Convex functions

A supplementary note to Chapter 3 of *Convex Optimization* by S. Boyd and L. Vandenberghe

Optimization Lab.

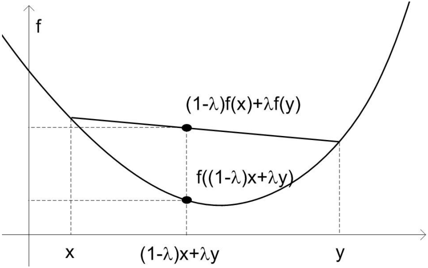
IE department
Seoul National University

28th September 2009

Definition

A function $f : \mathbb{R}^n \to \mathbb{R}$ is *convex* if its domain $\mathrm{dom} f$ is convex and if for all $x, y \in \mathrm{dom} f$, and $0 \leq \lambda \leq 1$, we have

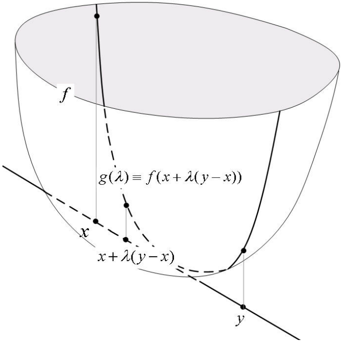
$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y).$$



- *Strictly convex* if strict inequality holds whenever $x \neq y$ and $0 < \lambda < 1$.
- We say f is *concave* if $-f$ is convex, and *affine* if both convex and concave.

"One-dimensionality" of convexity

From definition, a function f is convex iff its restriction to any line is convex: for any $x, y \in \mathrm{dom} f$, $g(\lambda) := f(x + \lambda(y - x))$ is convex over $\{\lambda | x + \lambda(y - x) \in \mathrm{dom} f\}$



Extended-value extensions

If f is convex we define its extended-value extension,

$$\tilde{f}(x) = \begin{cases} f(x) & x \in \mathrm{dom} f \\ \infty & x \notin \mathrm{dom} f \end{cases}$$

With the extended reals, this can simplify notation, since we do not need to explicitly describe the domain.

Example

For a convex set C, its indicator function I_C is defined to be $I_C(x) = 0$ for all $x \in C$. Then its extension is

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

Suppose $\mathrm{dom} f = \mathbb{R}^n$. Then, $\min\{f(x) : x \in C\}$ is equivalent to minimizing $f + \tilde{I}_C$.

First-order conditions

Theorem

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable. Then f is convex iff $dom f$ is convex and

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) \ \forall x, y \in dom f.$$

Proof Case 1: $n = 1$.

"Only if" Assume f is convex and $x, y \in \text{dom} f$ with $x \neq y$. Since $\text{dom} f$ is convex, we have for all $0 < \lambda \leq 1$, $x + \lambda(y - x) \in \text{dom} f$, and by convexity of f, $f(x + \lambda(y - x)) \leq (1 - \lambda)f(x) + \lambda f(y)$.

Dividing both sides by λ, we obtain

$$f(y) \geq f(x) + \frac{f(x + \lambda(y - x)) - f(x)}{\lambda}.$$

Taking limit as $\lambda \to 0$, we get $f(y) \geq f(x) + f'(x)(y - x)$.

First-order conditions (*cont'd*)

"If" Choose any x, $y \in \mathrm{dom} f$ and $0 \leq \lambda \leq 1$, and let $z = \lambda x + (1-\lambda)y$. Then, by the above,

$$f(x) \geq f(z) + f'(z)(x-z), \qquad f(y) \geq f(z) + f'(z)(y-z).$$

Multiplying the first inequality by λ, the second by $1-\lambda$, and adding them yields

$$\lambda f(x) + (1-\lambda)f(y) \geq f(z) = f(\lambda x + (1-\lambda)y).$$

Case 2: $n \geq 2$. Let $x, y \in \mathrm{dom} f$. Consider restriction of f to line through x and y: $g(\lambda) := f(x + \lambda(y-x))$, and apply the above case. $\square$

Second-order conditions (*cont'd*)

Proposition

Assume f is twice differentiable on dom$f$ which is open. Then f is convex if and only if dom$f$ is convex and its Hessian is positive semidefinite: $\forall x \in$ dom$f$,

$$\nabla^2 f(x) \succeq 0.$$

Remark that

> for $y \in \mathrm{dom} f$ and $z \in \mathbb{R}^n$, define $g(\lambda) := f(y + \lambda z)$. Then $g''(\lambda) = z^T \nabla^2 f(y + \lambda z) z$. Thus, $g''(\lambda) \geq 0$ on $\{\lambda | y + \lambda z \in \mathrm{dom} f\}$ if and only if $\nabla^2 f(x) \succeq 0 \; \forall x \in \mathrm{dom} f$.

Second-order conditions (*cont'd*)

Proof Case 1 $f : \mathbb{R} \to \mathbb{R}$
"Only if" If f is convex, then $f(y) \geq f(x) + f'(x)(y - x)$ for all $x, y \in \mathrm{dom} f$, where $x < y$. Thus,

$$\frac{f(y) - f(x)}{y - x} \geq f'(x).$$

Taking limit as $x \longrightarrow y$, we get $f'(y) \geq f'(x)$, which implies that f' is monotone nondecreasing. Hence, $f''(x) \geq 0, \forall x \in \mathrm{dom} f$.
"If" For all $x, y \in \mathrm{dom} f$, there exists $z \in \mathrm{dom} f$ satisfying

$$f(y) = f(x) + f'(x)(y - x) + \frac{1}{2} f''(z)(y - x)^2 \geq f(x) + f'(x)(y - x).$$

The second inequality follows from the hypothesis. Hence f is convex.

Second-order conditions (*cont'd*)

Case 2 $f : \mathbb{R}^n \to \mathbb{R}$
f is convex if and only if $g(\lambda) = f(x + \lambda y)$ is convex on $\{\lambda | x + \lambda y \in \mathrm{dom} f\}$, $\forall x \in \mathrm{dom} f$ and y. Then, by **Case 1**, the latter holds if and only if $g''(\lambda) \geq 0$ on $\{\lambda | x + \lambda y \in \mathrm{dom} f\}$:

$$\begin{aligned} g''(\lambda) &= \frac{d}{d\lambda} g'(t) = \frac{d}{d\lambda}\left(\sum_{i=1}^{n} f_i'(x + \lambda y) y_i\right) \\ &= \sum_{i=1}^{n} y_i \frac{d}{d\lambda} f_i(x + \lambda y) = \sum_{i=1}^{n} y_i \nabla^2 f(x + \lambda y)_{i\cdot} y \\ &= y^T \nabla^2 f(x + \lambda y) y \geq 0, \end{aligned}$$

where $\nabla^2 f(x)_{i\cdot}$ is the i-th row of $\nabla^2 f(x)$. Therefore, $\nabla^2 f(x) \succeq 0$ for all $x \in \mathrm{dom} f$. $\square$

Some simple examples

Example

- Exponential e^{ax} is convex on $\mathbb{R}$ for $a \in \mathbb{R}$.
- Powers x^a are convex on $\mathbb{R}_{++}$ for $a \geq 1$ or $a \leq 0$, and concave for $0 \leq a \leq 1$.
- Powers of absolute value, $|x|^p$ for $p \geq 1$, is convex on $\mathbb{R}$.
- Logarithm $\log x$ is convex on $\mathbb{R}_{++}$.
- Negative entropy $x \log x$ is convex on $\mathbb{R}_{++}$. (Also on $\mathbb{R}_+$ if defined as 0 for $x = 0$.)

Max function

Max function, $f(x) = \max\{x_1, \ldots, x_n\}$ is convex on $\mathbb{R}^n$.

Proof

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &= \max_i \{\lambda x_i + (1-\lambda) y_i\} \\ &\leq \lambda \max_i x_i + (1-\lambda) \max_i y_i \\ &= \lambda f(x) + (1-\lambda) f(y). \square \end{aligned}$$

Log-sum-exp

Log-sum-exp function $f(x) = \log(e^{x_1} + \cdots + e^{x_n})$ is convex on $\mathbb{R}^n$.

Proof The Hessian of the log-sum-exp function is

$$\nabla^2 f(x) = \frac{1}{(\mathbf{1}^T z)^2}\left((\mathbf{1}^T z)\mathrm{Diag}(z) - zz^T\right),$$

where $z = (e^{x_1}, \ldots, e^{x_n})$. We must show that for all v, $v^T \nabla^2 f(x) v \geq 0$, but

$$v^T \nabla^2 f(x) v = \frac{1}{(\mathbf{1}^T z)^2}\left(\left(\sum_{i=1}^n z_i\right)\left(\sum_{i=1}^n v_i^2 z_i\right) - \left(\sum_{i=1}^n v_i z_i\right)^2\right) \geq 0.$$

The inequality follows from the Cauchy-Schwarz inequality $(a^T a)(b^T b) \geq (a^T b)^2$ applied to $a_i = \sqrt{z_i}$ and $b_i = v_i\sqrt{z_i}$. □

Sublevel sets and graphs

Definition

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

$$C_\alpha = \{x \in \mathrm{dom} f | f(x) \le \alpha\}.$$

Sublevel sets of a convex function are convex. (Converse is false.)

Definition

The *graph* of a function $f : \mathbb{R}^n \to \mathbb{R}$ is $\{(x, f(x)) | x \in \mathrm{dom} f\}$.
The *epigraph* of f is $\mathrm{epi} f = \{(x, t) | x \in \mathrm{dom} f, f(x) \le t\}$.
The *hypograph* of f is $\mathrm{hyp} f = \{(x, t) | x \in \mathrm{dom} f, f(x) \ge t\}$.

A function is convex (concave) if and only if its epigraph (hypograph, resp.) is convex.

Epigraph and convex function

Consider the first-order condition for convexity: $\forall\ x, y \in \mathrm{dom} f$, $f(y) \geq f(x) + \nabla f(x)^T(y-x)$. Thus, if $(y,t) \in \mathrm{epi} f$, then $t \geq f(y) \geq f(x) + \nabla f(x)^T(y-x)$. Hence $\nabla f(x)^T(y-x) - (t - f(x)) \leq 0$. Thus,

$$(x,t) \in \mathrm{epi} f \Rightarrow \begin{bmatrix} \nabla f(x) \\ -1 \end{bmatrix}^T \left(\begin{bmatrix} y \\ t \end{bmatrix} - \begin{bmatrix} x \\ f(x) \end{bmatrix} \right) \leq 0,$$

which means hyperplane in $\mathbb{R}^{n+1}$ defined by $(\nabla f(x), -1)$ supports $\mathrm{epi} f$ at the boundary point $(x, f(x))$.

Nonnegative scaling preserves convexity.
Proof If $w \geq 0$ and f is convex, we have

$$\mathrm{epi}(wf) = \begin{bmatrix} I & 0 \\ 0 & w \end{bmatrix} \mathrm{epi} f,$$

which is convex because the image of a convex set under a linear mapping is convex. $\square$

If $f_1, \ldots, f_m$ are convex functions, then $\forall\ w_i \geq 0$, $i = 1, \ldots, m$, $f = w_1 f_1 + \cdots + w_m f_m$ is convex.

Suppose $f : \mathbb{R}^n \to \mathbb{R}$, $A \in \mathbb{R}^{n\times m}$, and $b \in \mathbb{R}^n$. Define $g : \mathbb{R}^m \to \mathbb{R}$ by

$$g(x) = f(Ax+b), \text{ with } \mathrm{dom} g = \{x | Ax+b \in \mathrm{dom} f\}.$$

Then, if f is convex, so is g; if f is concave, so is g.
Proof Suppose $(x,s)^T$, $(y,t)^T \in \mathrm{epi} g$ satisfy $f(Ax+b) \le s$ and $f(Ay+b) \le t$. Then,

$$f\Big(A((1-\lambda)x+\lambda y)+b\Big) = f\Big((1-\lambda)(Ax+b)+\lambda(Ay+b))\Big)$$
$$\le (1-\lambda)f(Ax+b)+\lambda f(Ay+b) \le (1-\lambda)s+\lambda t.$$

Thus $(1-\lambda)(x,s)^T + \lambda(y,t)^T \in \mathrm{epi} g$. □

If f_1 and f_2 are convex functions, then so is their *pointwise maximum*,

$$f(x) = \max\{f_1(x), f_2(x)\} \text{ with } \mathrm{dom} f = \mathrm{dom} f_1 \cap \mathrm{dom} f_2.$$

Proof $0 \leq \lambda \leq 1$ and $x, y \in \mathrm{dom} f$,

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &= \max\{f_1(\lambda x + (1-\lambda)y), f_2(\lambda x + (1-\lambda)y)\} \\ &\leq \max\{\lambda f_1(x) + (1-\lambda)f_1(y), \lambda f_2(x) + (1-\lambda)f_2(y)\} \\ &\leq \max\{\lambda f_1(x), \lambda f_2(x)\} + \max\{(1-\lambda)f_1(y), (1-\lambda)f_2(y)\} \\ &= \lambda \max\{f_1(x), f_2(x)\} + (1-\lambda)\max\{f_1(y), f_2(y)\} \\ &= \lambda f(x) + (1-\lambda)f(y). \quad \square \end{aligned}$$

Or, easy to see $\mathrm{epi} f = \mathrm{epi} f_1 \cap \mathrm{epi} f_2$.

If for each $y \in \mathcal{A}$, $f(x, y)$ is convex in x, then the function g, defined as

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

is convex in x. ($\mathrm{dom} g = \{x | \ (x, y) \in \mathrm{dom} f \ \forall y \in \mathcal{A}, \sup_{y \in \mathcal{A}} f(x, y) < \infty\}$)

Application

- Support function of a set, $S_C(x) = \sup\{x^T y | y \in C\}$ is convex.
- Distance to farthest point of a set, $f(x) = \sup_{y \in C} \|x - y\|$ is convex.
- Least-squares as function of weights $g(w) = \inf_x \sum_{i=1}^n w_i(a_i^T x - b_i)^2$ with $\mathrm{dom} g = \{w | \inf_x \sum_{i=1}^n w_i(a_i^T x - b_i)^2 > -\infty\}$. Needs proof.
- Max eigenvalue of symm matrices $f(X) = \sup\{y^T X y | \ \|y\|_2 = 1\}$.
- Norm of a matrix

Convex as pointwise affine supremum

If $f : \mathbb{R}^n \to \mathbb{R}$ is convex, with $\mathrm{dom} f = \mathbb{R}^n$, then we have

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

Proof $(\geq)$ Easy.
$(\leq)$ For any x we can find a supporting hyperplane of $\mathrm{epi} f$ at $(x, f(x))$: $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$ with $(a, b) \neq 0$ such that $\forall\ (z, t) \in \mathrm{epi} f$,

$$\begin{bmatrix} a \\ b \end{bmatrix}^T \begin{bmatrix} x - z \\ f(x) - t \end{bmatrix} \leq 0. \text{ Or, } \ a^T(x - z) + b(f(x) - f(z) - s) \leq 0,$$

for all $z \in \mathrm{dom} f = \mathbb{R}^n$ and all $s \geq 0$. This implies $b > 0$ as easily seen. Therefore,

$$g(z) = f(x) + (a/b)^T(x - z) \leq f(z)$$

for all z. The function g is an affine underestimator of f and satisfies $g(x) = f(x)$. $\square$

Chain rule: Review

Consider a twice differentiable $f : \mathbb{R}^n \to \mathbb{R}^m$ whose dom$f$ is assumed to be open for simplicity.

- For $m = 1$, the *derivative* $Df : \mathbb{R}^n \to \mathbb{R}$ of f at x is defined to be

$$Df(x) = \left[\; D_1 f(x) \cdots D_n f(x) \; \right].$$

  A linear transformation from $\mathbb{R}^n$ to $\mathbb{R}$ which linearly approximates f at x.

- For $m \geq 2$, the *derivative* of f at x is defined to be

$$Df(x) = \begin{bmatrix} Df_1(x) \\ \vdots \\ Df_m(x) \end{bmatrix}.$$

  A linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ which linearly approximates f at x.

Chain rule: Review(*cont'd*)

- For $m = 1$, we define the *gradient* of f is a column-wise representation of its derivative:

$$\nabla f(x) = \begin{bmatrix} D_1 f(x) \\ \vdots \\ D_n f(x) \end{bmatrix},$$

  a function from $\mathbb{R}^n \to \mathbb{R}^n$.

- For $m = 1$, the *Hessian* $\nabla^2 f(x)$ of f is defined to be the derivative of the gradient ∇f

$$\nabla^2 f(x) = \begin{bmatrix} D_{11} f(x) & \cdots & D_{1n} f(x) \\ \vdots & \ddots & \vdots \\ D_{n1} f(x) & \cdots & D_{nn} f(x) \end{bmatrix}.$$

Chain rule: Review(*cont'd*)

Suppose that $h : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable at $x \in \mathrm{dom} h$, and that $g : \mathbb{R}^m \longrightarrow \mathbb{R}^p$ is differentiable at $h(x) \in \mathrm{dom} g$. (Assume domains are open.) Let $f := g \circ h : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ by $(g \circ h)(x) = g(h(x))$. Then, f is differentiable at x and its derivative is

$$Df(x) = D(g \circ h)(x) = Dg(h(x))Dh(x).$$

Convexity conditions of composition

- Let $p = 1$ and we consider case $m = 1$. For the convexity conditions of composition, it suffices to consider one-dimensional cases: $n = 1$. Assume g, h twice differ'ble, $\mathrm{dom} g = \mathrm{dom} h = \mathbb{R}$. Then

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

  - g convex, nondecreasing, h convex $\Rightarrow$ f convex,
  - g convex, nonincreasing, h concave $\Rightarrow$ f convex,
  - g concave, nondecreasing, h concave $\Rightarrow$ f concave,
  - g concave, nonincreasing, h convex $\Rightarrow$ f concave.

Composition(*cont'd*)

- In general,
  - g convex, $\tilde{g}$ nondecreasing, h convex $\Rightarrow f$ convex,
  - g convex, $\tilde{g}$ nonincreasing, h concave $\Rightarrow f$ convex,
  - g concave, $\tilde{g}$ nondecreasing, h concave $\Rightarrow f$ concave,
  - g concave, $\tilde{g}$ nonincreasing, h convex $\Rightarrow f$ concave.

Example

- $g(x) = \log(x)$, then g concave, $\tilde{g}$ nondecreasing
- $g(x) = x^{1/2}$, then g concave, $\tilde{g}$ nondecreasing
- $g(x) = x^{3/2}$, then g convex, $\tilde{g}$ not nondecreasing
- $g(x) = x^{3/2}$ for $x \geq 0$, $= 0$ for $x < 0$ then g convex, $\tilde{g}$ nondecreasing.

Composition(*cont'd*)

Proposition

g convex, $\tilde{g}$ nondecreasing, h convex $\Rightarrow$ f convex.

Proof: $\square$

The monotonicity of $\tilde{g}$ is to guarantee convexity of $h^{-1}(\mathrm{dom} g)$. (Then $\mathrm{dom} f = \mathrm{dom} h \cap h^{-1}(\mathrm{dom} g)$ is convex.) Without it, $h^{-1}(\mathrm{dom} g)$ is not convex in general: for instance $h(x) = x^2$, $g(x) = x$ with domain $1 \leq x \leq 2$.

Composition(*cont'd*)

Example

- h convex $\Rightarrow \exp h$ convex.
- h concave, positive $\Rightarrow \log h$ concave.
- h concave, positive $\Rightarrow 1/h(x)$ concave.
- h convex, nonnegative, and $p \geq 1 \Rightarrow h(x)^p$ convex.
- h convex $\Rightarrow -\log(-h(x))$ convex on $\{x|h(x) < 0\}$.

Composition(*cont'd*)

Consider $g : \mathbb{R}^m \to \mathbb{R}$ and $h : \mathbb{R} \to \mathbb{R}^m$ with dom$g = \mathbb{R}^m$ dom$h = \mathbb{R}$. Then can derive similar conditions for convexity of $g \circ h$ for general m from

$$\nabla^2 f(x) = Dg(h(x))\nabla^2 h(x) + Dh(x)^T \nabla^2 g(h(x))Dh(x).$$

(Here, we understand $\nabla^2 h$ is $m \times 1$ matrix, $[f_1''(x), \dots, f_m''(x)]^T$.)
However, even without differentiability, we can observe the followings.

- g convex, $\tilde{g}$ nondecreasing in each argument, h_i convex $\Rightarrow$ f convex,
- g convex, $\tilde{g}$ nonincreasing in each argument, h_i concave $\Rightarrow$ f convex,
- g concave, $\tilde{g}$ nondecreasing in each argument, h_i concave $\Rightarrow$ f concave.

Composition(*cont'd*)

Example

- $g(z) = z_{[1]} + \cdots + z_{[r]}$, sum of r largest components of $z \in \mathbb{R}^m$. Then g is convex and nondecreasing in each z_i. Therefore, if $h_1, \ldots, h_m$ convex functions on $\mathbb{R}^n$, $f := g \circ h$ is convex.
- $g(z) = \log(\sum_{i=1}^m e^{z_i})$ is convex and nondecreasing in each z_i. Hence if h_i are convex, so is $g \circ h$.
- For $0 < p \leq 1$, $g(z) = (\sum_{i=1}^m z_i^p)^{1/p}$ is concave on $\mathbb{R}_+^m$ and its extension is nondecreasing in each z_i. Hence if h_i are concave and nonnegative $g \circ h$ is concave.
- For $p \geq 1$, if h_i are convex and nonnegative, $(\sum_{i=1}^m h_i(x)^p)^{1/p}$ is convex.
- $g(z) = (\prod_{i=1}^m z_i)^{1/m}$ on $\mathbb{R}_+^m$ is concave and its extension is nondecreasing in each z_i. If h_i are nonnegative concave function, so is $(\prod_{i=1}^m h_i)^{1/m}$.

Perspective of a function

If $f : \mathbb{R}^n \to \mathbb{R}$, then the *perspective* of f is the function $g : \mathbb{R}^{n+1} \to \mathbb{R}$ defined by

$$g(x, t) = tf(x/t),$$

with domain

$$\mathrm{dom} g = \{(x, t) | x/t \in \mathrm{dom} f, t > 0\}$$

Proposition

If f is convex (concave, resp.), so is its perspective.

Proof: $\square$

Perspective of a function(*cont'd*)

Example

$g(x,t) = \frac{x^T x}{t}$ is convex on $t > 0$.

Example

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is convex, then is

$$g(x) = (c^T x + d) f(Ax + b)/(c^T x + d),$$

with $\mathrm{dom} g = \{x | c^T x + d > 0,\ Ax + b)/(c^T x + d) \in \mathrm{dom} f\}$.

Definition

A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is called quasiconvex if its domain and sublevel sets

$$S_\alpha = \{x \in \mathrm{dom} f | f(x) \leq \alpha\}$$

are convex $\forall\ \alpha \in \mathbb{R}$.

- A function is *quasiconcave* if $-f$ is quasiconvex.
- A function that is both quasiconvex and quasiconcave is called *quasilinear*.

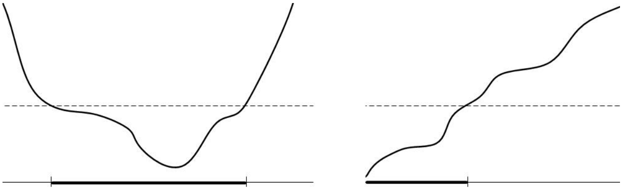


Figure: Quasiconcave and quasilinear function

Examples

- Logarithm $\log x$ is quasiconvex, quasiconcave, and hence quasilinear.
- Ceiling $\text{ceil}(x) = \min\{z \in \mathbb{Z} | x \geq z\}$ is quasilinear.
- Length of a vector x, $\max\{i | x_i \neq 0\}$ is quasiconvex.
- $f(x_1, x_2) = x_1 x_2$ on $\mathbb{R}_+^2$ is quasiconcave.
- $f(x) = \frac{a^T x + b}{c^T x + d}$ on $\{x | c^T x + d > 0\}$ is quasiconvex, quasiconcave and hence quasilinear.
- Distance ratio $f(x) = \frac{\|x - a\|_2}{\|x - b\|_2}$ is quaisconvex on halfspace $\|x - a\|_2 \leq \|x - b\|_2$.

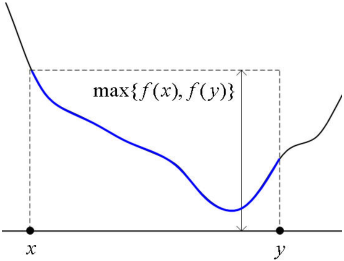
Basic properties

Proposition

A function f is quasiconvex if and only if dom$f$ is convex and for any $x, y \in$ dom$f$ and $0 \leq \lambda \leq 1$,

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

Proof □



Basic properties(*cont'd*)

- f is quasiconvex iff its restriction on line is quasiconvex.
- A continuous function $f : \mathbb{R} \to \mathbb{R}$ is quasiconvex iff one of the followings holds:
  - f is nondecreasing,
  - f is nonincreasing, or
  - $\exists\ c \in \text{dom} f$: f is nonincreasing on $x \leq c$, and nondecreasing on $x \geq c$.

First-order condition

Proposition

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 $x, y \in \operatorname{dom} f$

$$f(y) \leq f(x) \Rightarrow \nabla f(x)^T(y - x) \leq 0.$$

(Thus $\nabla f(x)$ defines supporting hyperplane of $\{y | f(y) \leq f(x)\}$.)

Proof Case 1: $f : \mathbb{R} \to \mathbb{R}$.
"If" Take any x, $y \in \operatorname{dom} f$ (assumed open) and $0 < \lambda < 1$. We need to show that $f\big((1-\lambda)x + \lambda y\big) \leq \max\{f(x), f(y)\}$. Assume $f(x) \geq f(y)$ and $f\big((1-\lambda)x + \lambda y\big) > f(x)$. Then there is $x < z < y$ such that $f(z) > f(x)$ and $f'(z) > 0$ and hence $(z - x)f'(z) > 0$. A contradiction. $\square$

Second-order condition

Supposer f is twice differentiable. If f is quasiconvex, then for all $x \in \operatorname{dom} f$, and all $y \in \mathbb{R}^n$, we have

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

When $\nabla f(x) \neq 0$, $\nabla^2 f(x) \succeq 0$ on $\nabla f(x)^\perp$, and hence may have at most 1 neg eigenvalue. As a (partial) converse, f is quasiconvex if f satisfies

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

- Nonnegative weighted maximum
  A nonnegative weighted maximum of quasiconvex functions

$$f = \max\{w_1 f_1, \ldots, w_m f_m\}$$

  with $w_i \geq 0$ and f_i quasiconvex, is quasiconvex.

- Composition
  - If $g : \mathbb{R}^n \to \mathbb{R}$ is quasiconvex and $h : \mathbb{R} \to \mathbb{R}$ is nondecreasing, then $f = h \circ g$ is quasiconvex.
  - Composition of quasiconvex function with affine or linear-fractional transform is quasiconvex: if f is quasiconvex, so are $f(Ax + b)$ and $f(\frac{Ax+b}{c^T x+d})$ on $\{x | c^T x + d > 0, \frac{Ax+b}{c^T x+d} \in \mathrm{dom} f\}$.

- Minimization. If $f(x, y)$ is quasiconvex jointly in x and y and C is a convex set, then the function

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

is quasiconvex.

Homework

3.1, 3.2, 3.3, 3.6, 3.7, 3.9, 3.17, 3.20, 3.22, 3.32, 3.43