Chapter 3 Convex functions

Last update on 2025-02-12 14:32:35+08:00

Table of contents

Properties and examples

Operations preserving convexity

Quasiconvex functions

Log-concave and log-convex functions

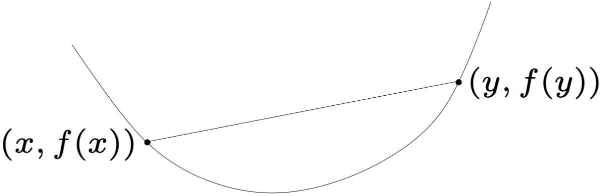
Convexity with respect to generalized inequalities

Convex function

- $f : \mathbb{R}^n \to \mathbb{R}$ is **convex** if $\mathbf{dom}\, f$ is a convex set and

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$$

for all $x, y \in \mathbf{dom}\, f$ and $0 \leq \theta \leq 1$



- $f : \mathbb{R}^n \to \mathbb{R}$ is **strictly convex** if $\mathbf{dom}\, f$ is a convex set and

$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta) f(y)$$

  for all $x, y \in \mathbf{dom}\, f$ with $x \neq y$ and $0 < \theta < 1$

- $f : \mathbb{R}^n \to \mathbb{R}$ is **concave** if $-f$ is convex

- $f : \mathbb{R}^n \to \mathbb{R}$ is **strictly concave** if $-f$ is strictly convex

Extended-value extension

∞-**extension** of a function $f\colon \mathbb{R}^n \to \mathbb{R}$ is

$$\tilde{f}\colon \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}; \qquad \mathbf{dom}\, \tilde{f} = \mathbb{R}^n$$

defined as

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

Extended-value extension

∞-**extension** of a function $f\colon \mathbb{R}^n \to \mathbb{R}$ is

$$\tilde{f}\colon \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}; \qquad \mathbf{dom}\, \tilde{f} = \mathbb{R}^n$$

defined as

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

lemma $\quad f\colon \mathbb{R}^n \to \mathbb{R}$ is convex $\iff$ for all $x, y \in \mathbb{R}^n$ and $0 < \theta < 1$

$$\tilde{f}(\theta x + (1-\theta)y) \le \theta \tilde{f}(x) + (1-\theta)\tilde{f}(y)$$

as an inequality in $\mathbb{R} \cup \{\infty\}$

Extended-value extension

∞-**extension** of a function $f : \mathbb{R}^n \to \mathbb{R}$ is

$$\tilde{f} : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}; \qquad \mathbf{dom}\, \tilde{f} = \mathbb{R}^n$$

defined as

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

lemma $\quad f : \mathbb{R}^n \to \mathbb{R}$ is convex $\iff$ for all $x, y \in \mathbb{R}^n$ and $0 < \theta < 1$

$$\tilde{f}(\theta x + (1-\theta)y) \leq \theta \tilde{f}(x) + (1-\theta)\tilde{f}(y)$$

as an inequality in $\mathbb{R} \cup \{\infty\}$

remark $\quad$ we can similarly define $(-\infty)$-extension of a function

Elementary techniques for establishing convexity

- definition
- restriction to lines
- first-order condition
- second-order condition

More advanced methods will be discussed in next section.

Restriction to a line

$f\colon \mathbb{R}^n \to \mathbb{R}$ is convex $\quad\Longleftrightarrow\quad$ the function $g\colon \mathbb{R} \to \mathbb{R}$ defined as

$$g(t) = f(x + tv), \qquad \mathbf{dom}\, g = \{t \mid x + tv \in \mathbf{dom}\, f\}$$

is convex in t for all $x \in \mathbf{dom}\, f$ and $v \in \mathbb{R}^n$

Useful upshot: we can check convexity of f by checking convexity of functions in one variable. Geometrically, it allows us to check whether a function is convex by restricting it to a line

Differentiability

- f is **differentiable** if $\mathbf{dom}\, f$ is open and the gradient

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \cdots, \frac{\partial f(x)}{\partial x_n} \right)$$

exists at each $x \in \mathbf{dom}\, f$

Differentiability

- f is **differentiable** if $\mathbf{dom}\, f$ is open and the gradient

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \cdots, \frac{\partial f(x)}{\partial x_n} \right)$$

  exists at each $x \in \mathbf{dom}\, f$

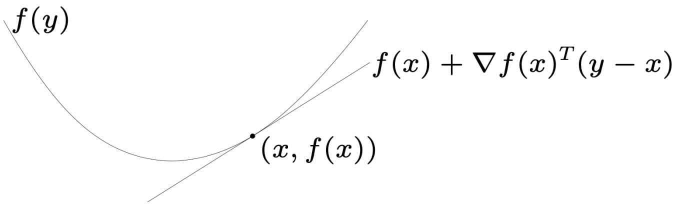
- f is **twice differentiable** if $\mathbf{dom}\, f$ is open and the Hessian

$$\nabla^2 f(x) = \left[\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right]_{1 \le i,j \le n}$$

  exists at each $x \in \mathbf{dom}\, f$

First-order condition

Suppose f is differentiable, then

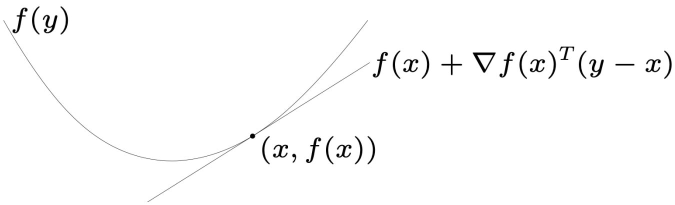


- f is convex $\iff$ $\mathbf{dom}\, f$ is convex and

$$f(y) \geq f(x) + \nabla f(x)^T(y - x) \qquad \text{for all } x, y \in \mathbf{dom}\, f$$

First-order condition

Suppose f is differentiable, then



- f is convex $\iff$ $\mathbf{dom}\, f$ is convex and

$$f(y) \geq f(x) + \nabla f(x)^T(y-x) \qquad \text{for all } x, y \in \mathbf{dom}\, f$$

- f is strictly convex $\iff$ $\mathbf{dom}\, f$ is convex and

$$f(y) > f(x) + \nabla f(x)^T(y-x) \qquad \text{for all } x, y \in \mathbf{dom}\, f \text{ and } x \neq y$$

- Proof of first-order convexity condition

proof of first/second-order condition

step 1. Establish the condition for $n = 1$ (standard calculus)

step 2. Prove the general case by restriction to lines

Second-order condition

Suppose f is twice differentiable, then

- f is convex $\iff$ $\mathbf{dom}\, f$ is convex and

$$\nabla^2 f(x) \succeq 0 \qquad \text{for all } x \in \mathbf{dom}\, f$$

Second-order condition

Suppose f is twice differentiable, then

- f is convex $\iff$ $\mathbf{dom}\, f$ is convex and

$$\nabla^2 f(x) \succeq 0 \qquad \text{for all } x \in \mathbf{dom}\, f$$

- f is strictly convex $\Longleftarrow$ $\mathbf{dom}\, f$ is convex and

$$\nabla^2 f(x) \succ 0 \qquad \text{for all } x \in \mathbf{dom}\, f$$

Examples

affine functions

- $f: \mathbb{R}^n \to \mathbb{R}$; $f(x) = a^T x + b$ where $a \in \mathbb{R}^n,\ b \in \mathbb{R}$
- $f: \mathbb{R}^{m \times n} \to \mathbb{R}$; $f(X) = \mathbf{tr}(A^T X) + b$ where $A \in \mathbb{R}^{m \times n},\ b \in \mathbb{R}$

affine functions are both convex and concave

powers of absolute values

- $|x|^p$, for $p \geq 1$, convex on $\mathbb{R}$

max functions

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

norms

$$\|\cdot\|: \qquad \mathbb{R}^n \longrightarrow \mathbb{R}$$

norms are convex functions (e.g. ℓ_p, Frobenius, spectral, nuclear, ...)

norms

$$\|\cdot\|: \qquad \mathbb{R}^n \longrightarrow \mathbb{R}$$

norms are convex functions (e.g. ℓ_p, Frobenius, spectral, nuclear, ...)

proof

- the domain
$$\mathbf{dom}\,(\|\cdot\|) = \mathbb{R}^n$$
is convex;
- for all $x, y \in \mathbb{R}^n$ and $\theta \in [0, 1]$
$$\|\theta x + (1-\theta)y\| \leq \|\theta x\| + \|(1-\theta)y\| = \theta\|x\| + (1-\theta)\|y\|$$

log-determinant

$$f : \mathbb{S}^n \to \mathbb{R}; \qquad f(X) = \log\det X; \qquad \mathbf{dom}\, f = \mathbb{S}^n_{++}$$

is concave

log-determinant

$$f : \mathbb{S}^n \to \mathbf{R}; \qquad f(X) = \log\det X; \qquad \mathbf{dom}\, f = \mathbb{S}^n_{++}$$

is concave

proof for every $X \in \mathbb{S}^n_{++}$ and every $V \in \mathbb{S}^n$

$$\begin{aligned} g(t) &= \log\det(X + tV) \\ &= \log\det X + \log\det(I + tX^{-1/2}VX^{-1/2}) \\ &= \log\det X + \sum_{i=1}^{n} \log(1 + t\lambda_i) \end{aligned}$$

where λ_i's are eigenvalues of $X^{-1/2}VX^{-1/2}$

$g(t)$ is concave for every choice of X and V, hence f is concave

quadratic function: $f(x) = (1/2)x^T P x + q^T x + r$ with $P \in \mathbb{S}^n$

quadratic function: $f(x) = (1/2)x^T Px + q^T x + r$ with $P \in \mathbb{S}^n$

$$\mathbf{dom}\, f = \mathbb{R}^n, \qquad \nabla f(x) = Px + q, \qquad \nabla^2 f(x) = P$$

convex iff $P \succeq 0$

quadratic function: $f(x) = (1/2)x^T Px + q^T x + r$ with $P \in \mathbb{S}^n$

$$\mathbf{dom}\, f = \mathbb{R}^n, \qquad \nabla f(x) = Px + q, \qquad \nabla^2 f(x) = P$$

convex iff $P \succeq 0$

least-square objective: $f(x) = \|Ax - b\|_2^2$

quadratic function: $f(x) = (1/2)x^T Px + q^T x + r$ with $P \in \mathbb{S}^n$

$$\mathbf{dom}\, f = \mathbb{R}^n, \qquad \nabla f(x) = Px + q, \qquad \nabla^2 f(x) = P$$

convex iff $P \succeq 0$

least-square objective: $f(x) = \|Ax - b\|_2^2$

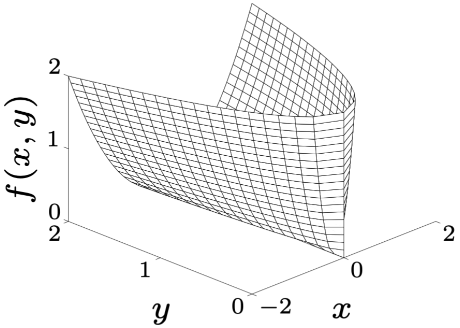
$$\mathbf{dom}\, f = \mathbb{R}^n, \qquad \nabla f(x) = 2A^T(Ax - b), \qquad \nabla^2 f(x) = 2A^T A \succeq 0$$

convex for any A and b

quadratic-over-linear: $f(x, y) = x^2/y$, $\mathbf{dom}\, f = \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$ is convex

quadratic-over-linear: $f(x,y) = x^2/y$, $\quad \mathbf{dom}\, f = \{(x,y) \in \mathbb{R}^2 \mid y > 0\}$ is convex

$$\nabla^2 f(x,y) = \frac{2}{y^3} \begin{bmatrix} y \\ -x \end{bmatrix} \begin{bmatrix} y \\ -x \end{bmatrix}^T \succeq 0$$



log-sum-exp: $f(x) = \log\left(\sum_{k=1}^{n} e^{x_k}\right)$ is convex

log-sum-exp: $f(x) = \log \left(\sum_{k=1}^{n} e^{x_k} \right)$ is convex

$\mathbf{dom}\, f = \mathbb{R}^n$; for convenience let $z_k = e^{x_k}$; and let $z = (z_1, \ldots, z_n)$

$$\nabla^2 f(x) = \cdots = \frac{1}{\sum z_k} \begin{bmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{bmatrix} - \frac{z z^T}{\left(\sum z_k \right)^2}$$

log-sum-exp: $f(x) = \log \left(\sum_{k=1}^{n} e^{x_k} \right)$ is convex

$\mathbf{dom}\, f = \mathbb{R}^n$; for convenience let $z_k = e^{x_k}$; and let $z = (z_1, \ldots, z_n)$

$$\nabla^2 f(x) = \cdots = \frac{1}{\sum z_k} \begin{bmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{bmatrix} - \frac{zz^T}{\left(\sum z_k\right)^2}$$

for every $v \in \mathbb{R}^n$

$$v^T \nabla^2 f(x) v = \frac{\left(\sum z_k v_k^2\right)\left(\sum z_k\right) - \left(\sum z_k v_k\right)^2}{\left(\sum z_k\right)^2} \geq 0$$

by Cauchy inequality, hence $\nabla^2 f(x) \succeq 0$ for every $x \in \mathbb{R}^n$

geometric mean: $$f(x) = \left(\prod_{k=1}^{n} x_k\right)^{\frac{1}{n}}$$ concave on $\mathbb{R}^n_{++}$

geometric mean: $$f(x) = \left(\prod_{k=1}^{n} x_k\right)^{\frac{1}{n}}$$ concave on $\mathbb{R}^n_{++}$

proof is similar to that of log-sum-exp

Properties of convex functions

- sublevel sets
- epigraphs
- Jensen's inequality

Sublevel set

α-**sublevel set** of $f\colon \mathbb{R}^n \to \mathbb{R}$

$$S_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$$

fact: f is convex $\implies$ all sublevel sets of f are convex (converse is false)

similar definition for **superlevel set**: f is concave $\implies$ all superlevel sets of f are convex

The geometric and arithmetic means of $x \in \mathbb{R}^n_+$ are, respectively,

$$G(x) = \left(\prod_{k=1}^{n} x_k \right)^{\frac{1}{n}}, \quad A(x) = \frac{1}{n} \sum_{k=1}^{n} x_k.$$

For $0 \leq \alpha \leq 1$, the set

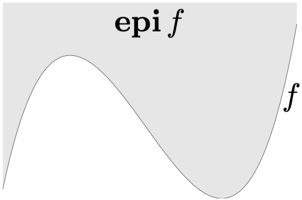
$$\{x \in \mathbb{R}^n_+ | G(x) \geq \alpha A(x)\},$$

is convex, since it is the 0-superlevel set of the concave function $G(x) - \alpha A(x)$

Epigraph

epigraph of $f\colon \mathbb{R}^n \to \mathbb{R}$

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



fact: f is convex $\iff$ $\mathbf{epi}\, f$ is a convex set

Jensen's inequality

basic version

if f is convex, then for $x, y \in \mathbf{dom}\, f$, $0 \leq \theta \leq 1$

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$$

if f is convex, then for $x_1, \cdots, x_k \in \mathbf{dom}\, f$, $\theta_1, \ldots, \theta_k \geq 0$ with $\theta_1 + \cdots + \theta_k = 1$

$$f(\theta_1 x_1 + \cdots + \theta_k x_k) \leq \theta_1 f(x_1) + \cdots + \theta_k f(x_k)$$

fancy version

if f is convex, then for $p(x) \geq 0$ on $S \subseteq \mathbf{dom}\, f$ with $\int_S p(x)\,\mathrm{d}\,x = 1$

$$f\left(\int_S xp(x)\,\mathrm{d}\,x\right) \leq \int_S f(x)p(x)\,\mathrm{d}\,x$$

in other words, for any random variable x taking values in $\mathbf{dom}\, f$

$$f(\mathbf{E}x) \leq \mathbf{E}f(x)$$

fancy version

if f is convex, then for $p(x) \geq 0$ on $S \subseteq \mathbf{dom}\, f$ with $\int_S p(x)\,\mathrm{d}\,x = 1$

$$f\left(\int_S x p(x)\,\mathrm{d}\,x\right) \leq \int_S f(x) p(x)\,\mathrm{d}\,x$$

in other words, for any random variable x taking values in $\mathbf{dom}\, f$

$$f(\mathbf{E}x) \leq \mathbf{E}f(x)$$

the above basic multi-point version is special case with discrete distribution

$$\mathbf{prob}(x_i) = \theta_i, \qquad i = 1, \cdots, k$$

Example: log-normal random variable

- suppose $X \sim N(\mu, \sigma^2)$
- with $f(u) = \exp u, Y = f(X)$ is log-normal
- we have $\mathbb{E}f(X) = \exp(\mu + \sigma^2/2)$
- Jensen's inequality is

$$f(\mathbb{E}X) = \exp(\mu) \leq \mathbb{E}f(X) = \exp(\mu + \sigma^2/2)$$

which indeed holds since $\exp(\sigma^2/2) > 1$

practical methods for establishing convexity of a function

1. definition; restriction to lines
2. first/second order conditions
3. reconstruct f from simple convex functions by operations preserving convexity
    - nonnegative weighted sum
    - composition with affine function
    - pointwise maximum and supremum
    - partial minimization
    - perspective
    - composition

Nonnegative weighted sum & composition with affine function

nonnegative weighted sum

$$f_1, f_2 \text{ are convex}, \quad \alpha_1, \alpha_2 \geq 0 \qquad \Longrightarrow \qquad \alpha_1 f_1 + \alpha_2 f_2 \text{ is convex}$$

extends to finite and infinite sums, integrals

composition with affine function

$$f \text{ is convex} \qquad \Longrightarrow \qquad f(Ax + b) \text{ is convex}$$

examples

- log barrier for linear inequalities

$$f(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x), \qquad \mathbf{dom}\, f = \{x \mid a_i^T x < b_i, i = 1, \cdots, m\}$$

- any norm of affine function

$$f(x) = \|Ax + b\|$$

Pointwise maximum

$f_1, \cdots, f_m$ are convex $\Longrightarrow$ $f(x) = \mathbf{max}\{f_1(x), \cdots, f_m(x)\}$ is convex

Pointwise maximum

$$f_1, \cdots, f_m \text{ are convex} \quad \Longrightarrow \quad f(x) = \mathbf{max}\{f_1(x), \cdots, f_m(x)\} \text{ is convex}$$

proof on page 80

examples

- piecewise-linear function

$$f(x) = \mathbf{max}\{a_i^T x + b_i \mid 1 \leq i \leq m\}$$

- sum of r largest components of $x \in \mathbb{R}^n$

$$f(x) = x_{[1]} + \cdots + x_{[r]}$$

Pointwise maximum

$$f_1, \cdots, f_m \text{ are convex} \quad \Longrightarrow \quad f(x) = \mathbf{max}\{f_1(x), \cdots, f_m(x)\} \text{ is convex}$$

proof on page 80

examples

- piecewise-linear function

$$f(x) = \mathbf{max}\{a_i^T x + b_i \mid 1 \leq i \leq m\}$$

- sum of r largest components of $x \in \mathbb{R}^n$

$$f(x) = x_{[1]} + \cdots + x_{[r]}$$

proof

$$f(x) = \mathbf{max}\{x_{i_1} + \cdots + x_{i_r} \mid 1 \leq i_1 < \cdots < i_r \leq n\}$$

Pointwise supremum

$$f(x, \lambda) \text{ is convex in } x \text{ for each } \lambda \in \Lambda \qquad \Longrightarrow \qquad g(x) = \sup_{\lambda \in \Lambda} f(x, \lambda) \text{ is convex}$$

Pointwise supremum

$f(x, \lambda)$ is convex in x for each $\lambda \in \Lambda \quad \Longrightarrow \quad g(x) = \sup_{\lambda \in \Lambda} f(x, \lambda)$ is convex

examples

- distance to farthest point in a set C

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

- maximum eigenvalue of symmetric matrices

$$\lambda_{\max}(X)$$

Pointwise supremum

$f(x, \lambda)$ is convex in x for each $\lambda \in \Lambda \quad \Longrightarrow \quad g(x) = \sup_{\lambda \in \Lambda} f(x, \lambda)$ is convex

examples

- distance to farthest point in a set C

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

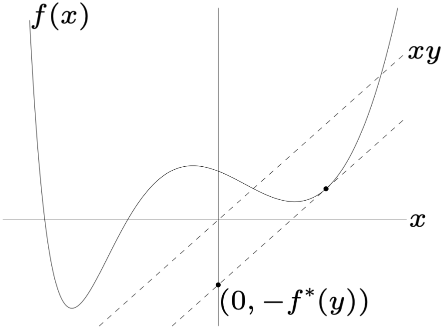
- maximum eigenvalue of symmetric matrices

$$\lambda_{\max}(X) = \sup_{\|y\|_2 = 1} y^T X y$$

The conjugate function

conjugate of any function f is

$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} (y^T x - f(x))$$



examples

- for $f(x) = -\log x$

$$
\begin{aligned}
f^*(y) &= \mathbf{sup}_{x>0}(xy + \log x) \\
&= \begin{cases} -1 - \log(-y), & y < 0 \\ \infty, & y \geq 0 \end{cases}
\end{aligned}
$$

- for $f(x) = (1/2)x^T Qx$ with $Q \in \mathbb{S}^n_{++}$

$$
\begin{aligned}
f^*(y) &= \mathbf{sup}_x(y^T x - (1/2)x^T Qx) \\
&= (1/2)y^T Q^{-1} y
\end{aligned}
$$

property

- $$f \text{ is convex and closed } \ (\mathbf{epi}\, f \text{ is closed}) \qquad \Longrightarrow \qquad f^{**} = f$$
- The conjugate of a differentiable function f is called the Legendre transform of f. We have
$$f^*(y) = x^{*T} \nabla f(x^*) - f(x^*),$$
where x^* is any maximizer of $y^T x - f(x)$ satisfying $y = \nabla f(x^*)$
- The conjugate of $g(x) = af(x) + b$ is $g^*(y) = af^*(y/a) - b$
- sums of independent functions. If $f(u,v) = f_1(u) + f_2(v)$ with f_1 and f_2 convex, then $f^*(w,z) = f_1^*(w) + f_2^*(z)$.

Minimization

$f(x, y)$ is convex in (x, y) and C is a convex set $\implies$ $g(x) = \inf_{y \in C} f(x, y)$ is convex

Minimization

$$f(x, y) \text{ is convex in } (x, y) \text{ and } C \text{ is a convex set} \quad \Longrightarrow \quad g(x) = \inf_{y \in C} f(x, y) \text{ is convex}$$

proof on page 88

examples

- if C is a convex set, then

$$\mathbf{dist}(x, C) = \inf_{y \in C} \|x - y\|$$

  is convex

Perspective

perspective of a function $f\colon \mathbb{R}^n \to \mathbb{R}$ is the function $g\colon \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$

$$g(x,t) = tf(x/t), \qquad \mathbf{dom}\, g = \{(x,t) \mid x/t \in \mathbf{dom}\, f, t > 0\}$$

$$f \text{ is convex} \qquad \Longrightarrow \qquad g \text{ is convex}$$

Perspective

perspective of a function $f\colon \mathbb{R}^n \to \mathbb{R}$ is the function $g\colon \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$

$$g(x,t) = tf(x/t), \qquad \mathbf{dom}\, g = \{(x,t) \mid x/t \in \mathbf{dom}\, f, t > 0\}$$

$$f \text{ is convex} \qquad \Longrightarrow \qquad g \text{ is convex}$$

proof on page 89

examples

- $f(x) = x^T x$ is convex, hence

$$g(x,t) = x^T x/t$$

  is convex for $t > 0$

- $f(x) = -\log x$ is convex, hence

$$g(x,t) = t\log t - t\log x$$

is convex on $\mathbb{R}^2_{++}$ (relative entropy)

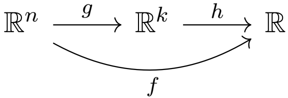
- if f is convex, then

$$g(x) = (c^T x + d) \cdot f\left(\frac{Ax+b}{c^T x + d}\right)$$

is convex on

$$\left\{ x \;\middle|\; c^T x + d > 0,\ \frac{Ax+b}{c^T x + d} \in \mathbf{dom}\, f \right\}$$

Composition



$$f(x) = h(g(x)) = h(g_1(x), \cdots, g_k(x))$$

where

$$\mathbf{dom}\, g = \bigcap_{i=1}^{k} \mathbf{dom}\, g_i$$

$$\mathbf{dom}\, f = \{x \in \mathbf{dom}\, g \mid g(x) \in \mathbf{dom}\, h\}$$

proposition assume $f(x) = h(g(x)) = h(g_1(x), \cdots, g_k(x))$

h convex, $\tilde{h}$ nondecreasing in each argument, each g_i convex $\quad \Longrightarrow \quad f$ convex

proposition assume $f(x) = h(g(x)) = h(g_1(x), \cdots, g_k(x))$

$$h \text{ convex}, \ \tilde{h} \text{ nondecreasing in each argument}, \ \text{each } g_i \text{ convex} \quad \Longrightarrow \quad f \text{ convex}$$

corollary assume $f(x) = h(g(x)) = h(g_1(x), \cdots, g_k(x))$

h	$\tilde{h}$	each g_i		$f = h \circ g$
convex	$\nearrow$	convex		convex
convex	$\searrow$	concave	$\Longrightarrow$	convex
concave	$\searrow$	convex		concave
concave	$\nearrow$	concave		concave

remark

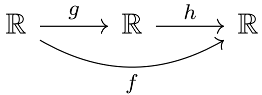
- $\tilde{h}$ is ∞-extension if h is convex and $(-\infty)$-extension if h is concave;
- the monotonicity of $\tilde{h}$ is non-strict and in each argument.

warning monotonicity of $\tilde{h}$ has to hold for $\tilde{h}$ instead of h

counterexample

- $g(x) = x^2$ is convex
- $h(x) = 0$ with $\mathbf{dom}\, h = [1, 2]$ is convex
- $h(x)$ is non-decreasing, but $\tilde{h}$ is not
- $f(x) = h(g(x)) = 0$ with $\mathbf{dom}\, f = [-\sqrt{2}, -1] \cup [1, \sqrt{2}]$ is not convex

fake proof



assume all functions are twice differentiable

$$f(x) = h(g(x)) \qquad \Longrightarrow \qquad f''(x) = h''(g(x))g'(x)^2 + h'(g(x))g''(x)$$

then we have

$$h'' \geq 0,\ h' \geq 0,\ g'' \geq 0 \qquad \Longrightarrow \qquad f'' \geq 0$$

without worrying about domains, we conclude

$$h \text{ convex},\ \tilde{h} \text{ nondecreasing},\ g \text{ convex} \quad \Longrightarrow \quad f \text{ convex}$$

proof Assume $\theta \in [0,1]$ and $x, y \in \mathbf{dom}\, f$

step 1 Show $\mathbf{dom}\, f$ is convex

$$\text{each } g_i \text{ convex} \implies \begin{cases} \theta x + (1-\theta)y \in \mathbf{dom}\, g \\ g(\theta x + (1-\theta)y) \preceq \theta g(x) + (1-\theta)g(y) \end{cases}$$

$$h \text{ convex} \implies \theta g(x) + (1-\theta)g(y) \in \mathbf{dom}\, h$$

$$\tilde{h} \text{ non-decreasing} \implies \tilde{h}(g(\theta x + (1-\theta)y)) \leq \tilde{h}(\theta g(x) + (1-\theta)g(y)) < \infty$$

$$\implies g(\theta x + (1-\theta)y) \in \mathbf{dom}\, h$$

Therefore we conclude

$$\theta x + (1-\theta)y \in \mathbf{dom}\, f$$

step 2 Show $f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$

$$\begin{aligned} f(\theta x + (1-\theta)y) &= h(g(\theta x + (1-\theta)y)) \\ &\leq h(\theta g(x) + (1-\theta)g(y)) \\ &\leq \theta h(g(x)) + (1-\theta)h(g(x)) = \theta f(x) + (1-\theta)f(y) \end{aligned}$$

examples

- if $g(x)$ is concave and positive, then

$$1/g(x)$$

  is convex

- if all $g_i(x)$ are convex, then

$$\log\left(\sum_{i=1}^{m} e^{g_i(x)}\right)$$

  is convex

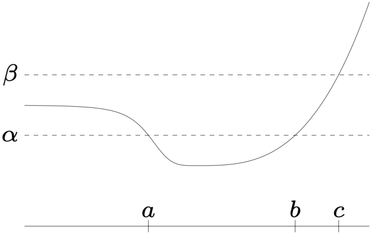
Quasiconvex functions

Quasiconvex function

- $f : \mathbb{R}^n \to \mathbb{R}$ is **quasiconvex** if $\mathbf{dom}\, f$ is convex and the sublevel sets

$$C_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \leq \alpha\}$$

  are convex for all $\alpha \in \mathbb{R}$



- f is quasiconcave if $-f$ is quasiconvex
- f is quasilinear if it is quasiconvex and quasiconcave

Examples

- $\sqrt{|x|}$ is quasiconvex on $\mathbb{R}$
- $\text{ceil}(x) = \mathbf{inf}\{z \in \mathbb{Z} \mid z \geq x\}$ is quasilinear on $\mathbb{R}$
- $\log x$ is quasilinear on $\mathbb{R}_{++}$
- **Note that quasiconvex functions can be concave, or discontinuous.**
- $f(x_1, x_2) = x_1 x_2$ is quasiconcave on $\mathbb{R}^2_{++}$
- linear-fractional function

$$f(x) = \frac{a^T x + b}{c^T x + d}, \qquad \mathbf{dom}\, f = \{x \mid c^T x + d > 0\}$$

  is quasilinear
- distance ratio

$$f(x) = \frac{\|x - a\|_2}{\|x - b\|_2}, \qquad \mathbf{dom}\, f = \{x \mid \|x - a\|_2 \leq \|x - b\|_2\}$$

  is quasiconvex

Techniques for establishing quasi-convexity

1. direct approaches
   - definition
   - modified Jensen inequality
   - restriction to lines
   - first order condition
   - second order condition
2. construct new from old

Properties

Modified Jensen inequality

f is quasi-convex iff $\mathbf{dom}\, f$ is convex and

$$f(\theta x + (1-\theta)y) \leq \mathbf{max}\{f(x), f(y)\}$$

for all $x, y \in \mathbf{dom}\, f$ and $\theta \in [0, 1]$.

Properties

Modified Jensen inequality

f is quasi-convex iff $\mathbf{dom}\, f$ is convex and

$$f(\theta x + (1-\theta)y) \leq \mathbf{max}\{f(x), f(y)\}$$

for all $x, y \in \mathbf{dom}\, f$ and $\theta \in [0, 1]$.

Warning

many properties of convex functions are false for quasiconvex functions;

e.g. sums of quasiconvex functions are not necessarily quasiconvex.

Log-concave and log-convex function

A positive function $f : \mathbb{R}^n \to \mathbb{R}$ is

- **log-concave** if $\log f$ is concave: $\mathbf{dom}\, f$ is convex and

$$f(\theta x + (1-\theta)y) \geq f(x)^{\theta} f(y)^{1-\theta} \qquad \text{for all } x, y \in \mathbf{dom}\, f \text{ and } 0 \leq \theta \leq 1$$

- **log-convex** of $\log f$ is convex: $\mathbf{dom}\, f$ is convex and

$$f(\theta x + (1-\theta)y) \leq f(x)^{\theta} f(y)^{1-\theta} \qquad \text{for all } x, y \in \mathbf{dom}\, f \text{ and } 0 \leq \theta \leq 1$$

Examples

- powers: x^a on $\mathbb{R}_{++}$ is log-convex for $a \leq 0$, log-concave for $a \geq 0$
- Gamma function is log-convex

$$\Gamma(x) = \int_0^\infty u^{x-1} e^{-u}\, \mathrm{d}u, \qquad \mathbf{dom}\, \Gamma = [1, \infty)$$

- many common probability density functions are log-concave, e.g. Gaussian

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp\left(-\frac{1}{2}(x-\bar{x})^T \Sigma^{-1} (x-\bar{x})\right)$$

- cumulative Gaussian distribution function is log-concave

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2}\, \mathrm{d}\, u$$

Operations preserving log-convexity

- If $f(x)$ and $g(x)$ are log-convex, then

$$f(x)g(x) \qquad \text{and} \qquad f(x) + g(x)$$

  are log-convex.

- If $f(x, \lambda)$ is log-convex in x for each $\lambda \in C$, then

$$g(x) = \int_C f(x, \lambda)\, \mathrm{d}\lambda$$

  is also log-convex.

- Laplace transform of a nonnegative function, Moment generating function

Operations preserving log-concavity

- If $f(x)$ and $g(x)$ are log-concave, then

$$f(x)g(x)$$

  is log-concave, but

$$f(x) + g(x)$$

  is not necessarily log-concave.

- **(Integration Theorem)** If $f\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is log-concave in (x, y), then

$$g(x) = \int f(x, y)\,\mathrm{d}y$$

  is log-concave. (Very useful, but difficult to prove.)

Applications of integration theorem

Convolution product If $f(x)$ and $g(x)$ are log-concave, then

$$(f * g)(x) = \int f(x-y)g(y)\,\mathrm{d}y$$

is log-concave.

Yield function $Y(x) = \mathbf{prob}(x + w \in S)$

- $x \in \mathbb{R}^n$: nominal (target) parameter values for product,
- $w \in \mathbb{R}^n$: random variations of parameters, with probability density function $p(w)$,
- $S \subseteq \mathbb{R}^n$: set of acceptable values.

Yield function $Y(x) = \textbf{prob}(x + w \in S)$

- $x \in \mathbb{R}^n$: nominal (target) parameter values for product,
- $w \in \mathbb{R}^n$: random variations of parameters, with probability density function $p(w)$,
- $S \subseteq \mathbb{R}^n$: set of acceptable values.

Assume

S is a convex set and $p(w)$ is log-concave

then

$Y(x)$ is log-concave and the yield regions $\{x \mid Y(x) \geq \alpha\}$ are convex

Yield function $\qquad Y(x) = \mathbf{prob}(x + w \in S)$

- $x \in \mathbb{R}^n$: nominal (target) parameter values for product,
- $w \in \mathbb{R}^n$: random variations of parameters, with probability density function $p(w)$,
- $S \subseteq \mathbb{R}^n$: set of acceptable values.

Assume

$$S \text{ is a convex set} \qquad \text{and} \qquad p(w) \text{ is log-concave}$$

then

$$Y(x) \text{ is log-concave} \quad \text{and} \quad \text{the yield regions } \{x \mid Y(x) \geq \alpha\} \text{ are convex}$$

Proof

$$Y(x) = \int g(x+w)p(w)\,\mathrm{d}w, \quad \text{where} \quad g(u) = \begin{cases} 1 & u \in S \\ 0 & u \notin S \end{cases} \quad \text{is log-concave.}$$

From density function to distribution function

Assume the density function $p(x)$ is log-concave, then the distribution function is

$$
\begin{aligned}
F(x) = \mathbf{prob}(w \preceq x) &= \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} p(w)\, \mathrm{d}w_1 \ldots \mathrm{d}w_n \\
&= \int_{\mathbb{R}^n} g(w - x) p(w)\, \mathrm{d}w
\end{aligned}
$$

where

$$
S = (-\infty, 0]^n \qquad \text{and} \qquad g(u) = \begin{cases} 1 & u \in S, \\ 0 & u \notin S. \end{cases}
$$

S is convex, $\quad p(x)$ is log-concave $\quad \Longrightarrow \quad F(x)$ is log-concave.

Properties and examples

Operations preserving convexity

Quasiconvex functions

Log-concave and log-convex functions

Convexity with respect to generalized inequalities

Monotonicity with respect to generalized inequalities

- Suppose K is a proper cone. A function $f : \mathbb{R}^n \to \mathbb{R}$ is called K-nondecreasing if

$$x \preceq_K y \Rightarrow f(x) \leq f(y).$$

- Matrix monotone functions. A function $f : \mathbb{S}^n \to \mathbb{R}$ is called matrix monotone if it is monotone wrt the positive semidefinite cone.
- $\mathbf{tr}(WX)$ is matrix nondecreasing if $W \succeq 0$
- $\mathbf{tr}(X^{-1})$ is matrix decreasing on $\mathbb{S}^n_{++}$
- $\det X$ is matrix increasing on $\mathbb{S}^n_{++}$
- A differentiable function f, with convex domain, is K-nondecreasing if and only if

$$\nabla f(x) \succeq_{K^*} 0$$

for all $x \in \mathbf{dom} f$

Convexity with respect to generalized inequalities

- let $K \subseteq \mathbb{R}^m$ be a proper cone with associated generalized inequality $\preceq_K$.
- $f: \mathbb{R}^n \to \mathbb{R}^m$ is K-**convex** if $\mathbf{dom}\, f$ is convex and

$$f(\theta x + (1-\theta)y) \preceq_K \theta f(x) + (1-\theta)f(y)$$

  for every $x, y \in \mathbf{dom}\, f$ and $0 \leq \theta \leq 1$.

Example

$$f \colon \mathbb{S}^m \to \mathbb{S}^m, \qquad f(X) = X^2 \qquad \text{is } \mathbb{S}^m_+\text{-convex}$$

Observe: $\quad \mathbf{dom}\, f = \mathbb{S}^m$ is convex.

Need to show: $\quad$ for any $X, Y \in \mathbb{S}^m$ and $\theta \in [0, 1]$, we have

$$\begin{aligned} & (\theta X + (1-\theta)Y)^2 \preceq_{\mathbb{S}^m_+} \theta X^2 + (1-\theta)Y^2 \\ \iff \quad & z^T(\theta X + (1-\theta)Y)^2 z \le \theta z^T X^2 z + (1-\theta) z^T Y^2 z \quad \text{for all } z \in \mathbb{R}^m \\ \iff \quad & \|\theta X z + (1-\theta) Y z\|_2^2 \le \theta \|Xz\|_2^2 + (1-\theta) \|Yz\|_2^2 \end{aligned}$$

which holds since $\| \cdot \|_2^2$ is convex on $\mathbb{R}^m$.

Example Convexity wrt componentwise inequality (page 110)

Example The function X^p is matrix convex on $\mathbb{S}^n_{++}$ for $1 \le p \le 2$ and matrix concave for $0 \le p \le 1$

Example The function $f(X) = e^X$ is not matrix convex on $\mathbb{S}^n$, for $n \ge 2$

Dual characterization of K-convexity

Differentiable K-convex functions

Composition theorem