then we have

$$\begin{aligned}\lambda^\top r + (\rho/2)\left\|r\right\|_2^2 &= (\rho/2)\left\|r + (1/\rho)\lambda\right\|_2^2 - (1/(2\rho))\left\|\lambda\right\|_2^2 \\ &= (\rho/2)\left\|r + \mu\right\|_2^2 - (\rho/2)\left\|\mu\right\|_2^2.\end{aligned}$$

The *scaled form of ADMM* consists of the following steps:

$$\begin{aligned}x^{k+1} &= \arg\min_x \left(f(x) + (\rho/2)\left\|Ax + Bz^k - c + \mu^k\right\|_2^2\right) \\ z^{k+1} &= \arg\min_z \left(g(z) + (\rho/2)\left\|Ax^{k+1} + Bz - c + \mu^k\right\|_2^2\right) \\ \mu^{k+1} &= \mu^k + Ax^{k+1} + Bz^{k+1} - c.\end{aligned}$$

If we define the *residual $r^k$ at step $k$* as

$$r^k = Ax^k + Bz^k - c = \mu^k - \mu^{k-1} = (1/\rho)(\lambda^k - \lambda^{k-1}),$$

then we see that

$$r = u^0 + \sum_{j=1}^k r^j.$$

The formulae in the scaled form are often shorter than the formulae in the unscaled form.

We now discuss the convergence of ADMM.

## 52.4 Convergence of ADMM ⊛

Let us repeat the steps of ADMM: Given some initial $(z^0, \lambda^0)$, do:

$$x^{k+1} = \arg\min_x L_\rho(x, z^k, \lambda^k) \qquad (x\text{-update})$$

$$z^{k+1} = \arg\min_z L_\rho(x^{k+1}, z, \lambda^k) \qquad (z\text{-update})$$

$$\lambda^{k+1} = \lambda^k + \rho(Ax^{k+1} + Bz^{k+1} - c). \qquad (\lambda\text{-update})$$

The convergence of ADMM can be proven under the following three assumptions:

(1) The functions $f\colon \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ and $g\colon \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ are proper and closed convex functions (see Section 51.1) such that $\mathbf{relint}(\mathrm{dom}(f)) \cap \mathbf{relint}(\mathrm{dom}(g)) \neq \emptyset$.

(2) The $n \times n$ matrix $A^\top A$ is invertible and the $m \times m$ matrix $B^\top B$ is invertible. Equivalently, the $p \times n$ matrix $A$ has rank $n$ and the $p \times m$ matrix has rank $m$.

(3) The unaugmented Lagrangian $L_0(x, z, \lambda) = f(x) + g(z) + \lambda^\top(Ax + Bz - c)$ has a saddle point, which means there exists $x^*, z^*, \lambda^*$ (not necessarily unique) such that

$$L_0(x^*, z^*, \lambda) \leq L_0(x^*, z^*, \lambda^*) \leq L_0(x, z, \lambda^*)$$

for all $x, z, \lambda$.