$$
is quasi-convex
- **supremum:**
if $\forall y \in \mathcal{C}$ is the function $f(x, y)$ quasi-convex in x and $\sup_{y \in \mathcal{C}} f(x, y) < \infty$, then
$$g(x) = \sup_{y \in \mathcal{C}} f(x, y)$$
is quasi-convex
- **infimum:**
if $f(x, y)$ is quasi-convex in (x, y), the set $\mathcal{C} \neq \emptyset$ is **convex** and $\inf_{y \in \mathcal{C}} f(x, y) > -\infty$, then
$$g(x) = \inf_{y \in \mathcal{C}} f(x, y)$$
is quasi-convex

First order conditions

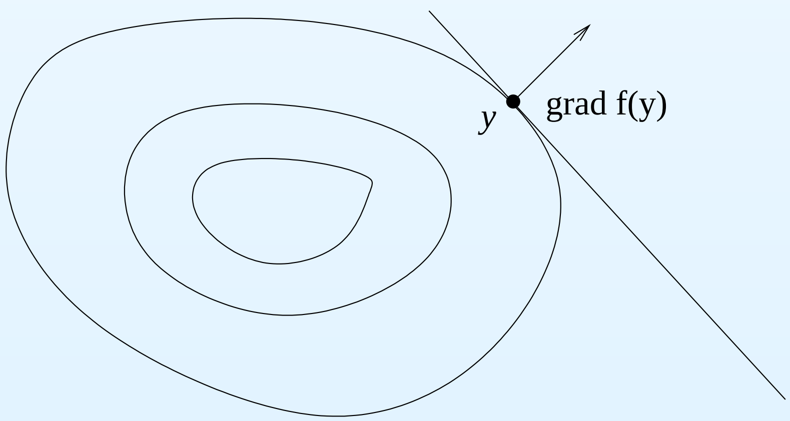
- The function $f : K \subset \mathbb{R}^n \to \mathbb{R}$ defined on a convex set K is quasi-convex **iff**

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

- Geometric interpretation: if $\nabla f(y) \neq 0$, then $\nabla f(y)$ is the normal of the supporting hyperplane of the sub-level set

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

at the point y.



Second order conditions

- If f is quasi-convex, then $\forall x, y$ it holds
$$y^T \nabla f(x) = 0 \;\Rightarrow\; y^T \nabla^2 f(x) y \geq 0$$
- If the function f satisfies $\forall x, y \neq 0$
$$y^T \nabla f(x) = 0 \;\Rightarrow\; y^T \nabla^2 f(x) y > 0,$$
then f is quasi-convex.
- $\forall x : \nabla f(x) = 0$ is $\nabla^2 f(x)$ positive (semi)definite
- If $\nabla f(x) \neq 0$, then $\nabla^2 f(x)$ is positive semidefinite on the subspace $\nabla f(x)^\perp$ - the matrix
$$H(x) = \begin{pmatrix} \nabla^2 f(x) & \nabla f(x) \\ \nabla f(x)^T & 0 \end{pmatrix}$$
has exactly one negative eigenvalue.

Strong convexity

The function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is called strong convex, if there exists $\beta > 0$ such that $\forall x \neq y$ and $\forall \lambda \in (0, 1)$ it holds

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

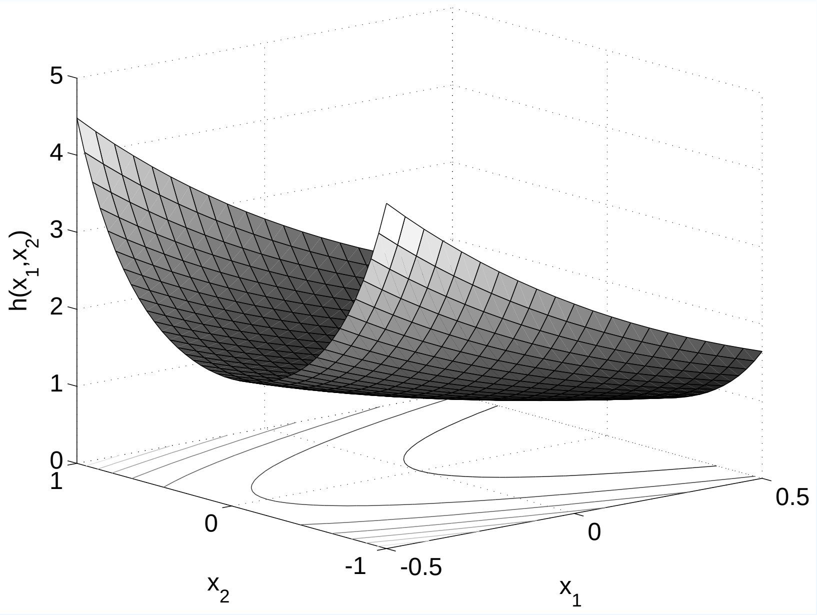
- $q(x) = \beta x^T x = \beta\|x\|_2^2,\ \beta > 0$ is the weakest strong convex function

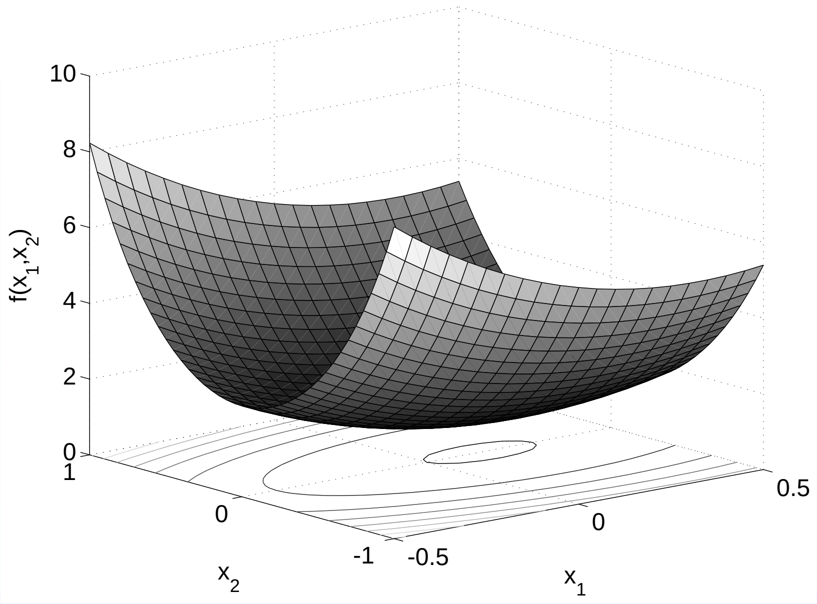
- $f(x)$ is strong convex $\Leftrightarrow$ there exists $\beta > 0$ such that the function $h(x) = f(x) - q(x)$ is convex.

- If $f(x)$ is strong convex, $\forall \alpha$ are the sub-level sets

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

  convex and compact.





Convex function

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

and strong convex function

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

First and second order conditions

The function $f(x)$ is strong convex $\Leftrightarrow$ if there exists $\beta > 0$ such that

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

Generalized convexity

The cone $\mathcal{K}$ is called **proper** if it has the following properties:

- $\mathcal{K}$ is convex;
- $\mathcal{K}$ is closed;
- $\mathcal{K}$ is solid - $\mathrm{int}(\mathcal{K}) \neq \emptyset$;
- $\mathcal{K}$ is pointed - $x \in \mathcal{K} \wedge -x \in \mathcal{K} \Rightarrow x = 0$

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

The partial ordering associated with the cone $\mathcal{K}$:

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

Properties of the generalized inequalities

property	$\preceq_K$	$\prec_K$
invariant	$x \preceq_K y,\ u \preceq_K v \Rightarrow$ $x + u \preceq_K y + v$ $x \preceq_K y, \alpha \geq 0 \Rightarrow \alpha x \preceq_K \alpha y$	$x \prec_K y,\ u \preceq_K v \Rightarrow$ $x + u \prec_K y + v$ $x \prec_K y, \alpha > 0 \Rightarrow \alpha x \prec_K \alpha y$
reflexive	$x \preceq_K x$	! $x \not\prec_K x$
transitive	$x \preceq_K y,\ y \preceq_K z \Rightarrow$ $x \preceq_K y$	$x \prec_K y,\ y \prec_K z \Rightarrow$ $x \prec_K z$
antisymmetric	$x \preceq_K y,\ y \preceq_K x \Rightarrow x = y$	—
	—	$x \prec_K y,\ \exists u, v$ small enough: $x + u \prec_K y + v$
	$x_i \preceq_K y_i, \forall i,\ x_i \to x, y_i \to y$ $\Rightarrow\ x \preceq_K y$	—

- $\mathcal{K} = \mathbb{R}^n_+$
  $x \preceq_{\mathcal{K}} y \ \Leftrightarrow \ x_i \leq y_i \ \forall i = 1, 2, \ldots, n$

- $\mathcal{K} = \mathcal{S}^n_+$
  Löwner partial ordering of symmetric matrices $\preceq$:

  - $A = Q \Lambda Q^T \preceq \alpha I$ - spectrum of the matrix A is bounded above with the constant α
  - for the positive semidefinite matrices - the inequality $0 \preceq A \preceq B$ implies $h(A) \leq h(B)$ and $\det(A) \leq \det(B)$

Generalized convexity:

Let $\mathcal{K} \subseteq \mathbb{R}^m$ be a proper cone and $\preceq_{\mathcal{K}}$ is the associated generalized inequality. Then the function $f : \mathbb{R}^n \to \mathbb{R}^m$ is called $\mathcal{K}$-**convex** if $\forall x, y \in \mathbb{R}^n$ and $\forall \lambda \in [0, 1]$ it holds

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

Example: Matrix convexity The function $f : \mathbb{R}^n \to \mathcal{S}^m$ is called matrix convex if

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

$\forall x, y \in \mathbb{R}^n$ and $\forall \lambda \in [0, 1]$.
Equivalent definition: the function $z^T f(x) z$ is convex $\forall z \in \mathbb{R}^m$. E. g. the function $f : \mathbb{R}^{n\times m} \to \mathcal{S}^n$, $f(X) = XX^T$ is matrix convex since for fixed z is the function $z^T XX^T z = \|X^T z\|^2$ convex quadratic.