Chapter 3 Convex functions

Last update on 2024-03-19 09:43

Table of contents

Properties and examples

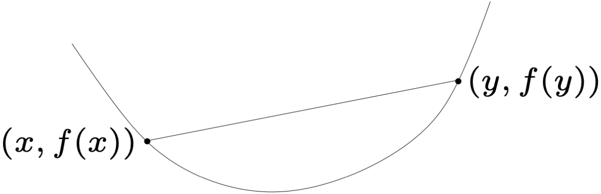
Operations preserving convexity

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

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

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

remark 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

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

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

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

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

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

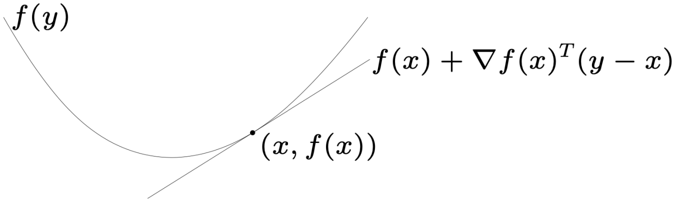
- f is **twice differentiable** if $\mathbf{dom}\, f$ is open and the Hessian

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

  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

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

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

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

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

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

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

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

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

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

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

convex iff $P \succeq 0$

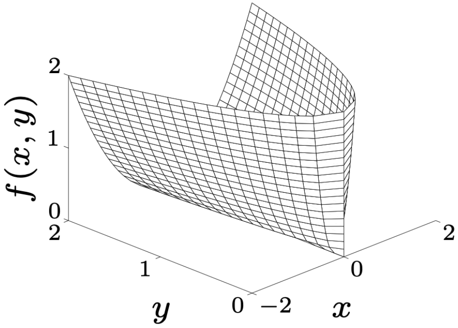
least-square objective: $f(x) = \|Ax - b\|_2^2$

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

convex for any A and b

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

$\mathbf{dom}\, f = \mathbb{R}^n$; 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 $\mathbb{R}^n_{++}$

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

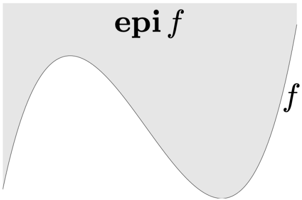
$$S_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$$

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

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 $\int_S p(x)\, \mathrm{d}\, x = 1$

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

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

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

Properties and examples

Operations preserving convexity

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
    - 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

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

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

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

examples

- distance to farthest point in a set C

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

- maximum eigenvalue of symmetric matrices

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