Convex Optimization Basics

Optimization terminology

Reminder: a convex optimization problem (or program) is

$$\begin{aligned} \min_{x \in D} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \le 0, \ i = 1, \dots m \\ & Ax = b \end{aligned}$$

where f and g_i, $i = 1, \dots m$ are all convex, and the optimization domain is $D = \mathrm{dom}(f) \cap \bigcap_{i=1}^m \mathrm{dom}(g_i)$ (often we do not write D)

- f is called criterion or objective function
- g_i is called inequality constraint function
- If $x \in D$, $g_i(x) \le 0$, $i = 1, \dots m$, and $Ax = b$ then x is called a feasible point
- The minimum of $f(x)$ over all feasible points x is called the optimal value, written $f^\star$

- If x is feasible and $f(x) = f^\star$, then x is called optimal; also called a solution, or a minimizer[1]
- If x is feasible and $f(x) \leq f^\star + \epsilon$, then x is called ϵ-suboptimal
- If x is feasible and $g_i(x) = 0$, then we say g_i is active at x
- Convex minimization can be reposed as concave maximization

$$
\begin{array}{ll}
\min\limits_x & f(x) \\
\text{subject to} & g_i(x) \leq 0, \\
 & \quad i = 1, \ldots m \\
 & Ax = b
\end{array}
\quad \Longleftrightarrow \quad
\begin{array}{ll}
\max\limits_x & -f(x) \\
\text{subject to} & g_i(x) \leq 0, \\
 & \quad i = 1, \ldots m \\
 & Ax = b
\end{array}
$$

Both are called convex optimization problems

[1] Note: a convex optimization problem need not have solutions, i.e., need not attain its minimum, but we will not be careful about this

Convex solution sets

Let X_{opt} be the set of all solutions of convex problem, written

$$\begin{array}{rll} X_{\text{opt}} = & \operatorname{argmin} & f(x) \\ & \text{subject to} & g_i(x) \le 0, \ i = 1, \ldots m \\ & & Ax = b \end{array}$$

Key property: X_{opt} is a convex set

Proof: use definitions. If x, y are solutions, then for $0 \le t \le 1$,

- $tx + (1-t)y \in D$
- $g_i(tx + (1-t)y) \le tg_i(x) + (1-t)g_i(y) \le 0$
- $A(tx + (1-t)y) = tAx + (1-t)Ay = b$
- $f(tx + (1-t)y) \le tf(x) + (1-t)f(y) = f^\star$

Therefore $tx + (1-t)y$ is also a solution

Another key property: if f is strictly convex, then the solution is unique, i.e., X_{opt} contains one element

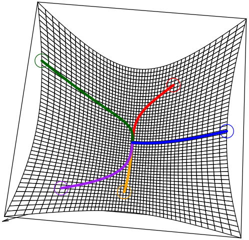
Local minima are global minima

For a convex problem, a feasible point x is called locally optimal is there is some $R > 0$ such that

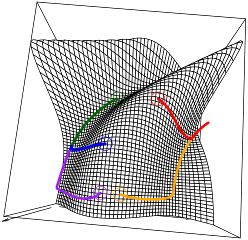
$$f(x) \le f(y) \quad \text{for all feasible } y \text{ such that } \|x - y\|_2 \le R$$

Reminder: for convex optimization problems, local optima are global optima

Proof simply follows from definitions



Convex



Nonconvex

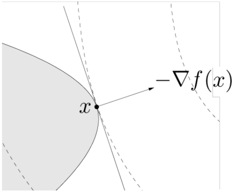
First-order optimality condition

For a convex problem

$$\min_x \; f(x) \;\; \text{subject to} \;\; x \in C$$

and differentiable f, a feasible point x is optimal if and only if

$$\nabla f(x)^T (y - x) \geq 0 \quad \text{for all } y \in C$$



This is called the first-order condition for optimality

In words: all feasible directions from x are aligned with gradient $\nabla f(x)$

Important special case: if $C = \mathbb{R}^n$ (unconstrained optimization), then optimality condition reduces to familiar $\nabla f(x) = 0$

Partial optimization

Reminder: $g(x) = \min_{y \in C} \; f(x,y)$ is convex in x, provided that f is convex in (x,y) and C is a convex set

Therefore we can always partially optimize a convex problem and retain convexity

E.g., if we decompose $x = (x_1, x_2) \in \mathbb{R}^{n_1+n_2}$, then

$$\begin{array}{ll} \displaystyle\min_{x_1,x_2} & f(x_1,x_2) \\ \text{subject to} & g_1(x_1) \le 0 \\ & g_2(x_2) \le 0 \end{array} \quad \Longleftrightarrow \quad \begin{array}{ll} \displaystyle\min_{x_1} & \tilde{f}(x_1) \\ \text{subject to} & g_1(x_1) \le 0 \end{array}$$

where $\tilde{f}(x_1) = \min\{f(x_1,x_2) : g_2(x_2) \le 0\}$. The right problem is convex if the left problem is

Transformations and change of variables

If $h : \mathbb{R} \to \mathbb{R}$ is a monotone increasing transformation, then

$$\begin{aligned} & \min_x \; f(x) \;\; \text{subject to} \;\; x \in C \\ \iff & \min_x \; h(f(x)) \;\; \text{subject to} \;\; x \in C \end{aligned}$$

Similarly, inequality or equality constraints can be transformed and yield equivalent optimization problems. Can use this to reveal the "hidden convexity" of a problem

If $\phi : \mathbb{R}^n \to \mathbb{R}^m$ is one-to-one, and its image covers feasible set C, then we can change variables in an optimization problem:

$$\begin{aligned} & \min_x \; f(x) \;\; \text{subject to} \;\; x \in C \\ \iff & \min_y \; f(\phi(y)) \;\; \text{subject to} \;\; \phi(y) \in C \end{aligned}$$

Eliminating equality constraints

Important special case of change of variables: eliminating equality constraints. Given the problem

$$\begin{aligned} \min_x \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0, \ i = 1, \ldots m \\ & Ax = b \end{aligned}$$

we can always express any feasible point as $x = My + x_0$, where $Ax_0 = b$ and $\mathrm{col}(M) = \mathrm{null}(A)$. Hence the above is equivalent to

$$\begin{aligned} \min_y \quad & f(My + x_0) \\ \text{subject to} \quad & g_i(My + x_0) \leq 0, \ i = 1, \ldots m \end{aligned}$$

Note: this is fully general but not always a good idea (practically)

Introducing slack variables

Essentially opposite to eliminating equality contraints: introducing slack variables. Given the problem

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0, \ i = 1, \dots m \\ & Ax = b \end{aligned}$$

we can transform the inequality constraints via

$$\begin{aligned} \min_{x,s} \quad & f(x) \\ \text{subject to} \quad & s_i \geq 0, \ i = 1, \dots m \\ & g_i(x) + s_i = 0, \ i = 1, \dots m \\ & Ax = b \end{aligned}$$

Note: this is no longer convex unless g_i, $i = 1, \dots, n$ are affine

Relaxing nonaffine equality constraints

Given an optimization problem

$$\min_x \; f(x) \;\; \text{subject to} \;\; x \in C$$

we can always take an enlarged constraint set $\tilde{C} \supseteq C$ and consider

$$\min_x \; f(x) \;\; \text{subject to} \;\; x \in \tilde{C}$$

This is called a relaxation and its optimal value is always smaller or equal to that of the original problem

Important special case: relaxing nonaffine equality constraints, i.e.,

$$h_j(x) = 0, \; j = 1, \ldots r$$

where h_j, $j = 1, \ldots r$ are convex but nonaffine, are replaced with

$$h_j(x) \le 0, \; j = 1, \ldots r$$

Linear program

A linear program or LP is an optimization problem of the form

$$\begin{aligned} \min_{x} \quad & c^T x \\ \text{subject to} \quad & Dx \le d \\ & Ax = b \end{aligned}$$

Observe that this is always a convex optimization problem

- Fundamental problem in convex optimization. Many diverse applications, rich history

Standard form

A linear program is said to be in standard form when it is written as

$$\begin{aligned} \min_{x} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

Any linear program can be rewritten in standard form (check this!)

Convex quadratic program

A convex quadratic program or QP is an optimization problem of the form

$$\begin{aligned} \min_x \quad & c^T x + \frac{1}{2} x^T Q x \\ \text{subject to} \quad & Dx \le d \\ & Ax = b \end{aligned}$$

where $Q \succeq 0$, i.e., positive semidefinite

Note that this problem is not convex when $Q \not\succeq 0$

From now on, when we say quadratic program or QP, we implicitly assume that $Q \succeq 0$ (so the problem is convex)

Standard form

A quadratic program is in standard form if it is written as

$$\begin{aligned} \min_x \quad & c^T x + \frac{1}{2} x^T Q x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

Any quadratic program can be rewritten in standard form