The ADMM procedure (in scaled form) is

$$\begin{aligned} x^{k+1} &= \arg\min_x \left(l(x) + (\rho/2) \left\| x - z^k + u^k \right\|_2^2 \right) \\ z^{k+1} &= S_{\tau/\rho}(x^{k+1} + u^k) \\ u^{k+1} &= u^k + x^{k+1} - z^{k+1}. \end{aligned}$$

The x-update is a proximal operator evaluation. In general, one needs to apply a numerical procedure to compute x^{k+1}, for example, a version of Newton's method. The special case where $l(x) = (1/2) \left\| Ax - b \right\|_2^2$ is particularly important.

(4) *Lasso regularization.*

This is the following minimization problem:

$$\text{minimize} \quad (1/2) \left\| Ax - b \right\|_2^2 + \tau \left\| x \right\|_1 .$$

This is a linear regression with the regularizing term $\tau \left\| x \right\|_1$ instead of $\tau \left\| x \right\|_2$, to encourage a sparse solution. This method was first proposed by Tibshirani around 1996, under the name *lasso*, which stands for "least absolute selection and shrinkage operator." This method is also known as *ℓ^1-regularized regression*, but this is not as cute as "lasso," which is used predominantly. This method is discussed extensively in Hastie, Tibshirani, and Wainwright [89].

The lasso minimization is converted to the following problem in ADMM form:

$$\begin{aligned} &\text{minimize} && \frac{1}{2} \left\| Ax - b \right\|_2^2 + \tau \left\| z \right\|_1 \\ &\text{subject to} && x - z = 0. \end{aligned}$$

Then the ADMM procedure (in scaled form) is

$$\begin{aligned} x^{k+1} &= (A^\top A + \rho I)^{-1}(A^\top b + \rho(z^k - u^k)) \\ z^{k+1} &= S_{\tau/\rho}(x^{k+1} + u^k) \\ u^{k+1} &= u^k + x^{k+1} - z^{k+1}. \end{aligned}$$

Since $\rho > 0$, the matrix $A^\top A + \rho I$ is symmetric positive definite. Note that the x-update looks like a *ridge regression step* (see Section 55.1).

There are various generalizations of lasso.

(5) *Generalized Lasso regularization.*

This is the following minimization problem:

$$\text{minimize} \quad (1/2) \left\| Ax - b \right\|_2^2 + \tau \left\| Fx \right\|_1 ,$$