Convex Optimization and Modeling

Convex functions

Third lecture, 28.04.2010

Jun.-Prof. Matthias Hein

Convex Functions:

- basic definitions, properties and some examples
- differentiability and subdifferential
- operations that preserve convexity
- conjugate function
- quasi-convex functions

Convex Functions:

Definition 1. *A function* $f : \mathbb{R}^n \to \mathbb{R}$ *is* ***convex*** *if* $\operatorname{dom} f$ *is a convex set and if for all* $x, y \in \operatorname{dom} f$*, and* $\lambda \in \mathbb{R}$ *with* $0 \le \lambda \le 1$*, we have*

$$f\big(\lambda x + (1-\lambda)y\big) \quad \le \quad \lambda f(x) + (1-\lambda) f(y).$$

A function is ***strictly convex*** *if the inequality is strict if* $x \neq y$*.*

Further definitions:

- A function is **concave** if and only if $-f$ is convex,
- A function is **strictly concave** if and only if $-f$ is strictly convex.
- An affine function is convex and concave.

Extended value formalism: A convex function f with domain $\operatorname{dom} f$ can be extended to $\mathbb{R}^n$ via

$$f(x) = \begin{cases} f(x) & x \in \operatorname{dom} f, \\ \infty & x \notin \operatorname{dom} f. \end{cases}$$

In the extended value formalism one defines:

$$\infty + x = \infty, \quad \infty + \infty = \infty.$$

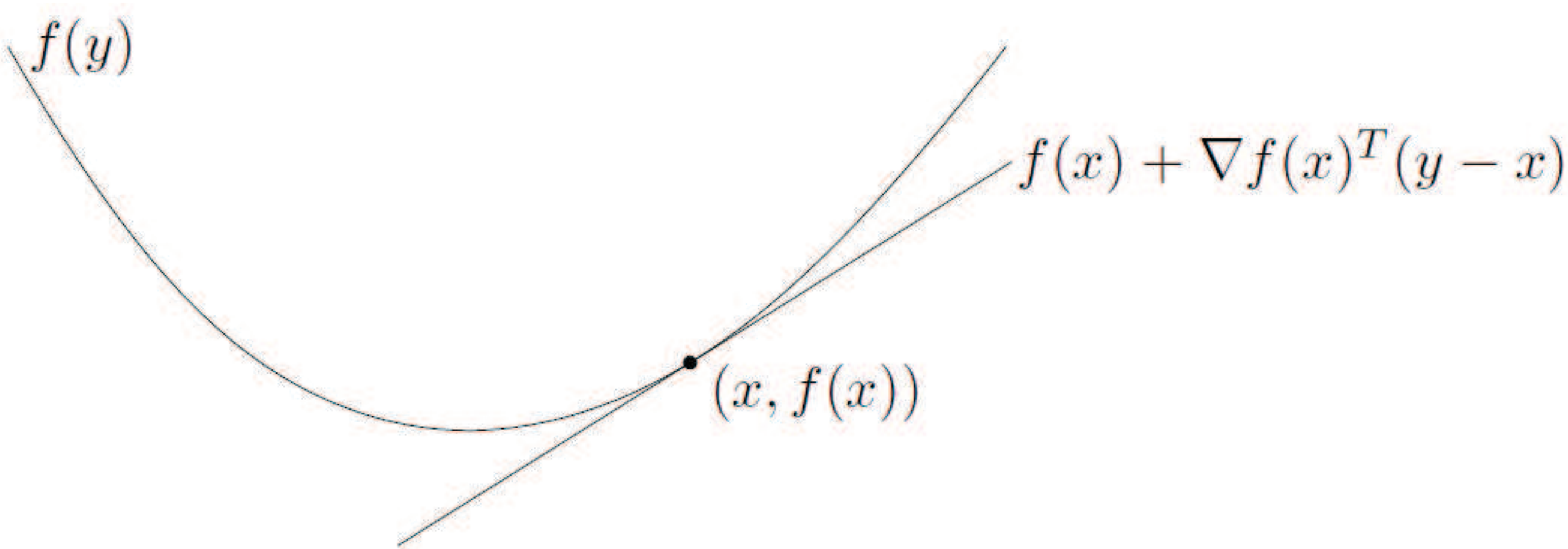
Definition 2. *The* ***indicator function*** I_C *of a convex set* C *is defined as*

$$I_C(x) = \begin{cases} 0 & x \in C \\ \infty & x \notin C. \end{cases}$$

First-Order condition - Illustration

Proposition 1. *Let f be differentiable and* $\operatorname{dom} f \subseteq \mathbb{R}^n$ *an open set. Then f is convex if and only if* $\operatorname{dom} f$ *is convex and*

$$f(y) \geq f(x) + \langle \nabla f|_x, y - x \rangle, \quad \forall y, x \in \operatorname{dom} f.$$



Proposition 2. *Let* f *be differentiable and* $\mathrm{dom}\, f \subseteq \mathbb{R}^n$ *an open set. Then* f *is convex if and only if* $\mathrm{dom}\, f$ *is convex and*

$$f(y) \geq f(x) + \langle \nabla f|_x, y - x \rangle\,, \quad \forall y, x \in \mathrm{dom}\, f.$$

Proof:

Suppose that f is convex and differentiable, then

$$f\big(x + \lambda\,(y - x)\big) \;\leq\; \lambda\, f(y) + (1 - \lambda)\, f(x) \;=\; \lambda(f(y) - f(x)) + f(x)$$

$$\Longrightarrow \quad \frac{f\big(x + \lambda\,(y - x)\big) - f(x)}{\lambda} \;\leq\; f(y) - f(x), \quad \forall\, 0 \leq \lambda \leq 1,$$

where we have used that since f is convex, $z = x + \lambda(y - x) \in \mathrm{dom}\, f$. Taking the limit $\lambda \to 0$ we obtain

$$f(y) \geq f(x) + \langle \nabla f|_x, y - x \rangle\,,$$

where we use that for $g(\lambda) = f(x + \lambda(y - x))$ we have $\frac{dg}{d\lambda}\Big|_{\lambda=0} = \langle \nabla f|_x, y - x \rangle\,.$

First-Order condition II

Proof (continued): In the other direction, let us consider $z = \lambda\, x + (1-\lambda) y$ which lies in the domain of f by assumption. Then,

$$f(x) \geq f(z) + \langle \nabla f|_z, (z-x) \rangle\,, \quad \text{and} \quad f(y) \geq f(z) + \langle \nabla f|_z, (z-y) \rangle.$$

Multiplying the first equation with λ and with $1-\lambda$ the second one and adding both yields,

$$\begin{aligned} \lambda\, f(x) + (1-\lambda)\, f(y) &\geq f(z) + \langle \nabla f|_z, z - \lambda x - (1-\lambda) y \rangle \\ &= f(z) = f(\lambda\, x + (1-\lambda) y) \end{aligned}$$

Second-order condition:

Proposition 3. *Let f be a twice continuously, differentiable function with open domain* $\operatorname{dom} f$*. Then f is convex if and only if* $\operatorname{dom} f$ *is a convex set and the Hessian of f is positive semi-definite.*

Proof: Suppose that f is convex and its domain $\operatorname{dom} f$ is convex and open. From the first-order condition,

$$f(y) \geq f(x) + \langle \nabla f|_x, y - x \rangle .$$

Using Taylor's theorem there exists θ with $0 \leq \theta \leq 1$ such that,

$$f(y) = f(x) + \langle \nabla f|_x, y - x \rangle + \frac{1}{2} \langle y - x, Hf(x + \theta(y - x))\,(y - x) \rangle .$$

Combining both results we have: $\frac{1}{2} \langle y - x, Hf(x + \theta(y - x))\,(y - x) \rangle \geq 0$.
Now, x, y are arbitrary $\Longrightarrow$ Hf is positive semi-definite on $\operatorname{dom} f$.

Proof (continued): Conversely: if $Hf \succeq 0$, then by Taylor's theorem:

$$\frac{1}{2}\langle y-x, Hf(x+\theta(y-x))(y-x)\rangle \geq 0,$$

for all $y, x \in \operatorname{dom} f$ and thus

$$f(y) \geq f(x) + \langle \nabla f|_x, y-x\rangle, \quad \forall\, x, y \in \operatorname{dom} f$$

which by the first order condition implies that f is convex.

Remarks:

- A quadratic function $f : \mathbb{R}^n \to \mathbb{R}$ with $\operatorname{dom} f = \mathbb{R}^n$, defined as

  $$f(x) = c + \langle w, x\rangle + \frac{1}{2}\langle x, Qx\rangle,$$

  is convex if and only if $Q \succeq 0$, since $Hf = Q$ for all x.
- If $Hf \succ 0$, $\forall\, x \in \operatorname{dom} f$, then f is strictly convex. However, the converse does not hold (example $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^4$).

Examples of convex functions on $\mathbb{R}$

- exponential: $e^{\alpha x}$ is convex on $\mathbb{R}$ for any $\alpha \in \mathbb{R}$,

$$f''(x) = \alpha^2 e^{\alpha x} > 0, \quad \forall x \in \mathbb{R}.$$

- powers: x^α is convex on $\mathbb{R}_{++}$ for $\alpha \geq 1$ and $\alpha \leq 0$ and concave otherwise,

$$f''(x) = (\alpha - 1)\alpha\, x^{\alpha-2} \geq 0, \quad \forall x > 0.$$

- powers of absolute value: $|x|^p$ for $p \geq 1$ is convex on $\mathbb{R}$,

$$\Rightarrow \text{sum of } f_1(x) = \begin{cases} (-x)^p & \text{for } x \leq 0 \\ 0 & x > 0 \end{cases} \text{ and } f_2(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ x^p & x > 0 \end{cases}$$

- logarithm: $f(x) = \log x$ is concave on $\mathbb{R}_{++}$ since $f''(x) = -x^{-2} < 0$ for $x > 0$,

Functions on convex subsets of $\mathbb{R}^n$

- norm: every norm on $\mathbb{R}^n$ is convex (triangle-inequality+homogenity),

$$\|\lambda x + (1-\lambda)y\| \;\le\; \|\lambda x\| + \|(1-\lambda)y\| \;\le\; \lambda\,\|x\| + (1-\lambda)\,\|y\|\,,$$

- max: $f(x) = \max_{i=1,\ldots,n} x_i$ is convex on $\mathbb{R}^n$,
- log-sum-exp: $f(x) = \log(\sum_{i=1}^n e^{x_i})$ is convex on $\mathbb{R}^n$, (differentiable approximation of the max function)
- geometric mean: $f(x) = \Big(\prod_{i=1}^n x_i\Big)^{\frac{1}{n}}$ is concave,
- log-determinant: $f(X) = \log\det X$ is concave on $\operatorname{dom} f = S^n_{++}$,

Proof that the geometric mean is concave: The Hessian of the geometric mean,

$$\frac{\partial^2 f}{\partial x_k \partial x_j} = \Big(\prod_{i=1}^n x_i\Big)^{\frac{1}{n}} \Big[\frac{1}{n^2}\frac{1}{x_j x_k} - \frac{1}{n}\frac{1}{x_j^2}\delta_{jk}\Big].$$

Now, we have a look a the quadratic form

$$\sum_{j,k=1}^n v_j v_k \frac{\partial^2 f}{\partial x_k \partial x_j} = \frac{1}{n^2}\Big(\prod_{i=1}^n x_i\Big)^{\frac{1}{n}} \Big[\sum_{j=1}^n \frac{v_j}{x_j}\sum_{k=1}^n \frac{v_k}{x_k} - n\sum_{j=1}^n \frac{v_j^2}{x_j^2}\Big] \leq 0,$$

where we used Cauchy-Schwarz inequality to the vector $a_i = (\frac{v_i}{x_i})$ and $b = \mathbf{1}$,

$$\Big(\sum_{j=1}^n \frac{v_j}{x_j}\Big)^2 = \Big(\sum_{j=1}^n \frac{v_j}{x_j}\,1\Big)^2 = \langle a, b\rangle^2 \leq \|a\|^2\,\|b\|^2 = \sum_{j=1}^n \frac{v_j^2}{x_j^2}\sum_{j=1}^n 1^2 = n\sum_{j=1}^n \frac{v_j^2}{x_j^2}.$$

Proof that the log-determinant is concave:

Consider the line $X + tV$ in S^n_{++}, that means

$$X + tV \succ 0,$$

for t small where $V \in S^n$. Then

$$g(t) = f(X + tV) = \log\det(X + tV) = \log\det(X^{\frac{1}{2}}(\mathbb{1} + tX^{-\frac{1}{2}}VX^{-\frac{1}{2}})X^{\frac{1}{2}})$$

$$= \log\det X + \log\det(\mathbb{1} + tX^{-\frac{1}{2}}VX^{-\frac{1}{2}}) = \log\det X + \sum_{i=1}^{n} \log(1 + t\lambda_i),$$

where λ_i are the eigenvalues of the symmetric matrix $X^{-\frac{1}{2}}VX^{-\frac{1}{2}}$. Thus,

$$g'(t) = \sum_{i=1}^{n} \frac{\lambda_i}{1 + t\lambda_i}, \qquad g''(t) = -\sum_{i=1}^{n} \frac{\lambda_i^2}{(1 + t\lambda_i)^2} \quad \Rightarrow \quad g''(0) \le 0.$$

Lemma 1. *Let* $f : \mathbb{R} \to \mathbb{R}$ *be convex. Then for every* $x, y, z \in \operatorname{int}\operatorname{dom} f$ *with* $x < y < z$ *we have*

$$\frac{f(y) - f(x)}{y - x} \leq \frac{f(z) - f(x)}{z - x} \leq \frac{f(z) - f(y)}{z - y}.$$

Proof. We express y as a convex combination of x and z,

$$y = \frac{z - y}{z - x} x + \frac{y - x}{z - x} z,$$

Then

$$f(y) \leq \frac{z - y}{z - x} f(x) + \frac{y - x}{z - x} f(z).$$

Thus $f(y) - f(x) \leq \frac{y-x}{z-x}(f(z) - f(x))$ which yields $\frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(x)}{z-x}$. $\square$

The lemma shows that $\frac{f(y)-f(x)}{y-x}$ is monotonically increasing for fixed y or x.

Proposition 4. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function, then* ***f is continuous in*** $\operatorname{relint} C$**.**

Proof:

- $x_0 \in \operatorname{relint} C$. Wlog let x_0 be the origin,
- e_i = corners of the $\|\cdot\|_\infty$-cube (of suff. small radius r)
- every vector x in the the cube: $x = \sum_{i=1}^m \lambda_i e_i$ with $\sum_{i=1}^m \lambda_i = 1$,

$$f(x) = f\Big(\sum_{i=1}^m \lambda_i e_i\Big) \le \sum_{i=1}^m \lambda_i f(e_i) \le M, \qquad \text{where } M = \max_{i=1,\ldots,2^m} f(e_i).$$

- Let $g(t) = f\Big(x_0 + t\frac{x-x_0}{\|x-x_0\|}\Big) \Rightarrow g(t) \le M$ for $|t| \le r$ and $g(t)$ is convex.

$$-\frac{M-g(0)}{r} \le \frac{g(-r)-g(0)}{0-r} \le \frac{g(\|x-x_0\|)-g(0)}{\|x-x_0\|-0} \le \frac{g(r)-g(0)}{r-0} \le \frac{M-g(0)}{r}$$

Thus we get $|f(x) - f(x_0)| \le \frac{M-f(x_0)}{r}\,\|x-x_0\|$.

The scalar case: The left- and right derivative at x are defined as

$$f_-(x) = \lim_{h \to 0} \frac{f(x) - f(x-h)}{h}, \qquad f_+(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

Theorem 1. *Let $f : \mathbb{R} \to \mathbb{R}$ be a convex function. Then,*

- $f_-(x) \leq f_+(x)$ *for all* $x \in \operatorname{int} \operatorname{dom} f$,
- $f_-(x)$ *and* $f_+(x)$ *are finite in the interior of* $\operatorname{dom} f$,
- *if* $x, z \in \operatorname{dom} f$ *and* $x < z$ *then* $f_+(x) \leq f_-(z)$,
- *the functions* f_- *and* f_+ *are monotonically increasing,*
- f_- *(f_+) is left (right)-continuous in the interior of* $\operatorname{dom} f$,

Corollary 1. *The* ***directional derivatives*** $\lim_{h \to 0} \frac{f(x+hv) - f(x)}{h}$ *of a convex function* $f : \mathbb{R}^n \to \mathbb{R}$ *exist for all* $x \in \operatorname{relint} \operatorname{dom} f$.

Epigraph

Definition 3. *The **graph** of a function* $f : \mathbb{R}^n \to \mathbb{R}$ *is defined as,*

$$\{(x, f(x)) \mid x \in \operatorname{dom} f\} \subseteq \mathbb{R}^{n+1}.$$

*The **epigraph** of a function* $f : \mathbb{R}^n \to \mathbb{R}$ *is defined as*

$$\operatorname{epi} f = \{(x, t) \mid x \in \operatorname{dom} f, \quad t \geq f(x)\} \subseteq \mathbb{R}^{n+1}.$$

Proposition 5. *A function* f *is convex if and only if the epigraph of* f *is a convex set.*

Proof. "$\Rightarrow$", Suppose that f is convex and let (x, t) and (y, s) be points in the epigraph, that is $t \geq f(x)$ and $s \geq f(y)$. Then

$$\lambda t + (1 - \lambda)s \geq \lambda f(x) + (1 - \lambda)f(y) \geq f(\lambda x + (1 - \lambda)y).$$

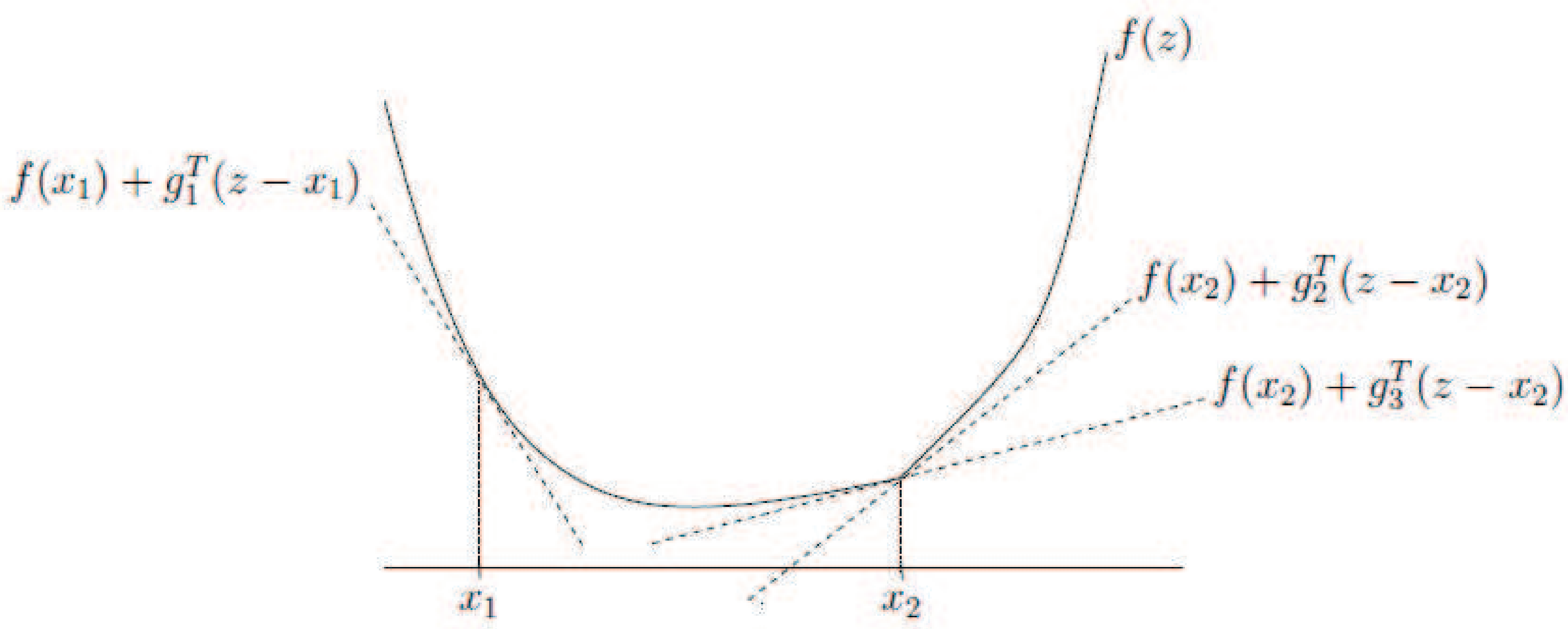
and together with the domain of f being convex the epigraph is a convex set. $\square$

Subgradient and Subdifferential: Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex.

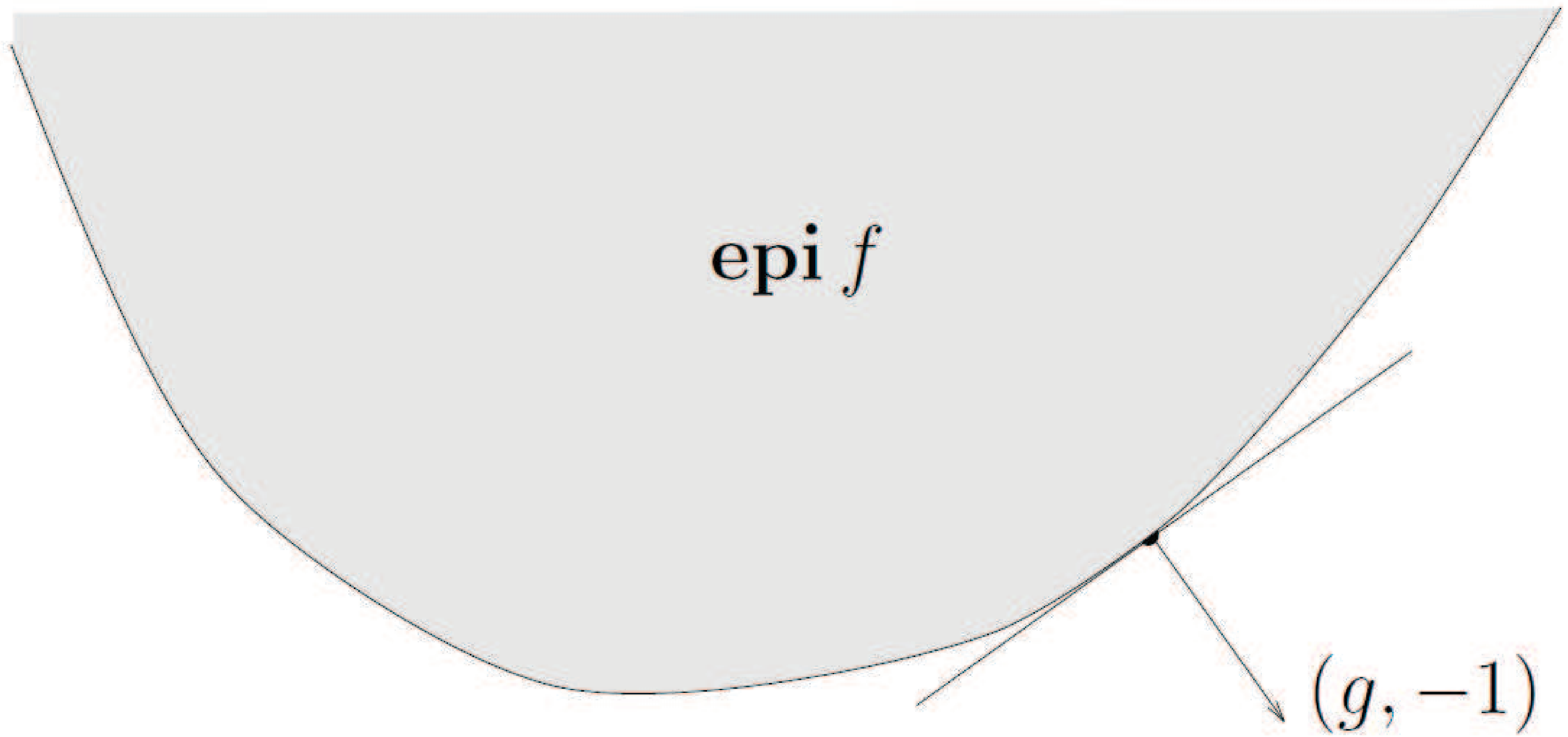
Definition 4. *A vector v is a* ***subgradient*** *of f at x if*

$$f(z) \geq f(x) + \langle v, z - x \rangle, \qquad \forall\, z \in \mathbb{R}^n.$$

The subdifferential $\partial f(x)$ of f at x is the set of all subgradients of f at x.



Subdifferential as supporting hyperplane of the epigraph



Examples

- $f(x) = |x|, \quad \partial f(x) = \begin{cases} -1 & \text{if } x < 0, \\ [-1, 1] & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases}.$

- $f(x) = \|x\|_2, \quad \partial f(x) = \begin{cases} \frac{x}{\|x\|}, & \text{if } x \neq 0, \\ \{u \in \mathbb{R}^n \mid \|u\| \leq 1\}, & \text{if } x = 0. \end{cases}.$

  This follows by Cauchy-Schwarz,

$$\|x\|_2 \geq \langle u, x \rangle = 0 + \langle u, x - 0 \rangle \,,$$

  for $\|u\|_2 \leq 1$.

Properties of the subdifferential

- The subdifferential $\partial f(x)$ is a closed convex set,
- If f is differentiable, $\partial f(x) = \{\nabla f(x)\}$,
- $f(x) = \sum_{i=1}^{k} \alpha_i f_i(x)$ with $f : \mathbb{R}^n \to \mathbb{R}$,

$$\Longrightarrow \partial f(x) = \alpha_1 \partial f_1(x) + \ldots + \alpha_k \partial f_k(x).$$

- chain rule $f(x) = g(Ax + b)$,

$$\partial f(x) = A^T \partial g(Ax + b).$$

- The subdifferential ∂f is non-empty in the relative interior of $\operatorname{dom} f$.

Sublevel sets

Definition 5. *The* ***α-sublevel set*** *of a function $f : \mathbb{R}^n \to \mathbb{R}$ is defined as*

$$L_\alpha = \{x \in \operatorname{dom} f \mid f(x) \leq \alpha\}.$$

Proposition 6. *The sublevel set of a convex function is convex.*

Proof. Suppose $x, y \in L_\alpha$ that is $f(x) \leq \alpha$ and $f(y) \leq \alpha$, then

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

and thus $\lambda x + (1-\lambda)y \in L_\alpha$ for each $0 \leq \lambda \leq 1$. □

The converse is in general false $\Longrightarrow$ sublevel sets do not characterize convex functions.

Jensen's inequality and extensions

The inequality for a convex function can be extended:

$$f\Big(\sum_{i=1}^{n} \lambda_i x_i\Big) \le \sum_{i=1}^{n} \lambda_i f(x_i) \quad (\textbf{Jensen's inequality}),$$

where $\sum_{i=1}^{n} \lambda_i = 1$ and $x_1, \ldots, x_n \in \mathrm{dom}\, f$.

Extension to probability measures on a convex domain S:

$$f\Big(\int_S x\, p(x) dx\Big) \le \int_S p(x) f(x) dx \quad \Longleftrightarrow \quad f(\mathbb{E}[X]) \le \mathbb{E} f(X)$$

Application of Jensen's inequality with some special convex functions yields interesting other inequalities, e.g. Hölder's inequality

$$\sum_{i=1}^{n} x_i y_i \le \|x\|_p \, \|y\|_q\,, \qquad \text{where } p > 1 \text{ and } \frac{1}{p} + \frac{1}{q} = 1.$$

Operations that conserve convexity

- **Weighted sum:** Let $f_i,\ i = 1, \ldots, n$ be convex, then $\sum_{i=1}^n w_i f_i(x)$ is convex again if $w_i \geq 0,\ i = 1, \ldots, n$. This can be extended to an integral formulation $g(x) = \int_S w(y) f(y, x)$.
- **Composition with an affine mapping:** $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $A \in \mathbb{R}^{n \times m}$ and $b \in \mathbb{R}^n$. Define

$$g(x) = f(Ax + b), \text{ where } \operatorname{dom} g = \{x \mid Ax + b \in \operatorname{dom} f\}.$$

  Then if f is convex, also g is convex.
- **Pointwise Maximum and Supremum:** If f_1, f_2 are convex, then

$$f(x) = \max\{f_1(x), f_2(x)\},$$

  is convex. Extension to **pointwise supremum:** Let $f(x, y)$ be convex for each $y \in S$, then $g(x) = \sup_{y \in S} f(x, y)$ is convex.

Operations that conserve convexity II

- **Composition:** Let $h : \mathbb{R}^k \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^k$ and $f = h \circ g : \mathbb{R}^n \to \mathbb{R}$,

$$f(x) = h(g(x)), \qquad \operatorname{dom} f = \{x \in \operatorname{dom} g \,|\, g(x) \in \operatorname{dom} h\}.$$

  Suppose: h and g are twice continuously differentiable and $n = 1$,

$$f''(x) = h''(g(x))g'(x)^2 + h'(g(x))g''(x),$$

  Thus for $f''(x) \geq 0$ that means f is convex if

  h is convex and nondecreasing   g is convex,

  h is convex and nonincreasing   g is concave.

- **Minimization:** If f is convex in (x, y) and C is a convex, nonempty set,

$$g(x) = \inf_{y \in C} f(x, y), \text{ is convex if } g(x) > -\infty \text{ for some } x.$$

Operations that conserve convexity III

Usage of these rules: Pointwise maximum: the sum of the r largest components. Let $x \in \mathbb{R}^n$ and $x_{[1]} \geq x_{[2]} \geq \ldots \geq x_{[n]}$. Then

$$f(x) = \sum_{i=1}^{r} x_{[i]},$$

is convex since f is the maximum of all linear combinations of r components.

Operations that conserve convexity III

Usage of these rules: Pointwise maximum: the sum of the r largest components. Let $x \in \mathbb{R}^n$ and $x_{[1]} \geq x_{[2]} \geq \ldots \geq x_{[n]}$. Then

$$f(x) = \sum_{i=1}^{r} x_{[i]},$$

is convex since f is the maximum of all linear combinations of r components.

Maximal eigenvalue of a symmetric matrix: $f(X) = \lambda_{\max}(X)$, since

$$f(X) = \sup\{\langle y, Xy \rangle \mid \|y\| = 1\}.$$

Usage of these rules: Pointwise maximum: the sum of the r largest components. Let $x \in \mathbb{R}^n$ and $x_{[1]} \geq x_{[2]} \geq \ldots \geq x_{[n]}$. Then

$$f(x) = \sum_{i=1}^{r} x_{[i]},$$

is convex since f is the maximum of all linear combinations of r components.

Maximal eigenvalue of a symmetric matrix: $f(X) = \lambda_{\max}(X)$, since

$$f(X) = \sup\{\langle y, Xy\rangle \mid \|y\| = 1\}.$$

Proposition 9. *For every* $x \in \operatorname{int} \operatorname{dom} f$ *we have*

$$f(x) = \sup\{g(x) \,|\, g \textit{ affine }, g(z) \leq f(z) \textit{ for all } z\}.$$

In the interior of dom f, f is the supremum of all affine functions which globally underestimate f.

The conjugate function

Definition 6. *Let* $f : \mathbb{R}^n \to \mathbb{R}$. *The function* $f^* : \mathbb{R}^n \to \mathbb{R}$ *defined as*

$$f^*(y) = \sup_{x \in \mathrm{dom}\, f} \Big(\langle y, x \rangle - f(x) \Big),$$

is the ***conjugate*** *of the function* f. *The domain of* f^* *consists of* $y \in \mathbb{R}^n$ *such* $f^*(y) < \infty$.

Remarks:

- f^* is convex since it is the pointwise supremum of a set of affine functions. This holds independently of the fact that f is convex or not,
- the conjugate function is also known as the **Legendre-Fenchel transform**,
- the conjugate of the conjugate function is denoted by f^{**}.

Why the name "conjugate" ?

Let $f \in C^2(\mathbb{R}^n)$ with a positive-definite Hessian and convex, then

$$g(x) = \langle y, x \rangle - f(x), \quad \nabla g = y - \nabla f, \quad Hg = -Hf,$$

and we know by convexity of f that g has a unique maximum x^* at which holds: $y = \nabla f$. Moreover,

$$f^*(y) = g(x^*) = \langle y, x^* \rangle - f(x^*).$$

and thus $\nabla f^*|_y = x^* \implies$ conjugation interchanges derivative and position. If f has a slope of y at x, then f^* has slope of x at y.

Theorem 2. *If* $\operatorname{epi} f$ *is closed and convex, then* $f = f^{**}$.

In particular: if f is convex and $\operatorname{dom} f = \mathbb{R}^n$ or if f is convex and continuous then $f^{**} = f$.

Examples:

- The conjugate of an affine function $f : \mathbb{R}^n \to \mathbb{R}$ with $f(x) = \langle w, x \rangle + b$ is given as

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \langle y, x \rangle - \langle w, x \rangle - b = \sup_{x \in \mathbb{R}^n} \langle y - w, x \rangle - b = \begin{cases} -b & \text{if } y = w, \\ \infty & \text{else.} \end{cases}.$$

- Let $f : \mathbb{R}^n \to \mathbb{R}$ with $f(x) = \frac{1}{2} \langle x, Qx \rangle$ where $Q \in S^n_{++}$, then

$$f^*(y) = \frac{1}{2} \langle y, Q^{-1} y \rangle .$$

Note, that $y = \nabla f = Qx^*$, so that $x^* = Q^{-1} y$, then
$f^*(y) = \langle y, Q^{-1} y \rangle - \frac{1}{2} \langle y, Q^{-1} y \rangle = \frac{1}{2} \langle y, Q^{-1} y \rangle$.

Conjugate Function II

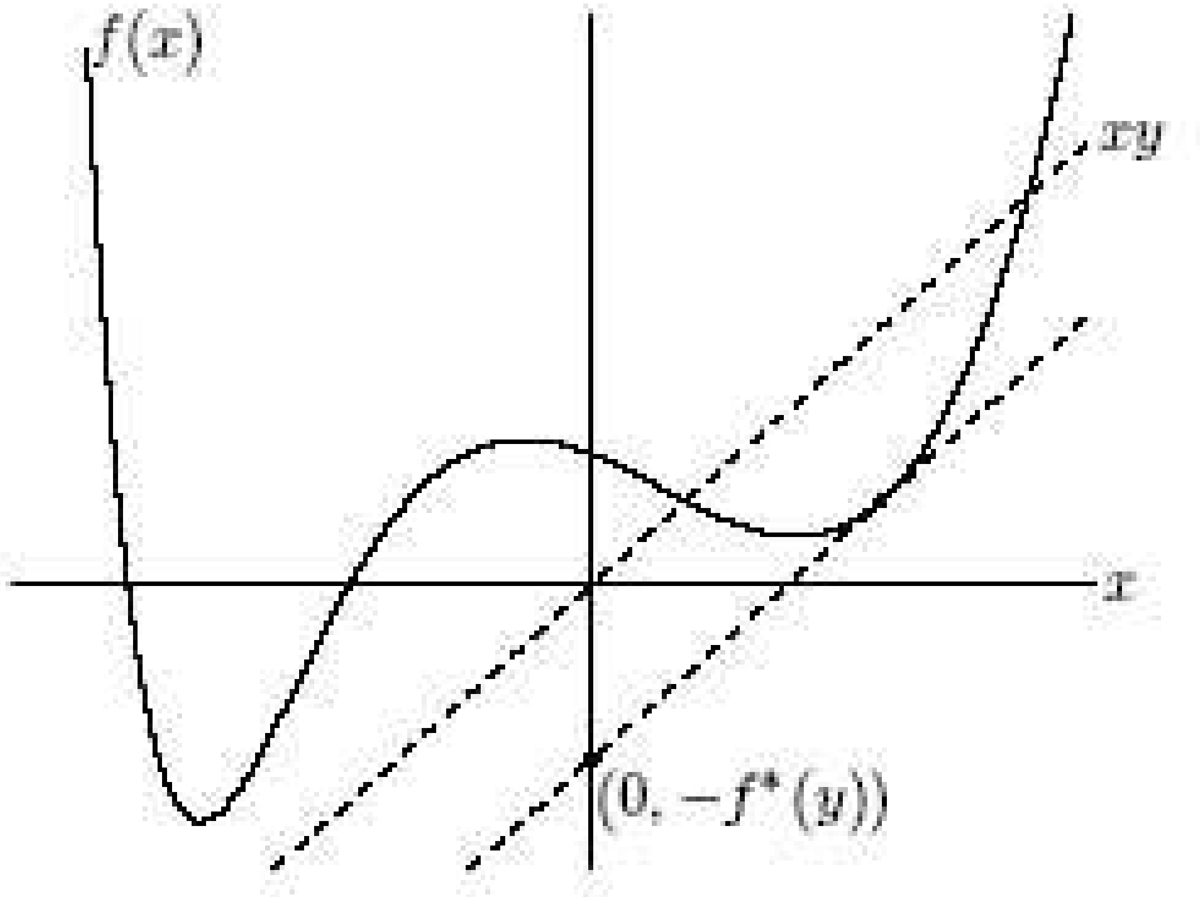


Figure 1: A function f together with the value of the conjugate function f^* at y.

Quasiconvex functions

Definition 7. *A function $f : \mathbb{R}^n \to \mathbb{R}$ is called* ***quasiconvex*** *if its domain and all sublevel sets,*

$$L_\alpha = \{x \in \operatorname{dom} f \mid f(x) \leq \alpha\},$$

for $\alpha \in \mathbb{R}$ are convex. A function is ***quasiconcave*** *if $-f$ is quasiconvex. A function that is both quasiconvex and quasiconcave is called* ***quasilinear****.*

Any convex function is quasiconvex but the converse does not hold.

Quasiconvex functions II

Properties of quasiconvex functions:

- Jensen's inequality for quasiconvex functions

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

- A continuous function $f : \mathbb{R} \to \mathbb{R}$ is quasiconvex if and only if at least one of the following conditions holds
  1. f is nondecreasing or f is nonincreasing,
  2. $\exists c \in \operatorname{dom} f$ s.th. for $t \leq c$, f is nonincreas. and for $t \geq c$, f is nondecreas.
  3. First-order condition: Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable. Then f is quasiconvex if and only if $\operatorname{dom} f$ is convex and for all $x, y \in \operatorname{dom} f$

$$f(y) \leq f(x) \Rightarrow \langle \nabla f|_x, y - x \rangle \leq 0.$$

  4. for a quasicvx function $\nabla f = 0$ at x^* does not imply that x^* is a global minimum of f.

Quasiconvex functions III

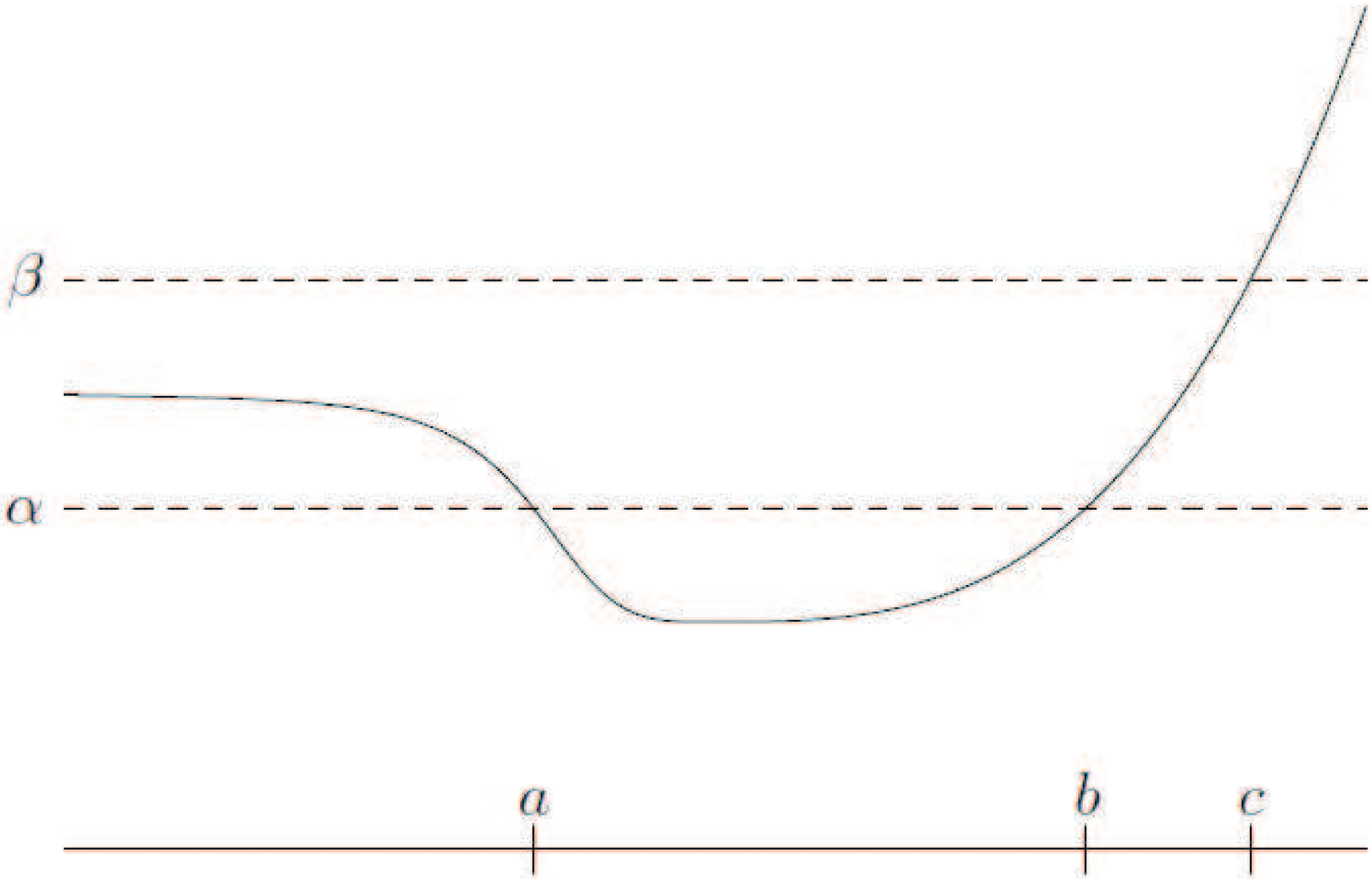


Figure 3.9 A quasiconvex function on $\mathbf{R}$. For each α, the α-sublevel set S_α is convex, *i.e.*, an interval. The sublevel set S_α is the interval $[a, b]$. The sublevel set S_β is the interval $(-\infty, c]$.

Quasiconvex functions IV

Representation via family of convex functions:

Seek a family of convex functions $\phi_t : \mathbb{R}^n \to \mathbb{R}$, indexed by $t \in \mathbb{R}$, with

$$f(x) \leq t \quad \Longrightarrow \quad \phi_t(x) \leq 0,$$

the t-sublevel set of the quasiconvex function f is the 0-sublevel set of the convex function ϕ_t. Thus ϕ_t must satisfy

$$\phi_t(x) \leq 0 \quad \Longrightarrow \quad \phi_s(x) \leq 0, \qquad \text{for } s \geq t,$$

e.g. this holds if $\phi_t(x)$ is a non-increasing function of t for all x.
Such representation always exists, with

$$\phi_t(x) = \begin{cases} 0 & \text{if } f(x) \leq t \\ \infty & \text{otherwise} \end{cases}.$$

Optimization: need for a family with nice properties e.g. differentiability.