CS703 - Optimization and Computing Notes

Nong Minh Hieu[1,2]

[1] School of Physical and Mathematical Sciences, Nanyang Technological University (NTU - Singapore)
[2] School of Computing and Information Systems, Singapore Management University (SMU - Singapore)

Contents

1 Introduction

Definition 1.1 (Optimization Problem).
Generally, an optimization problem is defined as follows:

$$\begin{aligned} \text{minimize}: &\ f_0(x) \\ \text{subject to}: &\ f_i(x) \leqslant 0, \quad i = 1, \ldots, m \\ &\ h_i(x) = 0, \quad i = 1, \ldots, p. \end{aligned} \tag{1}$$

Where we have:

1. $x \in \mathbb{R}^n$ is the optimization variable.
2. $f_0 : \mathbb{R}^n \to \mathbb{R}$ is the opjective (cost function).
3. $f_i : \mathbb{R}^n \to \mathbb{R}$ are inequality constraints.
4. $h_i : \mathbb{R}^n \to \mathbb{R}$ are equality constraints.

Definition 1.2 (Convex Optimization Problem).
An optimization problem is a **convex optimization problem** if:

1. $f_0, f_1, \ldots, f_m$ are convex.
2. Equality constraints are affine.

The reason why we need convex optimization problems are:

1. Convex optimization problems can be solved optimally (no local minima).
2. Time required to solve convex optimization problems is polynomial (in terms of number of variables and constraints).

1.1 Convex Sets

Definition 1.3 (Lines).
Let $x_1, x_2 \in \mathbb{R}^n$. A line passing through x_1, x_2 is defined as:

$$L(x_1, x_2) = \Big\{x \in \mathbb{R}^n : x = \theta x_1 + (1 - \theta)x_2, \theta \in \mathbb{R}\Big\}. \tag{2}$$

When $\theta \in (0, 1)$, we restrict the line to the points between x_1 and x_2 (exclusive).
Definition 1.4 (Affine Sets).
An affine set contains its elements' **affine combinations**: If $x_1, \ldots, x_k$ belongs to an affine set A, then it contains the affine combination

$$\sum_{i=1}^{k} \theta_i x_i \in A, \quad \theta_i \in \mathbb{R}, \sum_{i=1}^{k} \theta_i = 1. \tag{3}$$

For example,

1. An empty set is affine because there is no point.
2. A singleton is affine because there is only one point.
3. A line (extends indefinitely) is affine.
4. Any vector space is affine.
5. Linear subspaces of a vector space is affine.

Definition 1.5 (Convex Sets).
A convex set contains its elements' **convex combinations**: If $x_1, \ldots, x_k$ belongs to an affine set A, then it contains the convex combination

$$\sum_{i=1}^{k} \theta_i x_i \in A, \quad \theta_i \in [0,1], \sum_{i=1}^{k} \theta_i = 1. \tag{4}$$

For example,

1. Norm ball $\left\{x : \|x\| \leqslant r\right\}$ for a given norm $\|\cdot\|$, radius r.
2. Hyperplane $\left\{x : a^\top x = b\right\}$ for given a, b.
3. Halfspace $\left\{x : a^\top x \leqslant b\right\}$ for given a, b.
4. Affine space $\left\{x : Ax = b\right\}$ for given A, b.

Definition 1.6 (Convex Hull).
Given a discrete set $C = \{x_1, \ldots, x_k\}$. The convex hull of C, denoted $\mathrm{conv}(C)$, is the set of all convex combinations of points in C:

$$\mathrm{conv}(C) = \left\{ \sum_{i=1}^{k} \theta_i x_i : x_i \in C, \theta_i \geqslant 0, \sum_{i=1}^{k} \theta_i = 1 \right\}. \tag{5}$$

Convex hulls are always convex.

1.2 Convex Functions

Definition 1.7 (Convex Function).
A function $f : \mathbb{R} \to \mathbb{R}$ is convex if

(i) $\mathbf{dom} f$ is convex.

(ii) $f(\alpha x + (1-\alpha)y) \leqslant \alpha f(x) + (1-\alpha) f(y), \forall x, y \in \mathbf{dom} f, \alpha \in [0,1]$.

A function f is concave if $-f$ is convex.
Definition 1.8 (Strictly Convex Function).
A function $f : \mathbb{R} \to \mathbb{R}$ is convex if

(i) $\mathbf{dom} f$ is convex.

(ii) $f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha) f(y), \forall x, y \in \mathbf{dom} f, \alpha \in [0,1]$.

A strictly convex function implies a unique global minimum.

Test for Convexity of Function: We have the following tests for the convexity of any (real-valued vector) functions.

1. First-order derivative test.
2. Second-order derivative test.
3. Restriction to a line.
4. Epigraph.

Proposition 1.1: Restriction to A Line

A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex iff $g_{x,\nu} : \mathbb{R} \to \mathbb{R}$, defined as $g_{x,\nu}(t) = f(x + t\nu)$ is convex in t for any $x \in \mathbf{dom} f$ and $\nu \in \mathbb{R}^n$. Furthermore,

$$\mathbf{dom} g_{x,\nu} = \Big\{ t \in \mathbb{R} : x + t\nu \in \mathbf{dom} f \Big\} \subset \mathbb{R}. \tag{6}$$

Proof (Proposition 1.1).
We prove from both directions.

(i) f **convex** $\implies g_{x,\nu}$ **convex**: There are two sub-sections to prove - $\mathbf{dom} g_{x,\nu}$ is convex and $g_{x,\nu}(\alpha t_1 + (1-\alpha)t_2) \leqslant \alpha g_{x,\nu}(t_1) + (1-\alpha) g_{x,\nu}(t_2),\ \forall \alpha \in (0,1); t_1, t_2 \in \mathbf{dom} g_{x,\nu}$.

- $\mathbf{dom} g_{x,\nu}$ is convex: Let $t_1, t_2 \in \mathbf{dom} g_{x,\nu}$. Then, we have $x + t_1\nu, x + t_2\nu \in \mathbf{dom} f$. Then, by convexity of $\mathbf{dom} f$, for all $\alpha \in (0,1)$, we have $x + [\alpha t_1 + (1-\alpha)t_2]\nu \in \mathbf{dom} f$. Hence, we have $\alpha t_1 + (1-\alpha)t_2 \in \mathbf{dom} g_{x,\nu}$, making $\mathbf{dom} g_{x,\nu}$ convex.
- For $\alpha \in (0,1)$ and $t_1, t_2 \in \mathbf{dom} g_{x,\nu}$:

$$\begin{aligned} g_{x,\nu}(\alpha t_1 + (1-\alpha)t_2) &= f\Big(\alpha(x + t_1\nu) + (1-\alpha)(x + t_2\nu)\Big) \\ &\leqslant \alpha f(x + t_1\nu) + (1-\alpha) f(x + t_2\nu) \\ &= \alpha g(t_1) + (1-\alpha) g(t_2). \end{aligned}$$

  Hence, g is convex.

(ii) $g_{x,\nu}$ **convex** $\implies f$ **convex**: This direction is proven similarly without much deviation.

□.

Proposition 1.2: First Order Derivative Test

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. Then, f is convex iff the following are satisfied for all $x, y \in \mathbf{dom} f$:

$$f(y) \geqslant \nabla f(x)^\top (y - x) + f(x). \tag{7}$$

Proof (Proposition 1.2).
We prove both directions separately.

(i) f convex $\implies$ Eqn. (7) holds: Let $\alpha \in (0,1)$. Then for every $x, y \in \mathbf{dom} f$, we have:

$$\begin{aligned} f(\alpha y + (1-\alpha)x) &= f(x + \alpha(y - x)) \\ &\leqslant \alpha f(y) + (1-\alpha) f(x). \end{aligned}$$

Set $\Delta = y - x$, we have:

$$\begin{aligned} & f(x + \alpha\Delta) - f(x) \leqslant \alpha[f(y) - f(x)] \\ \implies & \frac{f(x + \alpha\Delta) - f(x)}{\alpha\Delta} \leqslant \frac{f(y) - f(x)}{\Delta}. \end{aligned}$$

Taking the limit as $\alpha \to 0$ from both sides, we have $\nabla f(x)^\top \leqslant \frac{f(y)-f(x)}{y-x}$. Hence, we proved Eqn. (7) holds.

(ii) Eqn. (7) holds $\implies f$ convex: For any $\alpha \in (0,1)$. For any $x, y \in \mathbf{dom} f$, let $z = \alpha x + (1-\alpha)y$. Then, since Eqn. (7) holds, we have:

$$\begin{aligned} f(x) &\geqslant f(z) + \nabla f(z)^\top (x - z), \\ f(y) &\geqslant f(z) + \nabla f(z)^\top (y - z). \end{aligned}$$

Then, we have:

$$\begin{aligned}\alpha f(x) + (1-\alpha) f(y) &\geqslant f(z) + \nabla f(z)^\top [\alpha x + (1-\alpha) y - z] \\ &= f(z) \\ &= f(\alpha x + (1-\alpha) y).\end{aligned}$$

Hence, f is convex.

$\square$.

Proposition 1.3: Second Order Derivative Test

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice-differentiable function. Then, f is convex iff $\nabla^2 f(x) \succeq 0$ (positive semi-definite Hessian).

Proof (Proposition 1.3)**.**
For each $x \in \mathbf{dom} f$, we denote $\mathbf{g}_x = \nabla f(x)$ and $\mathbf{H}_x = \nabla^2 f(x)$. $\square$.

A List of Definitions

B Important Theorems

C Important Corollaries

D Important Propositions

Return to Contents

E References

Return to Contents