The above suggests that a reasonable termination criterion is that $\left\|r^k\right\|$ and $\left\|s^k\right\|$ should be small, namely that

$$\left\|r^k\right\| \leq \epsilon^{\text{pri}} \quad \text{and} \quad \left\|s^k\right\| \leq \epsilon^{\text{dual}},$$

for some chosen feasibility tolerances ϵ^{pri} and ϵ^{dual}. Further discussion for choosing these parameters can be found in Boyd et al. [28] (Section 3.3.1).

Various extensions and variations of ADMM are discussed in Boyd et al. [28] (Section 3.4). In order to accelerate convergence of the method, one may choose a different ρ at each step (say ρ^k), although proving the convergence of such a method may be difficult. If we assume that ρ^k becomes constant after a number of iterations, then the proof that we gave still applies. A simple scheme is this:

$$\rho^{k+1} = \begin{cases} \tau^{\text{incr}}\rho^k & \text{if } \left\|r^k\right\| > \mu \left\|s^k\right\| \\ \rho^k/\tau^{\text{decr}} & \text{if } \left\|s^k\right\| > \mu \left\|r^k\right\| \\ \rho^k & \text{otherwise,} \end{cases}$$

where $\tau^{\text{incr}} > 1, \tau^{\text{decr}} > 1$, and $\mu > 1$ are some chosen parameters. Again, we refer the interested reader to Boyd et al. [28] (Section 3.4).

52.6 Some Applications of ADMM

Structure in f, g, A, and B can often be exploited to yield more efficient methods for performing the x-update and the z-update. We focus on the x-update, but the discussion applies just as well to the z-update. Since z and λ are held constant during minimization over x, it is more convenient to use the scaled form of ADMM. Recall that

$$x^{k+1} = \arg\min_x \left(f(x) + (\rho/2) \left\|Ax + Bz^k - c + u^k\right\|_2^2 \right)$$

(here we use u instead of μ), so we can express the x-update step as

$$x^+ = \arg\min_x \left(f(x) + (\rho/2) \left\|Ax - v\right\|_2^2 \right),$$

with $v = -Bz^k + c - u^k$.

Example 52.7. A first simplification arises when $A = I$, in which case the x-update is

$$x^+ = \arg\min_x \left(f(x) + (\rho/2) \left\|x - v\right\|_2^2 \right) = \mathbf{prox}_{f,\rho}(v).$$

The map $v \mapsto \mathbf{prox}_{f,\rho}(v)$ is known as the *proximity operator of f with penalty ρ*. The above minimization is generally referred to as *proximal minimization*.