Contents

1

Non-linear Programming Problems

1.1 Basic Terminology and Notations

Definition 1.1.1 ▸ General Non-linear Programming (NLP) Problems

Define the function $f : \mathbf{R}^n \to \mathbf{R}$. Let $\boldsymbol{x} \in \mathbf{R}^n$ be a vector, then a general NLP problem aims to **optimise** (i.e. maximise or minimise) $f(\boldsymbol{x})$ subject to the constraint $\boldsymbol{x} \in S \subseteq \mathbf{R}^n$, where

- f is known as the **objective function**;
- S is known as the **feasible set**;
- A solution (point) $\boldsymbol{x} \in S$ is known as a **feasible solution (point)**. Otherwise, it is known as an **infeasible solution(point)**.

Remark. Note that to maximise $f(\mathrm{x})$ is equivalent to minimising $-f(\mathrm{x})$, so it suffices to only study minimisation problems.

The word "optimal", however, can be ambiguous due to its qualitative nature. Thus, we shall define what it means to be optimal quantitatively with more rigorous terms.

Definition 1.1.2 ▸ Optimal Solution

Consider a minimisation problem subject to constraint $\boldsymbol{x} \in S \subseteq \boldsymbol{R}^n$ whose objective function is $f(\boldsymbol{x})$. A feasible solution $\boldsymbol{x}^*$ is called an **optimal solution** if $f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in S$. We can write

$$\boldsymbol{x}^* = \operatorname*{argmin}_{x \in S} f(\boldsymbol{x}).$$

$f(\boldsymbol{x}^*)$ is then known as the **optimal value**.

Remark. For maximisation problems, we can write

$$\boldsymbol{x}^* = \operatorname*{argmax}_{x \in S} f(\boldsymbol{x})$$

Note that not all optimisation problems have an optimal solution. We shall still expect to encounter problems for which no optimal solution nor value exists.

Definition 1.1.3 ▶ Unboundedness

Consider a minimisation problem subject to constraint $\boldsymbol{x} \in S \subseteq \boldsymbol{R}^n$ whose objective function is $f(\boldsymbol{x})$. The objective value is said to be **unbounded** if for all $K \in \mathbf{R}$, there exists some $\boldsymbol{x} \in S$ such that $f(\boldsymbol{x}) < K$.

1.2 Unconstrained Non-linear Programs

To introduce the notion of an unconstrained NLP, we shall first define the openness of a set.

Definition 1.2.1 ▶ Open Set

Let $S \subseteq \boldsymbol{R}^n$ be a set. S is called **open** if for all $\boldsymbol{x} \in S$ there exists $\epsilon > 0$ such that the ball

$$B(\boldsymbol{x}, \epsilon) := \{\boldsymbol{y} \in \boldsymbol{R}^n : \ \|\boldsymbol{y} - \boldsymbol{x}\| < \epsilon\}$$

is a subset of S.

Definition 1.2.2 ▶ Unconstrained NLP

An **unconstrained** NLP is an NLP whose feasible set $\mathcal{X}$ is an **open** subset of $\mathbf{R}^n$.

1.3 Constrained Non-linear Programs

Similarly, to introduce the notion of a constrained NLP, we shall first define the closed-ness of a set.

Definition 1.3.1 ▶ Closed Set

Let $S \subseteq \boldsymbol{R}^n$ be a non-empty set. S is said to be **closed** if for all convergent sequences $\{\boldsymbol{x}_i\}_{i=1}^{\infty}$ with $\boldsymbol{x}_i \in S$ for $i = 1, 2, \cdots$, the limit $\lim_{i \to \infty} \boldsymbol{x}_i \in S$.

The empty set and Euclidean spaces $\mathbb{R}^n$ are both open and closed.

> *Remark.* Note that a set which is not open may not necessarily be closed. However, a set is open if and only if its complement is closed.

Theorem 1.3.2 ▶ Intersection of Closed Sets

If C_1 and C_2 are both closed, then $C_1 \cap C_2$ is closed.

Proof. The case where $C_1 \cap C_2 = \varnothing$ is trivial. If $C_1 \cap C_2 \neq \varnothing$, let $\{\boldsymbol{x}_i\}_{i=1}^{\infty}$ be an arbitrary convergent sequence in $C_1 \cap C_2$. Since $\{\boldsymbol{x}_i\}_{i=1}^{\infty} \in C_1$ which is closed, we have $\lim_{i\to\infty} \boldsymbol{x}_i \in C_1$. Similarly, $\lim_{i\to\infty} \boldsymbol{x}_i \in C_2$. Therefore, $\lim_{i\to\infty} \boldsymbol{x}_i \in C_1 \cap C_2$.

Therefore, $C_1 \cap C_2$ is closed. □

We then follow up by introducing three important closed sets.

Theorem 1.3.3

Let $g : \mathbf{R}^n \to \mathbf{R}$ be a continuous function, then the sets

$$
\begin{aligned}
S_1 &= \{\boldsymbol{x} \in \mathbf{R}^n : g(\boldsymbol{x}) \leq 0\},\\
S_2 &= \{\boldsymbol{x} \in \mathbf{R}^n : g(\boldsymbol{x}) \geq 0\},\\
S_3 &= \{\boldsymbol{x} \in \mathbf{R}^n : g(\boldsymbol{x}) = 0\}
\end{aligned}
$$

are closed.

Proof. Consider S_1. Let $\{\boldsymbol{x}_i\}_{i=1}^{\infty}$ be any convergent sequence with $\boldsymbol{x}_i \in S_1$ for $i = 1, 2, \cdots$, then

$$
g\left(\lim_{i\to\infty} \boldsymbol{x}_i\right) \leq 0
$$

since $\boldsymbol{x}_i \leq 0$. Therefore, $\lim_{i\to\infty} \boldsymbol{x}_i \in S_1$ and so S_1 is closed.
S_2 and S_3 can be proved similarly. □

By Theorem 1.3.2, we know that $S_1 \cup S_2 \cup S_3$ is closed, which motivates the following definition:

Definition 1.3.4 ▶ Constrained NLP

A **constrained** NLP is an NLP whose feasible set

$$
S := \left\{\boldsymbol{x} \in \mathbf{R}^n : g_i(\boldsymbol{x}) = 0, i = 1, 2, \cdots, p, h_j(\boldsymbol{x}) \leq 0, j = 1, 2, \cdots, q\right\}
$$

is **closed**, where each of the g_i's is known as an equality constraint and each of the h_j's is known as an inequality constraint.

2

Convex Functions

2.1 Convex Sets and Functions

Definition 2.1.1 ▸ Convex Set

A set $D \subseteq \mathbf{R}^n$ is said to be **convex** if for all $\boldsymbol{x}, \boldsymbol{y} \in D$ and for all $\lambda \in [0, 1]$,

$$\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y} \in D.$$

Definition 2.1.2 ▸ Convex Function

A function $f : D \to \mathbf{R}^n$ is said to be **convex** if for all $\boldsymbol{x}, \boldsymbol{y} \in D$ and for all $\lambda \in [0, 1]$,

$$f(\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y}) \leq \lambda f(\boldsymbol{x}) + (1 - \lambda) f(\boldsymbol{y}).$$

Definition 2.1.3 ▸ Concave Function

A function $f : D \to \mathbf{R}^n$ is said to be **concave** if for all $\boldsymbol{x}, \boldsymbol{y} \in D$ and for all $\lambda \in [0, 1]$,

$$f(\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y}) \geq \lambda f(\boldsymbol{x}) + (1 - \lambda) f(\boldsymbol{y}).$$

Remark. A function which is not convex must be concave. However, a function which is convex may not be non-concave (consider $f(x) = x$).

Definition 2.1.4 ▸ Epigraph

Let $f : D \to \mathbf{R}$ be a function over a convex set $D \subseteq \mathbf{R}^n$. The **epigraph** of f is the set $E_f \subseteq \mathbf{R}^{n+1}$ defined by

$$E_f := \{(\boldsymbol{x}, \alpha) : \boldsymbol{x} \in D, \alpha \in \mathbf{R}, f(\boldsymbol{x}) \leq \alpha\}.$$

Theorem 2.1.5 ▸ Convexity of Epigraph

Let $f : D \to \mathbf{R}$ be a function over a convex set $D \subseteq \mathbf{R}^n$. The epigraph E_f is convex if and only if f is convex.

Theorem 2.1.6 ▶ Generalised Convex Combination

Let $f : S \to \mathbf{R}$ be a convex function on the convex set $S \subseteq \mathbf{R}^n$ and let $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k \in S$, then

$$f\left(\sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i\right) \leq \sum_{i=1}^{k} \lambda_i f(\boldsymbol{x}_i),$$

where $\sum_{i=1}^{k} \lambda_i = 1$ and $\lambda_i \geq 0$ for $i = 1, 2, \cdots, k$.

Theorem 2.1.7 ▶ An Alternative Expression for Directional Derivatives

Let f be a function over $D \subseteq \mathbf{R}^n$ and let $\boldsymbol{d} \in \mathbf{R}^n$ be non-zero, then

$$\nabla f(\boldsymbol{x})^{\mathrm{T}} \boldsymbol{d} = \lim_{\lambda \to 0} \frac{f(\boldsymbol{x} + \lambda \boldsymbol{d}) - f(\boldsymbol{x})}{\lambda}.$$

Theorem 2.1.8 ▶ Tangent Plane Characterisation of Convex Functions

Let f be a function over an open convex set $S \subseteq \mathbf{R}^n$ with continuous first partial derivatives, then f is convex if and only if

$$f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{y} - \boldsymbol{x}) \leq f(\boldsymbol{y})$$

for all $\boldsymbol{x}, \boldsymbol{y} \in S$. In particular, f is strictly convex if and only if the above inequality is strict.

Theorem 2.1.9 ▶ Global Minimiser of Convex Functions

Let $f : C \to \mathbf{R}$ be a convex and continuously differentiable function over a convex set $C \subseteq \mathbf{R}^n$. Then $\boldsymbol{x}^* \in C$ is a global minimiser for the minimisation problem

$$\min \{f(\boldsymbol{x}) : \boldsymbol{x} \in C\}$$

if and only if

$$\nabla f(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^*) \geq 0$$

for all $\boldsymbol{x} \in C$.