COMP 9602: Convex Optimization

Convex Function

Dr. C Wu

Department of Computer Science
The University of Hong Kong

Convex functions

- Definition and basic properties
- Operations that preserve convexity
- Conjugate function
- Quasi-convex function
- log-concave and log-convex functions

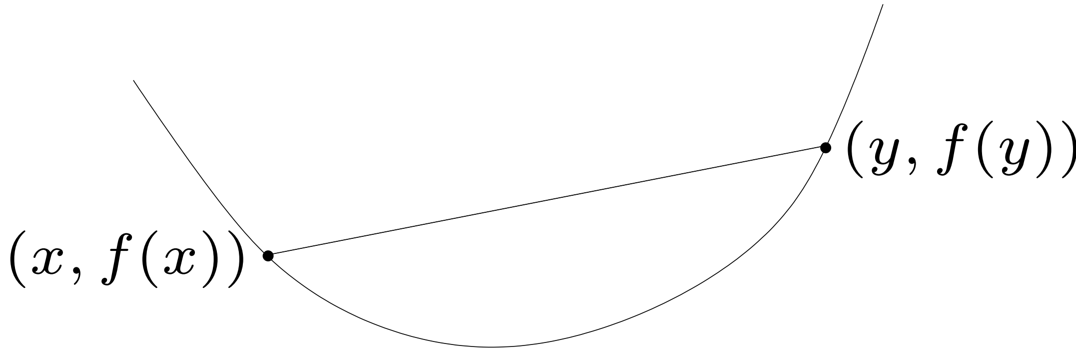
Definition

- Convex function

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



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

- Concave function
  - f is concave iff -f is convex

    $f(\theta x + (1 - \theta)y) \ge \theta f(x) + (1 - \theta)f(y)$
  - f is strict concave iff -f is strictly convex

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

Examples on 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 R^n

- Affine functions are both convex and concave
- all norms are convex

examples

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

Epigraph of a function

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

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

a subset ∈ dom f in R^n

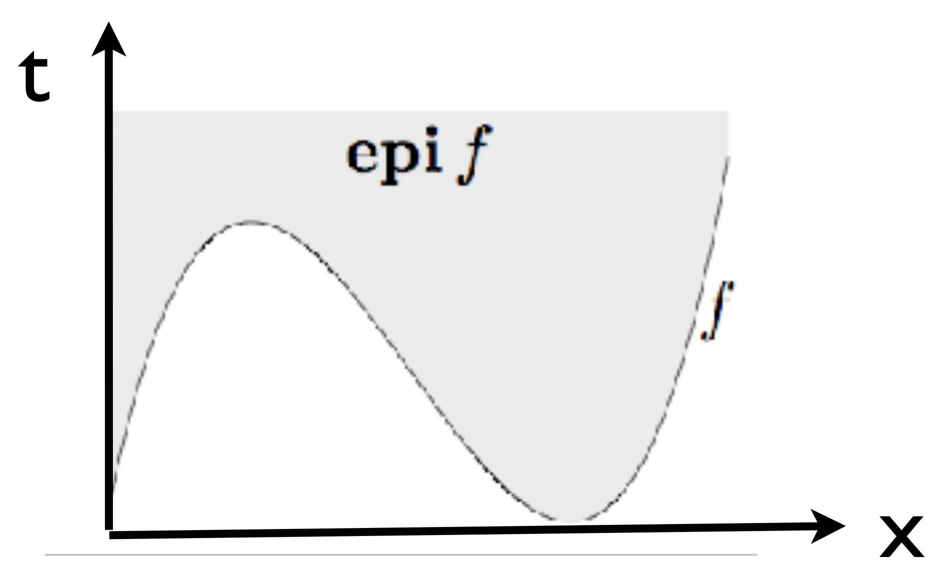
sublevel sets of convex functions are convex

if all sublevel sets of a function are convex, is the function necessarily convex?

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

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

a set in R^{n+1}



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

Differentiable convex functions

Differentiable functions

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

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

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

- f (defined on an open domain) is convex => f is continuous
- f is convex ~~=>~~ differentiable at all points

First-order condition

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

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

first-order approximation of f is always underestimator

- strictly convex: $\geq \Rightarrow >$
- proof:

Second-order condition

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

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

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

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

- f is convex if and only if

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

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

  converse not true, e.g., $f(x) = x^4$

Second-order condition

- Examples

quadratic functions

$$f(x) = x^2$$

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

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

convex iff $P \succeq 0$

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

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

convex (for any A)

Global minimum

(1) If f is convex, differentiable, and $\nabla f(x^*) = 0$, then f(y) ≥ f(x*), ∀ y, i.e., x* is the global minimum of f(x)

(2) If f is twice differentiable and $\nabla f(x^*) = 0, \nabla^2 f(x\ ) \succeq 0$, then f(y)≥ f(x*), ∀y, i.e., x* is a global minimum of f(x)

Note:

- a function may not have a global minimum even if it is convex,
  e.g. $f(x) = x \qquad f(x) = e^x$
- these properties only apply to unconstrained minimization,
  e.g., $f(x, y) = x^2 + y^2$

Local minimum

Local minimum: x* is a local minimum of unconstrained function f if it is no worse than its neighbors, i.e.,

$$\exists \epsilon > 0, \text{ s.t.}, f(x^*) \leq f(x), \forall x \text{ with } \| x - x^* \|_2 \leq \epsilon$$

- **Property**: Let x^* be a local minimum of $f : R^n \rightarrow R$. Suppose f is twice differentiable, then $\bigtriangledown f(x^*) = 0, \bigtriangledown^2 f(x^*) \succeq 0$

Operations that preserve convexity

practical methods for establishing convexity of a function

1. verify definition

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

Operations that preserve convexity

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 iff$f$ is convex

$\forall A \in R^{m \times n}, b \in R^m$

examples

- log barrier for linear inequalities

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

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

Operations that preserve convexity

pointwise maximization:

if $f_1, \ldots, f_m$ are convex, then $f(x) = \max\{f_1(x), \ldots, f_m(x)\}$ is convex

how about f(x)=min{f1(x),...,fm(x)}?

examples

- piecewise-linear function: $f(x) = \max_{i=1,\ldots,m}(a_i^T x + b_i)$ is convex
- sum of r largest components of $x \in \mathbf{R}^n$:

$$f(x) = x_{[1]} + x_{[2]} + \cdots + x_{[r]}$$

is convex ($x_{[i]}$ is ith largest component of x)

proof:

$$f(x) = \max\{x_{i_1} + x_{i_2} + \cdots + x_{i_r} \mid 1 \le i_1 < i_2 < \cdots < i_r \le n\}$$

Operations that preserve convexity

pointwise supremum:

if $f(x, y)$ is convex in x for each $y \in \mathcal{A}$, then

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

is convex

examples

- distance to farthest point in a set C:

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

Operations that preserve convexity

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

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

Operations that preserve convexity

convex

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

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

Composition of scalar functions:

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

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

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

- discover composition rules for n=1, twice differentiable g, h

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

(very similar rules hold in the general case, p. 84 textbook)

examples

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

Operations that preserve convexity

Vector composition:

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

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

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

- discover composition rules for n=1, twice differentiable g, h

see appendix A.4 for chain rule

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

(very similar rules hold in the general case, p. 86 textbook)

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

(p. 74)

Operations that preserve convexity

Perspective:

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

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

g is convex if f is convex

examples

- $f(x) = x^T x$ is convex; hence $g(x,t) = x^T x/t$ is convex for $t > 0$
- negative logarithm $f(x) = -\log x$ is convex; hence relative entropy $g(x,t) = t\log t - t\log x$ is convex on $\mathbf{R}^2_{++}$
- if f is convex, then

$$g(x) = (c^T x + d) f\left((Ax+b)/(c^T x + d)\right)$$

is convex on $\{x \mid c^T x + d > 0,\ (Ax+b)/(c^T x + d) \in \mathbf{dom}\, f\}$

Conjugate function

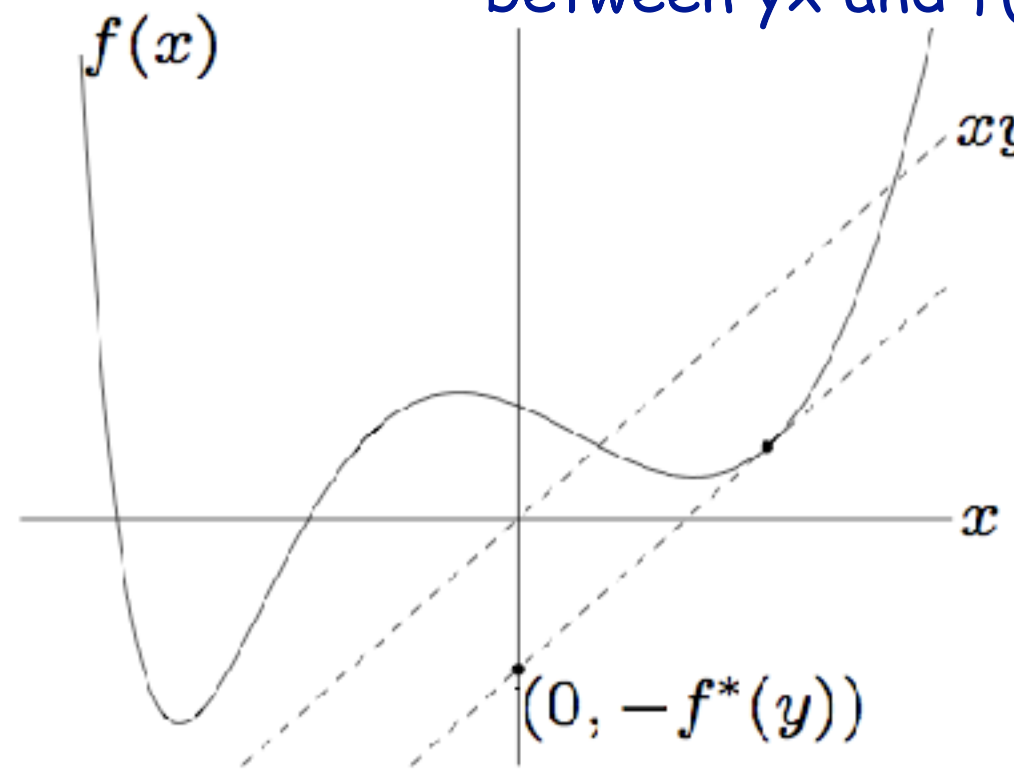
the **conjugate** of a function f is

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

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

  Pointwise supremum of a family of affine functions of y

f*(y) is the max. gap between yx and f(x)



Examples:

- *Affine function.* $f(x) = ax + b$. $\quad f^*(y) = \sup_x (yx - ax - b)$

- *Negative logarithm.* $f(x) = -\log x$, with $\mathbf{dom}\, f = \mathbf{R}_{++}$.

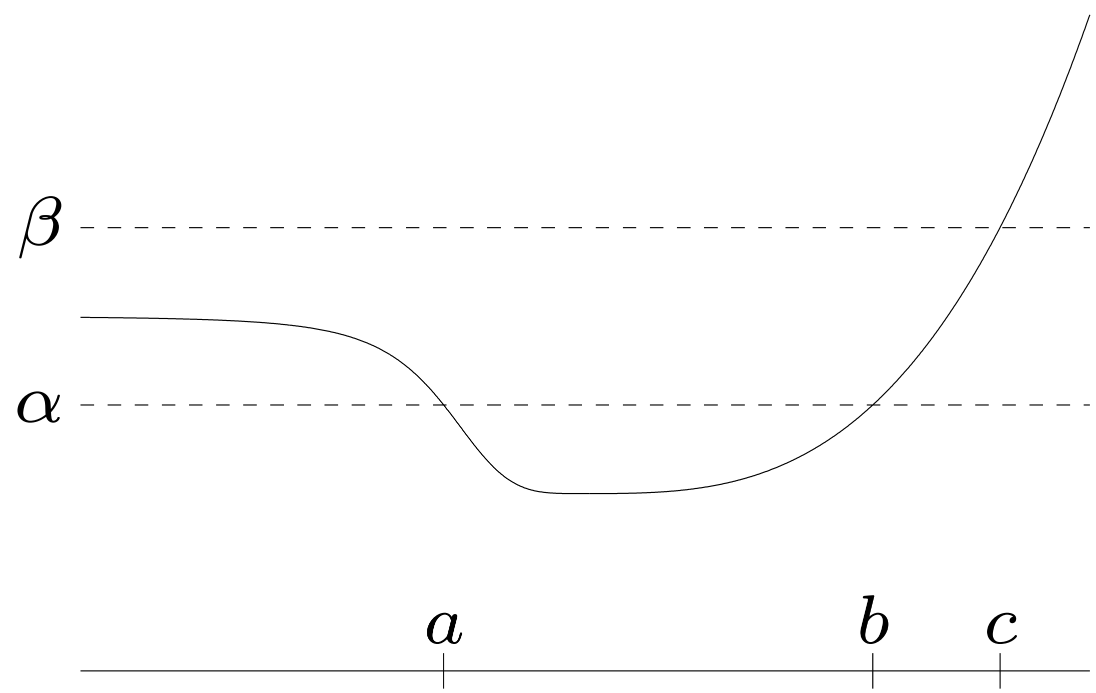
$$\begin{aligned} f^*(y) &= \sup_{x>0}(xy + \log x) \\ &= \begin{cases} -1 - \log(-y) & y < 0 \\ \infty & \text{otherwise} \end{cases} \end{aligned}$$

Quasi-convex functions

$f : \mathbf{R}^n \to \mathbf{R}$ 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 α



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

Quasi-convex functions

Examples

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

Quasi-convex functions

Properties

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

first-order condition: differentiable f with cvx domain is quasiconvex iff

$$f(y) \leq f(x) \quad \Longrightarrow \quad \nabla f(x)^T (y - x) \leq 0$$

Operations that preserve quasi-convexity: Chapter 3.4.4

Log-concave and log-convex functions

log-concave:

a positive function f is log-concave if $\log f$ is concave:

$$f(\theta x + (1-\theta)y) \geq f(x)^{\theta} f(y)^{1-\theta} \quad \text{for } 0 \leq \theta \leq 1$$

log-convex:

f is log-convex if $\log f$ is convex

example:

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

- Reference
  - Chapter 3, Convex Optimization.
- Acknowledgement
  - Some materials are extracted from the slides created by Prof. Stephen Boyd for the textbook
  - Some materials are extracted from the lecture notes of Convex Optimization by Prof. Wei Yu at the University of Toronto