Chapter 51

Subgradients and Subdifferentials of Convex Functions $\circledast$

In this chapter we consider some deeper aspects of the theory of convex functions that are not necessarily differentiable at every point of their domain. Some substitute for the gradient is needed. Fortunately, for convex functions, there is such a notion, namely subgradients. Geometrically, given a (proper) convex function f, the subgradients at x are vectors normal to supporting hyperplanes to the epigraph of the function at $(x, f(x))$. The subdifferential $\partial f(x)$ to f at x is the set of all subgradients at x. A crucial property is that f is differentiable at x iff $\partial f(x) = \{\nabla f_x\}$, where ∇f_x is the gradient of f at x. Another important property is that a (proper) convex function f attains its minimum at x iff $0 \in \partial f(x)$. A major motivation for developing this more sophisticated theory of "differentiation" of convex functions is to extend the Lagrangian framework to convex functions that are not necessarily differentiable.

Experience shows that the applicability of convex optimization is significantly increased by considering extended real-valued functions, namely functions $f\colon S \to \mathbb{R} \cup \{-\infty, +\infty\}$, where S is some subset of $\mathbb{R}^n$ (usually convex). This is reminiscent of what happens in measure theory, where it is natural to consider functions that take the value $+\infty$. We already encountered functions that take the value $-\infty$ as a result of a minimization that does not converge. For example, if $J(u, v) = u$, and we have the affine constraint $v = 0$, for any fixed λ, the minimization problem

$$\begin{aligned} &\text{minimize} && u + \lambda v \\ &\text{subject to} && v = 0, \end{aligned}$$

yields the solution $u = -\infty$ and $v = 0$.

Until now, we chose not to consider functions taking the value $-\infty$, and instead we considered partial functions, but it turns out to be convenient to admit functions taking the value $-\infty$.

Allowing functions to take the value $+\infty$ is also a convenient alternative to dealing with partial functions. This situation is well illustrated by the indicator function of a convex set.