BASICS OF SET-CONSTRAINED AND UNCONSTRAINED OPTIMIZATION

1. Introduction

In this chapter we consider the optimization problem

$$\begin{gathered} \text{minimize } f(x) \\ \text{subject to } x \in \Omega \end{gathered}$$

The function $f : R^n \to R$ that we wish to minimize is a real-valued function called the *objective function or the cost function.* The vector x is an n-vector of independent variables : $x = [x_1, x_2, \ldots, x_n]^T \in R^n$. The variables $x_1, \ldots, x_n$ are often referred to as *decision variables.* The set Ω is a subset of R^n called the *constraint set or feasible set.*

The optimization problem above can be viewed as a decision problem that involves finding the "best" vector x of the decision variables over all possible vectors in Ω. By the "best" vector we mean the one that results in the smallest value of the objective function. This vector is called the *minimizer* of f over Ω. It is possible that there may be many minimizers. In this case, findng any of the minimizers will suffice.

There are also optimization problems that require maximization of the objective function, in which case we seek *maximizers.* Minimizers and maximizers are also called *extremizers.* Maximization problems, however, can be represented equivalently in the minimization form above because maximizing f is equivalent to minimizing $-f$. Therefore, we can confine our attention to minimization problems without loss of generality.

The problem above is a general form of a *constrained optimization problem*, because the decision variables are constrained to be in the constraint Ω. If $\Omega = R^n$, then we refer to the problem as an *unconstrained optimization problem.* In this chapter we discuss the unconstrained case.

The constraint "$x \in \Omega$" is called a *set constraint.* Often the constraint set Ω takes the form $\Omega = \{x : h(x) = 0, g(x) \leq 0\}$, where h and g are given functions. We refer to such constraints as *functional constraints. The case where* $\Omega = R^n$ is called the unconstrained case.

Definition 6.1 : Suppose that $f: R^n \to R$ is a real-valued function defined on some set $\Omega \subset R^n$. A point $x^* \in \Omega$ is a *local minimizer of f over* Ω if there exists $\epsilon > 0$ such that $f(x) \geq f(x^*)$ for all $x \in \Omega \setminus \{x^*\}$ and $\| x - x* \| < \epsilon$. A point $x^* \in \Omega$ is a global minimizer of f over Ω if $f(x) \geq f(x^*)$ for all $x \in \Omega \setminus \{x^*\}$.

If in the definitions above we replace " $\geq$ " $with$ " $>$ ", then we have a strict *local minimizer and a strict global minimizer.*

If x^* is a global minimizer of f over Ω, we write
$f(x^*) = min_{x \in \Omega} f(x)$ *and* $x^* = argmin_{x \in \Omega} f(x)$.
If the minimization is unconstrained, we simply write
$x^* = argmin_{x \in \Omega} f(x)$ *or* $x^* = argmin\ f(x)$.
In other words, given a real-valued function f, the notation arg min $f(x)$ denotes the *argument* that minimizes the function f(a point in the domain f), assuming that such a points is unique.

2. Conditions for Local Minimizers