Optimization Fundamentals

- The generic optimization problem
- Properties of functions and sets
- Existence of optimal solutions
- Local vs. global optimality
- Convex programs

The Generic Optimization Problem

$$
\begin{aligned}
(P): \quad \min \quad & f(x), \\
\text{s.t.} \quad & x \in X \subseteq \mathbb{R}^n,
\end{aligned}
$$

where $f : \mathbb{R}^n \mapsto \mathbb{R}$ is the **objective function** and X is the **constraint set** or the set of **feasible solutions**.

W.l.o.g our discussion will be in terms of minimization problems. A maximization problem $\max\{f(x) \mid x \in X\}$ is equivalent to $-\min\{-f(x) \mid x \in X\}$.

Typically the constraint set is defined by inequality and equality constraints, as well as the domain of the values of the decision variables, for e.g.

$$
\begin{aligned}
X = \{x \in \mathbb{Z}^p \times \mathbb{R}^{n-p} \mid \; & g_i(x) \leq 0, i = 1, \ldots, I, \\
& h_j(x) = 0, j = 1, \ldots, J\}.
\end{aligned}
$$

The effort in solving (P) depends on the structure of f and X.

Important Properties of Functions

Continuity:

A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is **continuous at** x^0 if for every sequence $\{x_i\}$ such that $\lim_{i\to\infty} x_i = x^0$ we have $\lim_{i\to\infty} f(x_i) = f(x^0)$.

A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is **continuous** if it is continuous at all $x \in \mathbb{R}^n$.

Differentiability:

A function $f : \mathbb{R} \mapsto \mathbb{R}$ is **differentiable at** x^0 if $\lim_{t\downarrow 0} \frac{f(x^0+t)-f(x^0)}{t}$ exists.

A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is **differentiable at** x^0 if $\frac{\partial f(x)}{\partial x_i}|_{x=x^0}$ exists for all $i = 1, \ldots, n$.

A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is **differentiable** (or smooth) if it is differentiable at all $x \in \mathbb{R}^n$.

The **gradient** of a differentiable function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is defined as

$$\nabla f(x^0) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix}_{x=x^0}.$$

The **Hessian** of a (twice) differentiable function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is defined as

$$\nabla^2 f(x^0) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial^2 x_1} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \frac{\partial^2 f(x)}{\partial^2 x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f(x)}{\partial^2 x_n} \end{bmatrix}_{x=x^0}.$$

Taylor's approximation of a twice-differentiable function $f : \mathbb{R}^n \mapsto \mathbb{R}$ around x^0:

$$f(x) \approx f(x^0) + (x - x^0)^T \nabla f(x^0) + \frac{1}{2}(x - x^0)^T \nabla^2 f(x^0)(x - x^0).$$

Convexity:

A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is **convex** if

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2),$$

for every $x_1, x_2 \in \mathbb{R}^n$ and $\lambda \in [0, 1]$.

A (twice) differentiable function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is convex if its hessian $\nabla^2 f(x)$ is positive semi-definite for all $x \in \mathbb{R}^n$.

A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is **concave** if $-f$ is convex.

Important Properties of Sets

Closedness:

A set $X \in \mathbb{R}^n$ is **closed** if it includes its boundary points.

Formally, a set $X \in \mathbb{R}^n$ is closed if for every convergent sequence $\{x_i\} \subset X$, the limit point $\lim_{i\to\infty} x_i \in X$.

Boundedness:

A set $X \in \mathbb{R}^n$ is **bounded** if it can be enclosed in a large enough sphere.

Formally, a set $X \in \mathbb{R}^n$ is bounded if there exists a large number $M \geq 0$ such that $||x|| < M$ for all $x \in X$.

A set $X \in \mathbb{R}^n$ that is both closed and bounded is called **compact**.

Convexity:

A set $X \in \mathbb{R}^n$ is **convex** if for every $x_1, x_2 \in X$ and $\lambda \in [0,1]$, the point

$$(\lambda x_1 + (1-\lambda)x_2) \in X.$$

Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be a convex function, then the α-lower level set $X_{\leq \alpha} = \{x \mid f(x) \leq \alpha\}$ is a convex set.

Let the sets $X_1, \ldots, X_m$ be convex then the set $\cap_{i=1}^m X_i$ is also convex.

Let $X = \{x \mid g_i(x) \leq 0 \; i = 1, \ldots, m\}$. If g_i is a convex function for every i, then X is a convex set.

Existence of Optimal Solutions

A vector $x^* \in \mathbb{R}^n$ is a **feasible solution** of (P) if $x^* \in X$.

A vector $x^* \in \mathbb{R}^n$ is an **optimal solution** of (P) if $x^* \in X$ and $f(x^*) \leq f(x)$ for all $x \in X$.

Is (P) always guaranteed to have an optimal solution?

Example: Suppose $n = 1$, $f(x) = x$ and $X = \{x \mid x > 0\}$, i.e.

$$\min\{x \mid x > 0\}.$$

This problem does not have an optimal solution, because given any feasible solution x^*, you can construct another solution $x^* - \epsilon$ (where ϵ is a small positive number) which is feasible and has a better objective value.

Weirstrass' Theorem: *If the objective function f is continuous and the constraint set X is compact (i.e. closed and bounded), then problem (P) is guaranteed to have an optimal solution.*

The above condition is sufficient but not necessary.

Example: Let $f(x) = \begin{cases} x^2 & \text{if } x \geq 0, \\ 1 + |x| & \text{if } x < 0. \end{cases}$
In the problem $\min\{f(x) \mid -1 < x < \infty\}$, the objective function is discontinuous and the constraint set is neither closed nor bounded. However an optimal solution $(x = 0)$ exists.

Local vs. Global Optimality

A feasible solution $x^* \in X$ is a **local** optimal solution of (P) if $f(x^*) \leq f(x)$ for all $x \in N_\epsilon(x^*) \cap X$ for some $\epsilon > 0$. Here $N_\epsilon(x^*) = \{x \mid ||x - x^*|| \leq \epsilon\}$ is an ϵ-neighborhood of x^*.

A feasible solution $x^* \in X$ is a **global** optimal solution of (P) if $f(x^*) \leq f(x)$ for all $x \in X$.

Every global optima is a local optima, but not vice versa.

Convex Programs

A minimization problem is a **convex program** if the objective function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a convex function and the constraint set X is a convex set.

Theorem: *For a convex program every local optimal solution is also a global optimal solution.*