Math 6336 – Optimization Theory

Andreas Mang

1 Introduction

1.1 Mathematical Optimization

def 1.1 *Let $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 0, \ldots, m$, $x = [x_1, \ldots, x_n] \in \mathbb{R}^n$. The general form of an* ***optimization problem*** *is given by*

$$\begin{aligned} &\underset{x \in \mathbb{R}^n}{\text{minimize}} && f_0(x) \\ &\text{subject to} && f_i(x) \leq b_i, \quad i = 1, \ldots, m. \end{aligned} \tag{1}$$

We call $x \in \mathbb{R}^n$ ***optimization*** or ***decision variable***, $f_0 : \mathbb{R}^n \to \mathbb{R}$ ***objective function***, and $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, n$ (***inequality***) ***constraint functions***, and $b_i \in \mathbb{R}$ ***limits***/***bounds*** for the constraints. The problem in Def 1.1 can be classified according to the nature of the constraints and objective function (linear, nonlinear, convex), the number of variables (large or small scale), the smoothness of the functions (differentiable or non-differentiable).

def 1.2 *We call $x^\star \in \mathbb{R}^n$ (an)* ***optimal*** *(point) or a* ***solution*** *of (1) if it has the smallest objective value among all vectors that satisfy the constraints: $\forall z \in \mathbb{R}^n \setminus \{x^\star\} : f_i(z) \leq b_i, i = 1, \ldots, m, \quad f_0(x^\star) \leq f_0(z)$.*

def 1.3 *A point $x \in \mathbb{R}^n$ is called a* ***feasible*** *if $f_i(x) \leq b_i$, $i = 1, \ldots, m$.*

def 1.4 *A (feasible) point $x^\star \in \mathbb{R}^n$ is called a* ***global minimizer*** *if $f_0(x^\star) \leq f_0(x)$ for all $x \in \mathbb{R}^n$.*

def 1.5 *A (feasible) point $x^\star \in \mathbb{R}^n$ is called a* ***local minimizer*** *if there is a neighborhood $\mathcal{N}$ of $x^\star$ such that $f_0(x^\star) \leq f_0(x)$ for all $x \in \mathcal{N}$.*

rem 1.1 *Concept can be extended to* ***strict*** *(local) minimizer by replacing $\leq$ with $<$.*

We call (1) a ***linear program*** if the objective function $f_0 : \mathbb{R}^n \to \mathbb{R}$ and the constraint functions $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$ are linear, i.e., $f_j(\alpha x + \beta y) = \alpha f_j(x) + \beta f_j(y)$ for all $x, y \in \mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$, $j = 0, \ldots, m$. If the optimization is not linear, it is called a ***nonlinear program***.

1.2 Convex Optimization

In this lecture, we will focus on ***convex optimization problems***. Convex optimization studies the problem of minimizing a ***convex function*** over a convex set. That is, in (1) the objective $f_0 : \mathbb{R}^n \to \mathbb{R}$ and the constraint functions $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, are convex functions, i.e., satisfy

$$f_j(\alpha x + \beta y) \leq \alpha f_j(x) + \beta f_i(x), \quad j = 0, \ldots, m,$$

for all $x, y \in \mathbb{R}^n$, $\alpha \geq 0, \beta \geq 0$, $\alpha + \beta = 1$. Convex optimization is a generalization of linear programming.

General Formulation

prb 1.1 *Given a convex objective function $f_0 : \mathbb{R}^n \to \mathbb{R}$, find $x^\star \in \mathcal{X} \subseteq \mathbb{R}^n$ such that*

$$x^\star = \operatorname{argmin} \{f_0(x) : x \in \mathcal{X}\}.$$

Alternative Formulation

prb 1.2 *Given a convex objective function $f_0 : \mathbb{R}^n \to \mathbb{R}$ and a set of convex constraint functions $f_i : \mathbb{R}^n \to \mathbb{R}$ we solve for $x \in \mathbb{R}^n$ as follows:*

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad f_0(x) \qquad \text{subject to} \quad f_i(x) \leq 0, i = 1, \dots, m.$$

Standard Form

prb 1.3 *Given a convex objective function $f_0 : \mathbb{R}^n \to \mathbb{R}$ and a set of convex constraint functions $f_i : \mathbb{R}^n \to \mathbb{R}$, $h_i : \mathbb{R}^n \to \mathbb{R}$, we solve for $x \in \mathbb{R}^n$ as follows:*

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leq 0, \quad i = 1, \dots, m \\ & h_i(x) = 0, \quad i = 1, \dots, p. \end{aligned} \tag{2}$$

ex 1.1 Special case for convex optimization problem: *least squares problem*:

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad \|Ax - b\|_2^2$$

with $A \in \mathbb{R}^{k,n}$, $k > n$.

ex 1.2 Special case for convex optimization problem: *linear program*:

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad & c^\top x \\ \text{subject to} \quad & a_i^\top x \leq b_i, \quad i = 1, \dots, m, \end{aligned}$$

with $c \in \mathbb{R}^n$, $a_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$.

1.3 Nonlinear (Nonconvex) Optimization

We use the term ***nonlinear optimization*** (or ***nonlinear programming***; **NLP**) for optimization problems of the form (1) when objective f_0 or constraint functions f_i are not linear. Nonlinear optimization problems can be convex and nonconvex. Nonconvex, nonlinear optimization problems are difficult to solve (but in general easy to formulate).

$$\text{linear programming} \subset \text{convex optimization} \subset \text{nonlinear programming}$$

1.3.1 Local Optimization

Compromise is to give up on seeking optimal x that minimizes objective f_0 over all feasible points. We seek a local minimizer instead.

+ methods only require differentiability of objective

+ can be used to solve large scale problems ($n \geq 1e6$)
− solution is (possibly) not global optimum
− solver requires (good) initial guess
− methods are sensitive to parameters

1.3.2 Global Optimization

Compromise is to give up on efficiency.

+ computed solution is global optimum
− methods are not efficient

1.3.3 Convex Optimization for Nonconvex Problems

1. find initial guess
2. provide heuristics for solving nonconvex problems
3. compute lower bound for global optimization / optimal value of nonconvex problem