51.6 Generalization of the Lagrangian Framework

Essentially all the results presented in Section 50.3, Section 50.7, Section 50.8, and Section 50.9 about Lagrangians and Lagrangian duality generalize to programs involving a proper and convex objective function J, proper and convex inequality constraints, and affine equality constraints. The extra generality is that it is no longer assumed that these functions are differentiable. This theory is thoroughly discussed in Part VI, Section 28, of Rockafellar [138], for programs called ordinary convex programs. We do not have the space to even sketch this theory but we will spell out some of the key results.

We will be dealing with programs consisting of an objective function $J\colon \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ which is convex and proper, subject to $m \geq 0$ inequality contraints $\varphi_i(v) \leq 0$, and $p \geq 0$ affine equality constraints $\psi_j(v) = 0$. The constraint functions φ_i are also convex and proper, and we assume that

$$\mathbf{relint}(\mathrm{dom}(J)) \subseteq \mathbf{relint}(\mathrm{dom}(\varphi_i)), \qquad \mathrm{dom}(J) \subseteq \mathrm{dom}(\varphi_i), \quad i = 1, \ldots, m.$$

Such programs are *called ordinary convex programs*. Let

$$U = \mathrm{dom}(J) \cap \{v \in \mathbb{R}^n \mid \varphi_i(v) \leq 0,\ \psi_j(v) = 0,\ 1 \leq i \leq m,\ 1 \leq j \leq p\},$$

be the set of *feasible solutions*. We are seeking elements in $u \in U$ that minimize J over U.

A generalized version of Theorem 50.18 holds under the above hypotheses on J and the constraints φ_i and ψ_j, except that in the KKT conditions, the equation involving gradients must be replaced by the following condition involving subdifferentials:

$$0 \in \partial\left(J + \sum_{i=1}^m \lambda_i\varphi_i + \sum_{j=1}^p \mu_j\psi_j\right)(u),$$

with $\lambda_i \geq 0$ for $i = 1, \ldots, m$ and $\mu_j \in \mathbb{R}$ for $j = 1, \ldots, p$ (where $u \in U$ and J attains its minimum at u).

The *Lagrangian* $L(v, \lambda, \nu)$ of our problem is defined as follows: Let

$$E_m = \{x \in \mathbb{R}^{m+p} \mid x_i \geq 0,\ 1 \leq i \leq m\}.$$

Then

$$L(v, \lambda, \mu) = \begin{cases} J(v) + \sum_{i=1}^m \lambda_i\varphi_i(v) + \sum_{j=1}^p \mu_j\psi_j(v) & \text{if } (\lambda, \mu) \in E_m,\ v \in \mathrm{dom}(J) \\ -\infty & \text{if } (\lambda, \mu) \notin E_m,\ v \in \mathrm{dom}(J) \\ +\infty & \text{if } v \notin \mathrm{dom}(J). \end{cases}$$

For *fixed values* $(\lambda, \mu) \in \mathbb{R}^m_+ \times \mathbb{R}^p$, we also define the function $h\colon \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ given by

$$h(x) = J(x) + \sum_{i=1}^m \lambda_i\varphi_i(x) + \sum_{j=1}^p \mu_j\psi_j(x),$$