

3. Convex functions

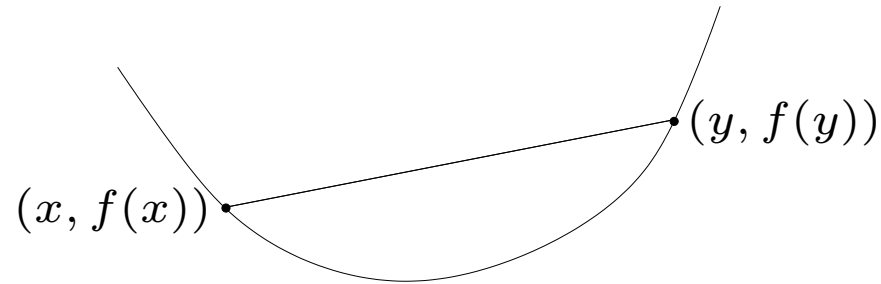
- basic properties and examples
- operations that preserve convexity
- the conjugate function
- quasiconvex functions
- log-concave and log-convex functions
- convexity with respect to generalized inequalities

Definition

$f : \mathbf{R}^n \rightarrow \mathbf{R}$ 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 $x, y \in \mathbf{dom} f$, $0 \leq \theta \leq 1$



- f is concave if $-f$ is convex
- f is strictly convex if $\mathbf{dom} f$ is convex and

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

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

Examples on \mathbf{R}

convex:

- affine: $ax + b$ on \mathbf{R} , for any $a, b \in \mathbf{R}$
- exponential: e^{ax} , for any $a \in \mathbf{R}$
- powers: x^α on \mathbf{R}_{++} , for $\alpha \geq 1$ or $\alpha \leq 0$
- powers of absolute value: $|x|^p$ on \mathbf{R} , for $p \geq 1$
- negative entropy: $x \log x$ on \mathbf{R}_{++}

concave:

- affine: $ax + b$ on \mathbf{R} , for any $a, b \in \mathbf{R}$
- powers: x^α on \mathbf{R}_{++} , for $0 \leq \alpha \leq 1$
- logarithm: $\log x$ on \mathbf{R}_{++}

Examples on \mathbf{R}^n and $\mathbf{R}^{m \times n}$

affine functions are convex and concave; all norms are convex

examples on \mathbf{R}^n

- affine function $f(x) = a^T x + b$
- norms: $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ for $p \geq 1$; $\|x\|_\infty = \max_k |x_k|$

examples on $\mathbf{R}^{m \times n}$ ($m \times n$ matrices)

- affine function

$$f(X) = \mathbf{tr}(A^T X) + b = \sum_{i=1}^m \sum_{j=1}^n A_{ij} X_{ij} + b$$

- spectral (maximum singular value) norm

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

Restriction of a convex function to a line

$f : \mathbf{R}^n \rightarrow \mathbf{R}$ is convex if and only if the function $g : \mathbf{R} \rightarrow \mathbf{R}$,

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

is convex (in t) for any $x \in \text{dom } f$, $v \in \mathbf{R}^n$

can check convexity of f by checking convexity of functions of one variable

example. $f : \mathbf{S}^n \rightarrow \mathbf{R}$ with $f(X) = \log \det X$, $\text{dom } f = \mathbf{S}_{++}^n$

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

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

g is concave in t (for any choice of $X \succ 0$, V); hence f is concave

Extended-value extension

extended-value extension \tilde{f} of f is

$$\tilde{f}(x) = f(x), \quad x \in \mathbf{dom} f, \quad \tilde{f}(x) = \infty, \quad x \notin \mathbf{dom} f$$

often simplifies notation; for example, the condition

$$0 \leq \theta \leq 1 \quad \implies \quad \tilde{f}(\theta x + (1 - \theta)y) \leq \theta \tilde{f}(x) + (1 - \theta) \tilde{f}(y)$$

(as an inequality in $\mathbf{R} \cup \{\infty\}$), means the same as the two conditions

- $\mathbf{dom} f$ is convex
- for $x, y \in \mathbf{dom} f$,

$$0 \leq \theta \leq 1 \quad \implies \quad f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y)$$

First-order condition

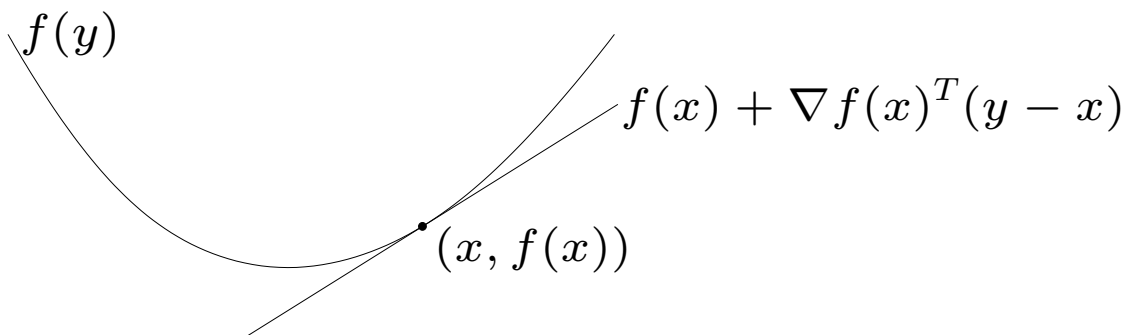
f is **differentiable** if $\text{dom } f$ is open and the gradient

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \dots, \frac{\partial f(x)}{\partial x_n} \right)$$

exists at each $x \in \text{dom } f$

1st-order condition: differentiable f with convex domain is convex iff

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



first-order approximation of f is global underestimator

Second-order conditions

f is **twice differentiable** if $\text{dom } f$ is open and the Hessian $\nabla^2 f(x) \in \mathbf{S}^n$,

$$\nabla^2 f(x)_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}, \quad i, j = 1, \dots, n,$$

exists at each $x \in \text{dom } f$

2nd-order conditions: for twice differentiable f with convex domain

- f is convex if and only if

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

- if $\nabla^2 f(x) \succ 0$ for all $x \in \text{dom } f$, then f is strictly convex

Examples

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

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

convex if $P \succeq 0$

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

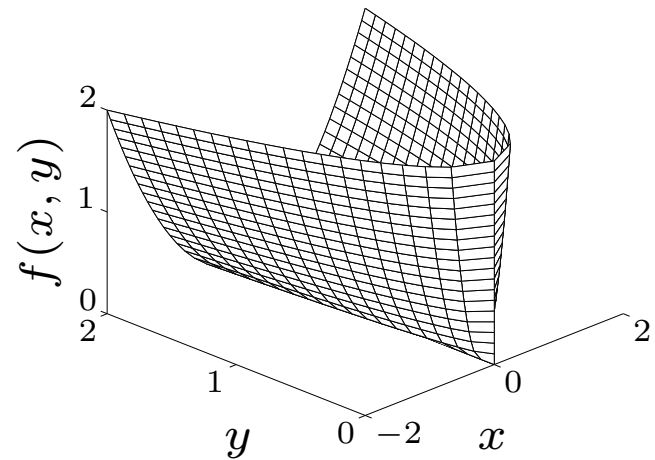
$$\nabla f(x) = 2A^T(Ax - b), \quad \nabla^2 f(x) = 2A^T A$$

convex (for any A)

quadratic-over-linear: $f(x, y) = x^2/y$

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

convex for $y > 0$



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

$$\nabla^2 f(x) = \frac{1}{\mathbf{1}^T z} \mathbf{diag}(z) - \frac{1}{(\mathbf{1}^T z)^2} z z^T \quad (z_k = \exp x_k)$$

to show $\nabla^2 f(x) \succeq 0$, we must verify that $v^T \nabla^2 f(x) v \geq 0$ for all v :

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

since $(\sum_k v_k z_k)^2 \leq (\sum_k z_k v_k^2)(\sum_k z_k)$ (from Cauchy-Schwarz inequality)

geometric mean: $f(x) = (\prod_{k=1}^n x_k)^{1/n}$ on \mathbf{R}_{++}^n is concave

(similar proof as for log-sum-exp)

Epigraph and sublevel set

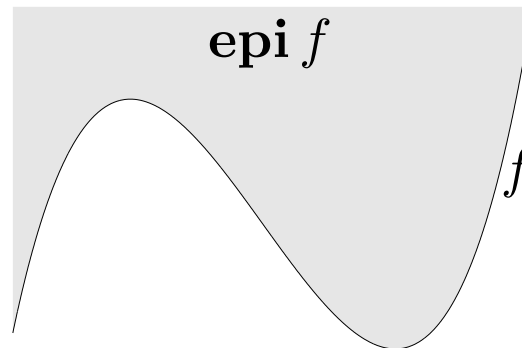
α -sublevel set of $f : \mathbf{R}^n \rightarrow \mathbf{R}$:

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

sublevel sets of convex functions are convex (converse is false)

epigraph of $f : \mathbf{R}^n \rightarrow \mathbf{R}$:

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



f is convex if and only if $\mathbf{epi} f$ is a convex set

Jensen's inequality

basic inequality: if f is convex, then for $0 \leq \theta \leq 1$,

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

extension: if f is convex, then

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

for any random variable z

basic inequality is special case with discrete distribution

$$\mathbf{prob}(z = x) = \theta, \quad \mathbf{prob}(z = y) = 1 - \theta$$

Operations that preserve convexity

practical methods for establishing convexity of a function

1. verify definition (often simplified by restricting to a line)
2. for twice differentiable functions, show $\nabla^2 f(x) \succeq 0$
3. show that f is obtained from simple convex functions by operations that preserve convexity
 - nonnegative weighted sum
 - composition with affine function
 - pointwise maximum and supremum
 - composition
 - minimization
 - perspective

Positive weighted sum & composition with affine function

nonnegative multiple: αf is convex if f is convex, $\alpha \geq 0$

sum: $f_1 + f_2$ convex if f_1, f_2 convex (extends to infinite sums, integrals)

composition with affine function: $f(Ax + b)$ is convex if f is convex

examples

- log barrier for linear inequalities

$$f(x) = -\sum_{i=1}^m \log(b_i - a_i^T x), \quad \text{dom } f = \{x \mid a_i^T x < b_i, i = 1, \dots, m\}$$

- (any) norm of affine function: $f(x) = \|Ax + b\|$

Composition with scalar functions

composition of $g : \mathbf{R}^n \rightarrow \mathbf{R}$ and $h : \mathbf{R} \rightarrow \mathbf{R}$:

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

f is convex if $\begin{array}{l} g \text{ convex, } h \text{ convex, } \tilde{h} \text{ nondecreasing} \\ g \text{ concave, } h \text{ convex, } \tilde{h} \text{ nonincreasing} \end{array}$

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

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

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

examples

- $\exp g(x)$ is convex if g is convex
- $1/g(x)$ is convex if g is concave and positive

Vector composition

composition of $g : \mathbf{R}^n \rightarrow \mathbf{R}^k$ and $h : \mathbf{R}^k \rightarrow \mathbf{R}$:

$$f(x) = h(g(x)) = h(g_1(x), g_2(x), \dots, g_k(x))$$

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

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

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

examples

- $\sum_{i=1}^m \log g_i(x)$ is concave if g_i are concave and positive
- $\log \sum_{i=1}^m \exp g_i(x)$ is convex if g_i are convex

Minimization

if $f(x, y)$ is convex in (x, y) and C is a convex set, then

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

is convex

examples

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

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

minimizing over y gives $g(x) = \inf_y f(x, y) = x^T (A - B C^{-1} B^T) x$

g is convex, hence Schur complement $A - B C^{-1} B^T \succeq 0$

- distance to a set: $\mathbf{dist}(x, S) = \inf_{y \in S} \|x - y\|$ is convex if S is convex