Notes on Convex Optimization

Fei Jie

Contents

Preface

This document is notes on the textbook Convex Optimization [1].

Chapter 1

Introduction

1.1 Mathematical optimization

A mathematical optimization problem has the form

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq b_i, \quad i = 1, \ldots, m \end{array} \tag{1.1}$$

Here the vector $x = (x_1, \ldots, x_n)$ is the optimization variable of the problem, the function $f_0 : \mathbb{R}^n \to \mathbb{R}$ is the objective function, the functions $f_i : \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots, m$ are the limits, or bounds, for the constraints. A vector x^* is called optimal, or a solution of the problem (1.1), if it has the smallest objective value among all vectors that satisfy the constraints: for any z with $f_1(z) \leq b_1, \ldots, f_m(z) \leq b_m$, we have $f_0(z) \geq f_0(x^*)$.

1.2 Least squares and linear programming

1.2.1 Least-squares problems

A least-squares problem is an optimization problem with no constraints and an objective which is a sum of squares of terms of the form $a_i^T x - b_i$:

$$\text{minimize} \quad f_0(x) = \|Ax - b\|_2^2 = \sum_{i=2}^{k} (a_i^T x - b_i)^2. \tag{1.2}$$

Here $A \in \mathbb{R}^{k \times n}$ (with $k \geq n$), a_i^T are the rows of A, and the vectors $x \in \mathbb{R}^n$ is the optimization variable.

1.2.2 Linear programming

In linear programming problem, the objective function and all constraint functions are linear:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & c_i^T x \leq b_i, \quad i = 1, \dots, m. \end{array} \tag{1.3}$$

Here the vectors $c, a_1, \dots, a_m \in \mathbb{R}^n$ and scalars $b_1, \dots, b_m \in \mathbb{R}$ are problem parameters that specify the obejctive and constraint functions.

1.3 Convex optimization

A convex optimization problem is one of the form

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq b_i, \quad i = 1, \dots, m, \end{array} \tag{1.4}$$

where the functions $f_0, \dots, f_m : \mathbb{R}^n \to \mathbb{R}$ are convex, i.e., satisfy

$$f_i(\alpha x + \beta y) \leq \alpha f(x) + \beta f(y) \tag{1.5}$$

for all $x, y \in \mathbb{R}^n$ and all $\alpha, \beta \in \mathbb{R}$ with $\alpha + \beta = 1, \alpha \geq 0, \beta \geq 0$.n

1.4 Nonlinear optimization

1.5 Outline

1.6 Notation

Part I

Theory

Chapter 2

Convex sets

2.1 Affine and convex sets

2.1.1 Lines and line segments

Suppose $x_1 \neq x_2$ are two different points in $\mathbb{R}^n$. Points of the form

$$y = \theta x_1 + (1-\theta)x_2,$$

where $\theta \in \mathbb{R}$, form the line passing through x_1 and x_2.

2.1.2 Affine sets

A set C is affine if the line through any two distinct points in C lies in C, i.e., for any $x_1, x_2 \in C$ and $\theta \in \mathbb{R}$, we have $\theta x_1 + (1-\theta)x_2 \in C$.

2.1.3 Affine dimension and relative interior

2.1.4 Convex sets

A set C is convex if the line segment between any two points in C lies in C, i.e., if for any $x_1, x_2 \in C$ and any θ with $0 \leq \theta \leq 1$, we have

$$\theta x_1 + (1-\theta)x_2 \in C.$$

2.1.5 Cones

A set C is called a cone, or nonnegative homogeneous, if for every $x \in C$ and $\theta \geq 0$, we have $\theta x \in C$. A set C is a convex cone if it is convex and a cone, which means that for any $x_1, x_2 \in C$ and $\theta_1, \theta_2 \geq 0$, we have

$$\theta_1 x_1 + \theta_2 x_2 \in C.$$

2.2 Some important examples

2.2.1 Hyperplanes and halfspaces

2.2.2 Euclidean balls and ellipsoids

2.2.3 Norm balls and norm cones

2.2.4 Polyhedra

2.2.5 The positive semidefinite cones

2.3 Operations that preserve convexity

2.3.1 Intersection

2.3.2 Affine functions

2.3.3 Linear-fractional and perspective functions

2.4 Generalized inequalities

2.4.1 Proper cones and generalized inequalities

A cone $K \subseteq \mathbb{R}^n$ is called a proper cone if it satisfies the following:

- K is convex.
- K is closed.
- K is solid, which means it has nonempty interior.
- K is pointed, which means that it contains no line (or equivalently, $x \in K, -x \in K \Rightarrow x = 0$).

A proper cone K can be used to define a generalized inequality, which is a partial order on $\mathbb{R}^n$ that has many of the properties of the standard ordering on $\mathbb{R}$.

2.4.2 Minimum and minimal elements

The most obvious difference between ordinary inequalities and generalized inequalities is that $\leq$ for $\mathbb{R}$ is a linear ordering: any two points are comparable, meaning either $x \leq y$ or $y \leq x$. This property does not hold for other generalized inequalities.

$x \in S$ is the minimum element of S (with respect to the generalized inequality $\preceq_K$) if for every $y \in S$ we have $x \preceq_K y$. If a set has minimum (maximum) element, then it is unique. $x \in S$ is a minimal element of S (with respect to the generalized inequality $\preceq_K$) if $y \in S$, $y \preceq_K x$ only if $y = x$.

2.5 Separating and supporting hyperplanes

2.6 Dual cones and generalized inequalities

Chapter 3

Convex functions

3.1 Basic properties and examples

3.2 Operations that preserve convexity

3.3 The conjugate function

3.4 Quasiconvex functions

3.5 Log-concave and log-convex functions

3.6 Convexity with respect to generalized inequalities

Chapter 4

Convex optimization problems

4.1 Optimization problems

4.2 Convex optimization

4.3 Linear optimization problems

4.4 Quadratic optimization problems

4.5 Geometric programming

4.6 Generalized inequality constraints

4.7 Vector optimization

Chapter 5

Duality

5.1 The Lagrange dual function

5.2 The Lagrange dual problem

5.3 Geometric interpretation

5.4 Saddle-point interpretation

5.5 Optimality conditions

5.6 Perturbation and sensitivity analysis

5.7 Examples

5.8 Theorems of alternatives

5.9 Generalized inequalities

Part II

Applications

Chapter 6

Approximation and fitting

Chapter 7

Statistical estimation

Chapter 8

Geometric problems

Part III

Algorithms

Chapter 9

Unconstrained minimization

Chapter 10

Equality constrained minimization

Chapter 11

Interior-point methods

Appendices

Appendix A

Mathematical background

Appendix B

Problems involving two quadratic functions

Appendix C

Numerical linear algebra background

Bibliography

[1] Boyd, S., and Vandenberghe, L. *Convex optimization.* Cambridge university press, 2004. iv

[2] Kipf, T. N., and Welling, M. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907* (2016).