Chapter 3 Convex functions

Last update on 2022-03-09 10:43

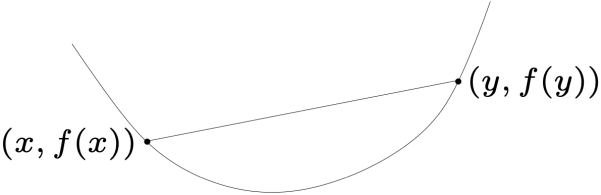
Table of contents

Convex function

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

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

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



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

Examples

convex/concave on $\mathbb{R}$

affine	$ax + b$	convex & concave on $\mathbb{R}$ for any $a, b \in \mathbb{R}$
exponential	e^{ax}	convex on $\mathbb{R}$ for any $a \in \mathbb{R}$
powers	x^{α}	convex on $\mathbb{R}_{++}$ for $\alpha \geq 1$ or $\alpha < 0$ concave on $\mathbb{R}_{++}$ for $0 \leq \alpha \leq 1$
powers of absolute value	$\lvert x \rvert^{p}$	convex on $\mathbb{R}$ for $p \geq 1$
logarithm	$\log x$	concave on $\mathbb{R}_{++}$
negative entropy	$x \log x$	convex on $\mathbb{R}_{++}$

convex on $\mathbb{R}^n$

- affine function

$$f(x) = a^T x + b \qquad \text{for any } a, b \in \mathbb{R}^n$$

- norms

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} \qquad \text{for } p \geq 1$$

$$\|x\|_\infty = \max_k |x_k|$$

convex on $\mathbb{R}^{m\times n}$ (for matrices)

- affine functions

$$f(X) = \mathbf{tr}(A^T X) + b$$

- spectral norm (maximal singlar value)

$$f(X) = \|X\|_2 = \sigma_{\max}(X) = \left(\lambda_{\max}(X^T X)\right)^{1/2}$$

Restriction of a convex function to a line

$f : \mathbb{R}^n \to \mathbb{R}$ is convex iff the function $g : \mathbb{R} \to \mathbb{R}$

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

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

upshot: we can check convexity of f by checking convexity of functions in one variable

Example

log-determinant

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

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

Extended-value extension

extended-value extension $\tilde{f}$ of a convex function f is

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

this often simplifies notation; for example

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

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

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

similarly, we can extend a concave function by defining it to be $-\infty$ outside its domain

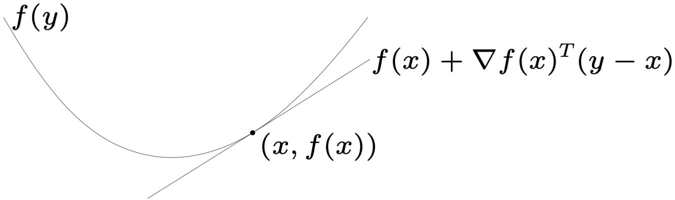
First-order condition

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 $x \in \mathbf{dom}\, f$

first-order condition: for differentiable function f



- 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

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

Second-order condition

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

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

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

second-order condition: for twice differentiable function f

- $\mathbf{dom}\, f$ is convex and $\nabla^2 f(x) \succeq 0$ for all $x \in \mathbf{dom}\, f \quad \Longleftrightarrow \quad f$ is convex
- $\mathbf{dom}\, f$ is convex and $\nabla^2 f(x) \succ 0$ for all $x \in \mathbf{dom}\, f \quad \Longrightarrow \quad f$ is strictly convex

Examples

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

$$\nabla f(x) = Px + q, \qquad \nabla^2 f(x) = P$$

convex iff $P \succeq 0$

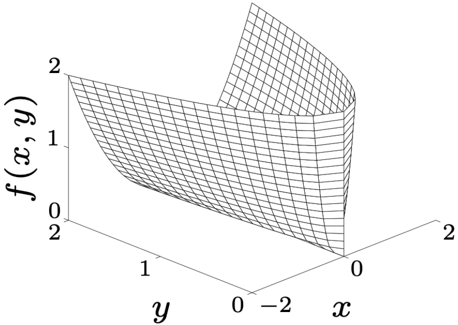
least-square objective: $f(x) = \|Ax - b\|_2^2$

$$\nabla f(x) = 2A^T(Ax - b), \qquad \nabla^2 f(x) = 2A^T A \succeq 0$$

convex for any A and b

quadratic-over-linear: $f(x,y) = x^2/y$, $\quad \mathbf{dom}\, f = \{(x,y) \in \mathbb{R}^2 \mid y > 0\}$ $\quad$ 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: $f(x) = \log\left(\sum_{k=1}^{n} e^{x_k}\right)$ convex

$$\nabla^2 f(x) = \frac{1}{\mathbf{1}^T z}\,\mathbf{diag}(z) - \frac{1}{(\mathbf{1}^T z)^2} zz^T \qquad \text{where } z_k = e^{x_k}$$

to verify $\nabla^2 f(x) \succeq 0$ we must show $v^T \nabla^2 f(x) v \geq 0$ for all $v \in \mathbb{R}^n$

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

since $(\sum_k z_k v_k)^2 \leq (\sum_k z_k v_k^2)(\sum_k z_k)$ by Cauchy inequality

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

Sublevel set

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

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

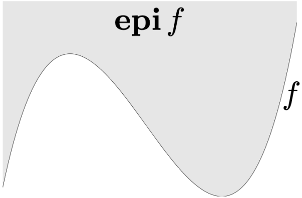
f is convex $\implies$ all sublevel sets of f are convex (converse is false)

similar definition for **superlevel set**

Epigraph

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

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



f is convex $\iff$ **epi** f is a convex set

similar definition for **hypograph**

Jensen's inequality

basic version

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

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

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

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

extended 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 $x \in \mathbf{dom}\, f$

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

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

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

practical methods for establishing convexity of a function

1. verify definition (often by restriction to lines)
2. for twice differentiable functions, show $\nabla^2 f(x) \succeq 0$
3. reconstruct f from simple convex functions by operations preserving convexity

   - nonnegative weighted sum
   - composition with affine function
   - pointwise maximum and supremum
   - composition
   - minimization
   - perspective

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

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

examples

- piecewise-linear function

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

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

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

proof

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

Pointwise supremum

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

examples

- support function of a set C

$$S_C(x) = \sup_{y \in C} y^T x$$

- distance to farthest point in a set C

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

- maximum eigenvalue of symmetric matrices

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

Composition

$$f(x) = h(g(x))$$

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "g"] \arrow[rr, bend right, "f"'] & \mathbb{R}^k \arrow[r, "h"] & \mathbb{R} \end{tikzcd}$$

more precisely

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

where

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

scalar composition: for $k = 1$

$$
\begin{array}{lllll}
g \text{ is convex,} & h \text{ is convex,} & \tilde{h} \text{ is nondecreasing} & \Longrightarrow & f \text{ is convex} \\
g \text{ is concave,} & h \text{ is convex,} & \tilde{h} \text{ is nonincreasing} & \Longrightarrow & f \text{ is convex}
\end{array}
$$

warning: monotonicity must hold for extended-value extension $\tilde{h}$

proof (for $n = 1$, differentiable g and h)

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

examples

- if $g(x)$ is convex, then $e^{g(x)}$ is convex
- if $g(x)$ is concave and positive, then $1/g(x)$ is convex

vector composition: for general $k \geq 1$

$$
\begin{array}{lcr}
g_i \text{ convex, } h \text{ convex, } \tilde{h} \text{ nondecreasing in each argument} & \Longrightarrow & f \text{ is convex} \\
g_i \text{ concave, } h \text{ convex, } \tilde{h} \text{ nonincreasing in each argument} & \Longrightarrow & f \text{ is convex}
\end{array}
$$

warning: monotonicity must hold for extended-value extension $\tilde{h}$

proof (for $n = 1$, differentiable g and h)

$$f''(x) = g'(x)^T \nabla^2 h(g(x)) g'(x) + \nabla h(g(x))^T g''(x)$$

examples

- if all $g_i(x)$ are concave and positive, then $\sum_{i=1}^{m} \log g_i(x)$ is concave
- if all $g_i(x)$ are convex, then $\log \left(\sum_{i=1}^{m} e^{g_i(x)} \right)$ is convex

Minimization

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

examples

- if S is a convex set, then

$$\textbf{dist}(x, S) = \inf_{y \in S} \|x - y\|$$

  is convex

- assume symmetric matrices A and C satisfy

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succeq 0, \qquad C \succ 0$$

then

$$f(x, y) = x^T A x + 2x^T B y + y^T C y$$

is convex, hence

$$g(x) = \inf_y f(x, y) = x^T (A - BC^{-1}B^T) x$$

is also convex, which implies

$$A - BC^{-1}B^T \succeq 0 \qquad \text{(Schur complement of } C\text{)}$$

Perspective

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

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

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

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

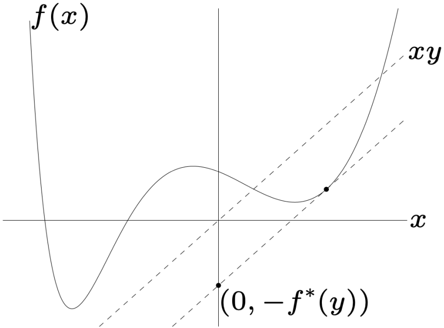
$$g(x) = (c^T x + d) \cdot f\left(\frac{Ax + b}{c^T x + d}\right)$$

is convex on

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

conjugate of a function f is

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



then f^* is convex (even if f is not)

examples

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

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

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

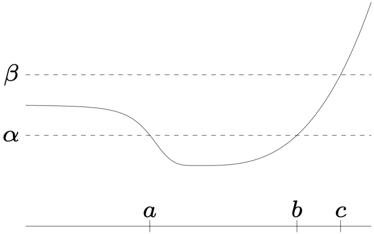
$$f^*(y) = \sup_x (y^T x - (1/2)x^T Qx)$$
$$= (1/2)y^T Q^{-1} y$$

Quasiconvex function

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

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

are convex for all α



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

Examples

- $\sqrt{|x|}$ is quasiconvex on $\mathbb{R}$
- $\text{ceil}(x) = \inf\{z \in \mathbb{Z} \mid z \geq x\}$ is quasilinear on $\mathbb{R}$
- $\log x$ is quasilinear on $\mathbb{R}_{++}$
- $f(x_1, x_2) = x_1 x_2$ is quasiconcave on $\mathbb{R}^2_{++}$
- linear-fractional function

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

  is quasilinear
- distance ratio

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

  is quasiconvex

Internal rate of return

- cash flow sequence $x = (x_0, \cdots, x_n)$, where x_i is payment in period i (to us if $x_i > 0$, from us if $x_i < 0$)
- assume $x_0 < 0$ and $x_0 + \cdots + x_n > 0$
- present value of cash flow x, for interest rate r

$$\mathrm{PV}(x, r) = \sum_{i=0}^{n} (1+r)^{-i} x_i$$

- clearly $\mathrm{PV}(x, 0) > 0$ and $\mathrm{PV}(x, r) \to x_0 < 0$ as $r \to \infty$
- **internal rate of return** is smallest interest rate for which $\mathrm{PV}(x, r) = 0$

$$\mathrm{IRR}(x) = \inf\{r \geq 0 \mid \mathrm{PV}(x, r) = 0\}$$

IRR is quasiconcave: superlevel set is intersection of open halfspaces

$$\mathrm{IRR}(x) \geq R \quad \Longleftrightarrow \quad \sum_{i=0}^{n} (1+r)^{-i} x_i > 0 \text{ for } 0 \leq r < R$$

Properties

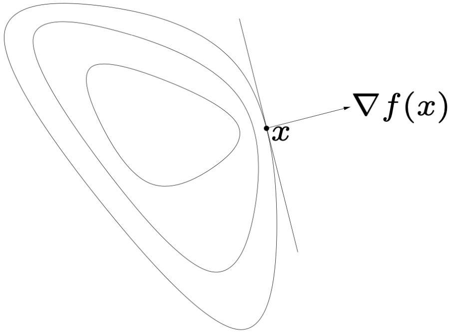
warning: many properties of convex functions are false for quasiconvex functions; e.g. sums of quasiconvex functions are not necessarily quasiconvex

modified Jensen inequality: for quasiconvex f

$$0 \leq \theta \leq 1 \quad \Longrightarrow \quad f(\theta x + (1-\theta)y) \leq \max\{f(x), f(y)\}$$

first-order condition: differentiable function f is quasiconvex if and only if

1. **dom** f is convex, and
2. for all $x, y \in \textbf{dom}\, f$, $\quad f(y) \leq f(x) \implies \nabla f(x)^T(y - x) \leq 0$



Log-concave and log-convex function

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

- **log-concave** if $\log f$ is concave: **dom** f is convex and

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

- **log-convex** of $\log f$ is convex: **dom** f is convex and

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

Examples

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

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

Properties

second-order condition: twice differentiable f is

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

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

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

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

product of log-concave/log-convex functions is log-concave/log-convex

sum of log-convex functions is log-convex; false for log-concave functions

Integration of log-concave functions

if $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is log-concave, then

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

is log-concave (not easy to show)

example

convolution of log-concave functions f and g

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

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 in manufactured product
- $S \subseteq \mathbb{R}^n$: set of acceptable values

assume S is convex and w has log-concave probability density $p(w)$ then

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

proof

$$Y(x) = \int g(x+w)p(w)\, d\, w, \qquad g(u) = \begin{cases} 1 & u \in S \\ 0 & u \notin S \end{cases}$$

Convexity with respect to generalized inequalities

$K \subseteq \mathbb{R}^n$ proper cone with associated generalized inequality $\preceq_K$

$f : \mathbb{R}^n \to \mathbb{R}^m$ is K-**convex** if $\textbf{dom}\, f$ is convex and

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

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

example

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

proof

for any $z \in \mathbb{R}^m$, $z^T X^2 z = \|Xz\|_2^2$ is convex in X

$$z^T(\theta X + (1-\theta)Y)^2 z \leq \theta z^T X^2 z + (1-\theta) z^T Y^2 z$$

for $X, Y \in \mathbb{S}^m$ and $0 \leq \theta \leq 1$, therefore

$$(\theta X + (1-\theta)Y)^2 \preceq_{\mathbb{S}^m_+} \theta X^2 + (1-\theta)Y^2$$