which is equivalent to the original problem.

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) is called the *method of multipliers* and is discussed in Section 52.2.

The alternating direction method of multipliers, for short ADMM, combines the decomposability of dual ascent with the superior convergence properties of the method of multipliers. The idea is to split the function J into two independent parts, as $J(x,z) = f(x) + g(z)$, and to consider the Minimization Problem (P_{admm}),

$$\begin{aligned} &\text{minimize} && f(x) + g(z) \\ &\text{subject to} && Ax + Bz = c, \end{aligned}$$

for some $p \times n$ matrix A, some $p \times m$ matrix B, and with $x \in \mathbb{R}^n$, $z \in \mathbb{R}^m$, and $c \in \mathbb{R}^p$. We also assume that f and g are convex. Further conditions will be added later.

As in the method of multipliers, we form the augmented Lagrangian

$$L_\rho(x,z,\lambda) = f(x) + g(z) + \lambda^\top(Ax + Bz - c) + (\rho/2)\left\|Ax + Bz - c\right\|_2^2,$$

with $\lambda \in \mathbb{R}^p$ and for some $\rho > 0$. The major difference with the method of multipliers is that instead of performing a minimization step jointly over x and z, ADMM *first performs an x-minimization step and then a z-minimization step*. Thus x and z are updated in an alternating or sequential fashion, which accounts for the term *alternating direction*. Because the Lagrangian is augmented, some mild conditions on A and B imply that these minimization steps are guaranteed to terminate. ADMM is presented in Section 52.3.

In Section 52.4 we prove the convergence of ADMM under the following 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^*, λ^* (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, λ.

By Theorem 51.41, Assumption (3) is equivalent to the fact that the KKT equations are satisfied by some triple (x^*, z^*, λ^*), namely

$$Ax^* + Bz^* - c = 0 \tag{$*$}$$