SDS 385: Exercises 8 - Spacial smoothing at scale

October 29, 2016

Professor Scott

Spencer Woody

Problem 1

Laplacian smoothing

Normalsize text

Problem 2

Graph fused lasso

Normalsize text

Problem 3

ADMM

We use ADMM to form Lasso estimates. Let A be our feature matrix, and x is our coefficient vector, and b is the response vector.

$$\min \frac{1}{2} \|Ax - b\|_2^2 + \lambda \|z\|_1 \tag{1}$$

$$\text{s.t. } x - z = 0 \tag{2}$$

The Lagrangian is

$$L_p(x, z, u) = \frac{1}{2} \|Ax - b\|_2^2 + \lambda \|z\|_1 + u^T(x - z) + \frac{\rho}{2} \|x - z\|_2^2 . \tag{3}$$

$$x^{k+1} = \arg\min_x L_p(x, z^k, u^k) \tag{4}$$

$$= \arg\min_x \left\{ \frac{1}{2} \|Ax - b\|_2^2 + \lambda \left\|z^k\right\|_1 + (u^k)^T(x - z) + \frac{\rho}{2} \left\|x - z^k\right\|_2^2 \right\} \tag{5}$$

We take the gradient of the objective function with respect to x and set it equal to 0,

$$\nabla_x L_p(x, z, u) = A^T(Ax^{k+1} - b) + u^k + \rho(x^{k+1} - z^k) \tag{6}$$

$$= (A^T A + \rho I)x^{k+1} - (A^T b + \rho z^k - u^k) = 0 \tag{7}$$

$$x^{k+1} = (A^T A + \rho I)^{-1}(A^T b + \rho z^k - u^k), \text{ let } v^k = u^k/\rho \tag{8}$$

$$x^{k+1} = (A^T A + \rho I)^{-1}(A^T b + \rho(z^k - v^k)) \tag{9}$$

$$z^{k+1} = \arg\min_z L_p(x^{k+1}, z, u^k) \tag{10}$$

$$= \arg\min_z \left\{ \frac{1}{2} \left\|Ax^{k+1} - b\right\|_2^2 + \lambda \|z\|_1 + (u^k)^T(x^{k+1} - z) + \frac{\rho}{2} \left\|x^{k+1} - z\right\|_2^2 \right\} \tag{11}$$

$$= \arg\min_z \left\{ \frac{\lambda}{\rho} \|z\|_1 + (v^k)^T(x^{k+1} - z) + \frac{1}{2} \left\|x^{k+1} - z\right\|_2^2 \right\} \tag{12}$$

$$= \arg\min_z \left\{ \frac{\lambda}{\rho} \|z\|_1 - (z - x^{k+1})^T v^k + \frac{1}{2} \left\|z - x^{k+1}\right\|_2^2 \right\} \tag{13}$$

$$= \operatorname*{prox}_{\gamma=1} \frac{\lambda}{\rho} \left\|w^k\right\| , w^k = x^{k+1} + v^k \tag{14}$$

Problem 3 continued on next page. . .

So z^{k+1} is subjected to the soft threshholding from the previous two exercises. Finally we update u^k, and, identically, v^k.

$$u^{k+1} = u^k + \rho(x^{k+1} - z^{k+1}) \tag{15}$$
$$\rho v^{k+1} = \rho v^k + \rho(x^{k+1} - z^{k+1}) \tag{16}$$
$$v^{k+1} = v^k + x^{k+1} - z^{k+1} \tag{17}$$

With $\lambda = 0.002420476$, the regular proximal gradient method takes 5887 iterations, the accelerated proximal gradient method takes 626 iterations, and the ADMM takes 146 iterations. See Figure 1.

Figure 1: Comparison of convergence for different methods

R script for `myfuns.R`

R script for `e6.R`