Chapter 49

General Results of Optimization Theory

This chapter is devoted to some general results of optimization theory. A main theme is to find sufficient conditions that ensure that an objective function has a minimum which is achieved. We define the notion of a coercive function. The most general result is Theorem 49.2, which applies to a coercive convex function on a convex subset of a separable Hilbert space. In the special case of a coercive quadratic functional, we obtain the Lions–Stampacchia theorem (Theorem 49.6), and the Lax–Milgram theorem (Theorem 49.7). We define elliptic functionals, which generalize quadratic functions defined by symmetric positive definite matrices. We define gradient descent methods, and discuss their convergence. A gradient descent method looks for a descent direction and a stepsize parameter, which is obtained either using an exact line search or a backtracking line search. A popular technique to find the search direction is steepest descent. In addition to steepest descent for the Euclidean norm, we discuss steepest descent for an arbitrary norm. We also consider a special case of steepest descent, Newton's method. This method converges faster than the other gradient descent methods, but it is quite expensive since it requires computing and storing Hessians. We also present the method of conjugate gradients and prove its correctness. We briefly discuss the method of gradient projection and the penalty method in the case of constrained optima.

49.1 Optimization Problems; Basic Terminology

The main goal of *optimization theory* is to construct *algorithms* to find solutions (often approximate) of problems of the form

$$\begin{aligned} &\text{find} \quad u \\ &\text{such that} \;\; u \in U \;\text{ and }\; J(u) = \inf_{v \in U} J(v), \end{aligned}$$

where U is a given subset of a (real) vector space V (possibly infinite dimensional) and $J\colon \Omega \to \mathbb{R}$ is a function defined on some open subset Ω of V such that $U \subseteq \Omega$.