52.2 Augmented Lagrangians and the Method of Multipliers

In order to make the minimization step of the dual ascent method more robust, one can use the trick of adding the penalty term $(\rho/2)\left\|Au-b\right\|_2^2$ to the Lagrangian.

Definition 52.1. Given the Optimization Problem (P),

$$\begin{aligned} &\text{minimize} && J(u) \\ &\text{subject to} && Au = b, \end{aligned}$$

the *augmented Lagrangian* is given by

$$L_\rho(u,\lambda) = J(u) + \lambda^\top(Au-b) + (\rho/2)\left\|Au-b\right\|_2^2,$$

with $\lambda \in \mathbb{R}^m$, and where $\rho > 0$ is called the *penalty parameter*.

The augmented Lagrangian $L_\rho(u,\lambda)$ can be viewed as the ordinary Lagrangian of the Minimization Problem (P_ρ),

$$\begin{aligned} &\text{minimize} && J(u) + (\rho/2)\left\|Au-b\right\|_2^2 \\ &\text{subject to} && Au = b. \end{aligned}$$

The above problem is equivalent to Program (P), since for any feasible solution of (P_ρ), we must have $Au - b = 0$.

The benefit of adding the penalty term $(\rho/2)\left\|Au-b\right\|_2^2$ is that by Proposition 51.37, Problem (P_ρ) has a unique optimal solution under mild conditions on A.

Dual ascent applied to the dual of (P_ρ) yields the the *method of multipliers*, which consists of the following steps, given some initial λ^0:

$$\begin{aligned} u^{k+1} &= \arg\min_u L_\rho(u,\lambda^k) \\ \lambda^{k+1} &= \lambda^k + \rho(Au^{k+1} - b). \end{aligned}$$

Observe that the second step uses the parameter ρ. The reason is that it can be shown that choosing $\alpha^k = \rho$ guarantees that (u^{k+1},λ^{k+1}) satisfies the equation

$$\nabla J_{u^{k+1}} + A^\top \lambda^{k+1} = 0,$$

which means that (u^{k+1},λ^{k+1}) is dual feasible; see Boyd, Parikh, Chu, Peleato and Eckstein [28], Section 2.3.

Example 52.2. Consider the minimization problem

$$\begin{aligned} &\text{minimize} && y^2 + 2x \\ &\text{subject to} && 2x - y = 0. \end{aligned}$$