Convex Optimization

Lecture 3 - Convex Functions

Instructor: Yuanzhang Xiao

University of Hawaii at Manoa

Fall 2017

Today's Lecture

Outline

Convex Functions

Definition of Convex Functions

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

- the domain $\text{dom} f$ is a convex set, and
- for all $x, y \in \text{dom} f$ and $\theta \in [0, 1]$, we have

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

Strictly Convex Functions and Concave Functions

strictly convex: if $\mathrm{dom} f$ is convex and

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

for all $x, y \in \mathrm{dom} f$ and $\theta \in (0,1)$

concave: if $\mathrm{dom} f$ is convex and $-f$ is convex

Equivalent Definition – Restriction to a Line

Restriction of Convex Function to a Line

$f : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if for any $x \in \mathrm{dom} f$ and $v \in \mathbb{R}^n$, the function $g : \mathbb{R} \to \mathbb{R}$, where

$$g(t) = f(x + tv), \ \mathrm{dom} g = \{t | x + tv \in \mathrm{dom} f\}$$

is convex in t.

check convexity of f $\to$ check convexity of g of one variable

Questions

- Prove the equivalence with the original definition.
  "$\Rightarrow$": apply original definition on g
  "$\Leftarrow$": for any $x, y \in \mathrm{dom} f$, choose x and $v = y - x$, and use convexity of g

Equivalent Definition – Restriction to a Line

Restriction of Convex Function to a Line

$f : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if for any $x \in \mathrm{dom} f$ and $v \in \mathbb{R}^n$, the function $g : \mathbb{R} \to \mathbb{R}$, where

$$g(t) = f(x + tv), \; \mathrm{dom} g = \{t | x + tv \in \mathrm{dom} f\}$$

is convex in t.

check convexity of f $\to$ check convexity of g of one variable

Questions

- Prove the equivalence with the original definition.
  "$\Rightarrow$": apply original definition on g
  "$\Leftarrow$": for any $x, y \in \mathrm{dom} f$, choose x and $v = y - x$, and use convexity of g

Equivalent Definition – First-Order Condition

f is differentiable if $\mathbf{dom} f$ is open and its gradient

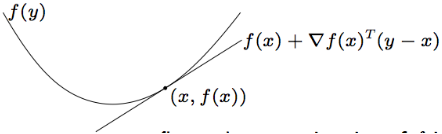
$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \ldots, \frac{\partial f(x)}{\partial x_n} \right)$$

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

First-Order Condition

f is convex if and only if $\mathbf{dom} f$ is convex, and for any $x, y \in \mathbf{dom} f$,

$$f(y) \geq f(x) + \nabla f(x)^T (y - x)$$



Equivalent Definition – First-Order Condition

Important implications for a convex function f

- local information (i.e., value and gradient) gives us global information (i.e., global underestimator)
- $\nabla f(x) = 0 \Leftrightarrow x$ is a global minimizer of f

Not true for non-convex functions

Questions

- Prove the equivalence with the original definition.
  prove it for $x \in \mathbb{R}$ first;
  use the restrictions to a line for general x

Equivalent Definition – First-Order Condition

Important implications for a convex function f

- local information (i.e., value and gradient) gives us global information (i.e., global underestimator)
- $\nabla f(x) = 0 \Leftrightarrow x$ is a global minimizer of f

Not true for non-convex functions

Questions

- Prove the equivalence with the original definition.
  prove it for $x \in \mathbb{R}$ first;
  use the restrictions to a line for general x

Equivalent Definition – Second-Order Condition

f is twice-differentiable if $\mathbf{dom} f$ is open and its Hessian (matrix)

$$\nabla^2 f(x) \in \mathbb{S}^n, \;\; [\nabla f(x)]_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}$$

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

Second-Order Condition

f is convex if and only if $\mathbf{dom} f$ is convex, and for any $x \in \mathbf{dom} f$,

$$\nabla^2 f(x) \succeq 0 \text{ (positive semidefinite)}$$

f convex, $x \in \mathbb{R} \Leftrightarrow \nabla f$ non-decreasing $\Leftrightarrow \nabla^2 f \geq 0$

Questions

- Can we drop the requirement of $\mathbf{dom} f$ being convex?
  No, $f(x) = \frac{1}{x^2}$ with $\mathbf{dom} f = \mathbb{R} \setminus \{0\}$

Equivalent Definition – Second-Order Condition

f is twice-differentiable if $\mathbf{dom} f$ is open and its Hessian (matrix)

$$\nabla^2 f(x) \in \mathbb{S}^n, \ \ [\nabla f(x)]_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}$$

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

Second-Order Condition

f is convex if and only if $\mathbf{dom} f$ is convex, and for any $x \in \mathbf{dom} f$,

$$\nabla^2 f(x) \succeq 0 \text{ (positive semidefinite)}$$

f convex, $x \in \mathbb{R} \Leftrightarrow \nabla f$ non-decreasing $\Leftrightarrow \nabla^2 f \geq 0$

Questions

- Can we drop the requirement of $\mathbf{dom} f$ being convex?
  No, $f(x) = \frac{1}{x^2}$ with $\mathbf{dom} f = \mathbb{R} \setminus \{0\}$

Outline

Examples on $\mathbb{R}$

convex functions on $\mathbb{R}$:

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

concave functions on $\mathbb{R}$:

- affine: $ax + b$ on $\mathbb{R}$ for any $a, b \in \mathbb{R}$
- powers: x^a on $\mathbb{R}_{++}$ for $a \in [0, 1]$
- logarithm: $\log x$ on $\mathbb{R}_{++}$

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

convex functions on $\mathbb{R}^n$ and $\mathbb{R}^{m\times n}$:

- affine: $a^T x + b$ on $\mathbb{R}^n$ for any $a \in \mathbb{R}^n,\ b \in \mathbb{R}$
- norms: $\|x\|$ on $\mathbb{R}^n$ (e.g., $\|x\|_p = \left(\sum_{i=1}^n x_i^p\right)^{1/p}$, $p \geq 1$; $\|x\|_\infty = \max_i |x_i|$)
- max: $f(x) = \max\{x_1, \ldots, x_n\}$
- quadratic-over-linear: $f(x, y) = x^2/y$ on $\mathbb{R} \times \mathbb{R}_{++}$
- log-sum-exp: $f(x) = \log\left(e^{x_1} + \cdots + e^{x_n}\right)$ on $\mathbb{R}^n$
- spectral norm (i.e., maximum singular value): $f(X) = \sigma_{\max}(X) = \left(\lambda_{\max}(X^T X)\right)^{1/2}$ on $\mathbb{R}^{m\times n}$

concave functions on $\mathbb{R}^n$ and $\mathbb{R}^{m\times n}$:

- affine: $a^T x + b$ on $\mathbb{R}^n$ for any $a \in \mathbb{R}^n,\ b \in \mathbb{R}$
- geometric mean: $f(x) = (\Pi_{i=1}^n x_i)$ on $\mathbb{R}^n_{++}$
- log-determinant: $f(X) = \log \det X$ on $\mathbb{S}^n_{++}$

Outline

Operations That Preserve Convexity

How to decide whether a function f is convex?

Method 1: By definition and equivalent conditions

- restriction to a line
- first-order conditions
- second-order conditions

Method 2: 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

Nonnegative Weighted Sum

if $f_1, \ldots, f_m$ are convex, the nonnegative weighted sum

$$f = w_1 f_1 + \cdots w_m f_m$$

is convex

Extension to infinite sums and integrals: if $f(x, y)$ is convex in x for any $y \in \mathcal{A}$, and $w(y) \geq 0$ for any $y \in \mathcal{A}$, then

$$g(x) = \int_{\mathcal{A}} w(y) f(x, y) dy$$

is convex (provided the integral exists)

Composition With Affine Function

if $f : \mathbb{R}^n \to \mathbb{R}$ is convex, and $A \in \mathbb{R}^{n\times m}$, $b \in \mathbb{R}^n$, then $g : \mathbb{R}^m \to \mathbb{R}$ defined as

$$g(x) = f(Ax + b),\ \mathbf{dom}\, g = \{x | Ax + b \in \mathbf{dom} f\}$$

is convex

useful examples:

- log barrier for linear inequalities: (in interior-point methods)

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

- norm of affine function

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

Pointwise Maximum

if $f_1, \ldots, f_m$ are convex, the pointwise maximum

$$f(x) = \max\{f_1(x), \ldots, f_m(x)\}$$

is convex

useful examples:

- pointwise linear function

$$f(x) = -\max_{i=1,\ldots,m}\left(a_i^T x + b_i\right),$$

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

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

where $x_{[i]}$ is the ith largest element of x

Pointwise Supremum

if $f(x, y)$ is convex in x for any $y \in \mathcal{A}$, the pointwise supremum

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

is convex

useful examples:

- support function of a set C

$$S_C(x) = \sup \left\{ x^T y | y \in C \right\},$$

- distance to the farthest point of a set C

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

Composition With Scalar Functions

composition of $h : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$: $f(x) = h[g(x)]$

f is convex if either one of the two holds:

- h convex, $\tilde{h}$ nondecreasing, g convex; or
- h convex, $\tilde{h}$ nonincreasing, g concave

where $\tilde{h}$ is extended-value extension of h:

$$\tilde{h}(x) = \begin{cases} h(x) & x \in \operatorname{dom} f \\ \infty & x \notin \operatorname{dom} f \end{cases}$$

f is concave if either one of the two holds:

- h concave, $\tilde{h}$ nondecreasing, g concave; or
- h concave, $\tilde{h}$ nonincreasing, g convex

where $\tilde{h}$ is extended-value extension of h:

$$\tilde{h}(x) = \begin{cases} h(x) & x \in \operatorname{dom} f \\ -\infty & x \notin \operatorname{dom} f \end{cases}$$

Composition With Scalar Functions

examples of h:

- $h(x) = \log x$ with $\mathbf{dom}\, h = \mathbb{R}_{++}$: concave, $\tilde{h}$ nondecreasing
- $h(x) = x^{1/2}$ with $\mathbf{dom}\, h = \mathbb{R}_{+}$: concave, $\tilde{h}$ nondecreasing
- $h(x) = x^{3/2}$ with $\mathbf{dom}\, h = \mathbb{R}_{+}$: convex, $\tilde{h}$ not nondecreasing
- $h(x) = \begin{cases} x^{3/2} & x \geq 0 \\ 0 & x < 0 \end{cases}$ with $\mathbf{dom}\, h = \mathbb{R}$: convex, $\tilde{h}$ nondecreasing

examples of simple compositions:

- $\exp g(x)$ is convex if g is convex
- $1/g(x)$ is convex if g is concave and positive
- $g(x)^p$ is convex if g is convex and nonnegative and $p \geq 1$
- $\log g(x)$ is concave if g is concave and positive

Questions

- Can we replace monotonicity of $\tilde{h}$ with monotonicity of h?
  No; $g(x) = x^2$ with $\mathbf{dom}\, g = \mathbb{R}$, $h(x) = 0$ with $\mathbf{dom}\, h = [1, 2]$

Composition With Vector Functions

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

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

f is convex if either one of the two holds:

- h convex, $\tilde{h}$ nondecreasing in each argument, g_i, $\forall i$ convex; or
- h convex, $\tilde{h}$ nonincreasing in each argument, g_i, $\forall i$ concave

f is concave if either one of the two holds:

- h concave, $\tilde{h}$ nondecreasing in each argument, g_i, $\forall i$ concave; or
- h concave, $\tilde{h}$ nonincreasing in each argument, g_i, $\forall i$ convex

Minimization

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

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

is convex

useful examples

- distance to a set:

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

  is convex if S is convex

note the difference from pointwise maximium

Perspective

if $f : \mathbb{R}^n \to \mathbb{R}$ is convex, its perspective

$$g(x,t) = t \cdot f(x/t), \text{ with dom} g = \{(x,t) | x/t \in \text{dom} f, t > 0\}$$

is convex in (x, t)

useful examples

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

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

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

Outline

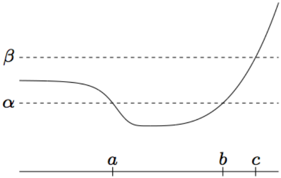
Quasiconvex Functions

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

$$S_\alpha = \{(x \in \mathrm{dom} f | f(x) \leq \alpha\}$$

$f : \mathbb{R}^n \to \mathbb{R}$ is quasiconvex if

- $\mathrm{dom} f$ is convex, and
- all subleval sets S_α are convex



f is quasiconcave if $-f$ is quasiconvex

f is quasilinear if f is quasiconvex and quasiconcave

Quasiconvex Functions

equivalent conditions: f is quasiconvex if and only if $\mathbf{dom} f$ is convex and

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

useful examples:

- $\sqrt{|x|}$ is quasiconvex on $\mathbb{R}$
- ceiling function $\mathrm{ceil}(x) = \inf\{z \in \mathbb{Z} | z \ge x\}$ is quasilinear
- $\log x$ is quasilinear on $\mathbb{R}_{++}$
- $f(x, y) = xy$ is quasiconcave on $\mathbb{R}^2_{++}$

operations that preserve quasiconvexity:

- pointwise maximum, composition, minimization

summation does not preserve quasiconvexity