HW 7

Evan Ott
UT EID: eao466

November 2, 2016

HW

Here, algorithms are based on §6.4 in (Boyd et al., 2011). For more information, see §3.3 and §5 in (Boyd and Vandenberghe, 2004).

The lasso is commonly written as

$$x^* = \arg\min_x \left(\frac{1}{2}\|Ax - b\|_2^2 + \lambda \|x\|_1 \right)$$

Similar to the last homework, we can split this up into the differentiable part and non-differentiable part, which gives rise to the ADMM form.

$$\begin{aligned} \text{minimize} \quad & \|Ax - b\|_2^2 + \lambda \|z\|_1 \\ \text{s.t.} \quad & x - z = 0 \end{aligned}$$

In other words,

$$\begin{aligned} f(x) &= \frac{1}{2}\|Ax - b\|_2^2 = \frac{1}{2}x^\top A^\top Ax - b^\top Ax + \frac{1}{2}y^\top y = \frac{1}{2}x^\top Px + q^\top x + r \\ P &= A^\top A \\ q &= -A^\top b \\ g(z) &= \lambda \|z\|_1 \\ Ix + (-I)z &= x - z = c = 0 \end{aligned}$$

So we can break up the minimization according to §4.2 to find the optimal value for x:

$$\begin{aligned} x^+ &= \arg\min_x \left(f(x) + \frac{\rho}{2} \|x - (z + 0 - u)\|_2^2 \right) \\ &= \arg\min_x \left(f(x) + \frac{\rho}{2} \|x - z + u\|_2^2 \right) \\ &= \arg\min_x \left(\frac{1}{2}x^\top Px + q^\top x + r + \frac{\rho}{2} \|x - z + u\|_2^2 \right) \\ &= \left(P + \rho I^\top I \right)^{-1} (\rho I v - q) \quad \text{[see equation (4.1) for quadratic functions]} \\ &= \left(A^\top A + \rho I \right)^{-1} \left(\rho(z - u) + A^\top b \right) \end{aligned}$$

Now, the update for z:

$$g(z) = \lambda \|z\|_1 = \sum_{i=1}^{p} \lambda |z_i| = \sum_{i=1}^{p} h(z_i)$$

so the objective for z is component separable, so we can minimize each element separately:

$$\begin{aligned} \text{minimize} \quad & g(z) \\ \text{s.t.} \quad & x - z = 0 \\ z_i^+ &= \arg\min_{z_i} \left(h(z_i) + \frac{\rho}{2}(x_i - z_i + u_i)^2 \right) \\ &= \arg\min_{z_i} \left(h(z_i) + \frac{\rho}{2}(z_i - (x_i + u_i))^2 \right) \\ v_i &= x_i + u_i \\ &= \arg\min_{z_i} \left(h(z_i) + \frac{\rho}{2}(z_i - v_i)^2 \right) \\ &= S_{\lambda/\rho}(v_i) \quad \text{[see §4.4.3]} \\ &= S_{\lambda/\rho}(x_i + u_i) \end{aligned}$$

Then finally for u,

$$u^+ = u + Ix + (-I)z - 0 = u + x - z \quad \text{[see equation (3.7)]}$$

TODO: refactor code to have “accelerated proximal gradient” and “ADMM” functions.

See R code on GitHub.