CSE584 HOMEWORK 3

CHEN SUN
CXS1031@PSU.EDU

1. Question 1

$$\arg\min_x \frac{1}{2}\|Ax-t\|_2^2 \text{ such that } \|x\|_2 \le \lambda$$

Since $\|x\|_2 \le \lambda$ is a convex set.

Denote indicator function $I_c(x) = \begin{cases} 0 & \text{if } x \in C \\ \infty & \text{if } x \notin C \end{cases}$,where $C = \{x | \|x\|_2 \le \lambda\}$

Original problem is equal to:

$$\arg\min_x \frac{1}{2}\|Ax-t\|_2^2 + I_c(x)$$

Which is equal to:

$$\arg\min_x \frac{1}{2}\|Ax-t\|_2^2 + I_c(z) \text{ subject to } x = z$$

Using ADMM:

$$D(x,z,\alpha) = \frac{1}{2}\|Ax-t\|_2^2 + I_c(z) + \frac{\rho}{2}\|x-z+\alpha\|_2^2$$

To derive update step of ADMM:

$$\begin{aligned} x_{k+1} &= \arg\min_x \frac{1}{2}\|Ax-t\|_2^2 + \frac{\rho}{2}\|x-z_k+\alpha_k\|_2^2 \\ z_{k+1} &= \arg\min_z I_c(x) + \frac{\rho}{2}\|x_{k+1}-z+\alpha_k\|_2^2 \\ \alpha_{k+1} &= \alpha_k + (x_{k+1}-z_{k+1}) \end{aligned}$$

So we have the update step of ADMM:

$$\begin{aligned} x_{k+1} &= (A^TA+\rho I)^{-1}(A^Tt+\rho z_k-\rho\alpha_k) \\ z_{k+1} &= \Pi_C(x_{k+1}+\alpha_k) \\ \alpha_{k+1} &= \alpha_k + (x_{k+1}-z_{k+1}) \end{aligned}$$

2. Question 2

2.1. ADMM.

$$\begin{aligned} y &= \arg\min_x \frac{1}{2}\|Ax-t\|_2^2 + \lambda_1\|x\|_2 + \lambda_2\|x\|_1 \\ y &= \arg\min_x \frac{1}{2}\|Ax-t\|_2^2 + \lambda_1\|y\|_2 + \lambda_2\|z\|_1, \text{such that } x = z, x = y. \end{aligned}$$

Apply scalar version of ADMM and form the dual function:

$$D(x, y, z, \alpha, \beta) = \frac{1}{2}\|Ax - t\|_2^2 + \lambda_1\|y\|_2 + \lambda_2\|z\|_1 + \frac{\rho}{2}\|x - y + \alpha\|_2^2 + \frac{\rho}{2}\|x - z + \beta\|_2^2$$

The update of ADMM:

$$\begin{aligned}
x_{k+1} &= \arg\min_x \frac{1}{2}\|Ax - t\|_2^2 + \frac{\rho}{2}\|x - y_k + \alpha_k\|_2^2 + \frac{\rho}{2}\|x - z_k + \beta_k\|_2^2 \\
y_{k+1} &= \arg\min_y \lambda_1\|y\|_2 + \frac{\rho}{2}\|x_{k+1} - y + \alpha_k\|_2^2 \\
z_{k+1} &= \arg\min_z \lambda_2\|z\|_1 + \frac{\rho}{2}\|x_{k+1} - z + \beta_k\|_2^2 \\
\alpha_{k+1} &= \alpha_k + (x_{k+1} - y_{k+1}) \\
\beta_{k+1} &= \beta_k + (x_{k+1} - z_{k+1})
\end{aligned}$$

So we have the update of ADMM:

$$\begin{aligned}
x_{k+1} &= (A^TA + 2\rho I)^{-1}(A^Tt + \rho(y_k + z_k - \alpha_k - \beta_k)) \\
y_{k+1} &= \begin{cases} 0 & \text{if } \|r_k\|_2 \le \frac{\lambda_1}{\rho} \\ (1 - \dfrac{\lambda_1}{\rho\|r_k\|_2})r_k & \text{otherwise} \end{cases} \text{,where } r_k = x_{k+1} + \alpha_k \\
z_{k+1} &= S_{\lambda_2/\rho}(x_{k+1} + \beta_k) = \begin{cases} x_{k+1} + \beta_k - \frac{\lambda_2}{\rho}\ \mathrm{sign}(x_{k+1} + \beta_k) & \text{if } \|x_{k+1} + \beta_k\|_2 > \lambda_2 \\ 0 & \text{otherwise} \end{cases} \\
\alpha_{k+1} &= \alpha_k + (x_{k+1} - y_{k+1}) \\
\beta_{k+1} &= \beta_k + (x_{k+1} - z_{k+1})
\end{aligned}$$

2.2. **Convergence.** Since y and z condition are automatically maintained by ADMM.

So we need to monitor the following quantities:

(1) $\rho(y_{k+1} - y_k) + \rho(z_{k+1} - z_k)$ close to 0
(2) $x - y$ close to 0
(3) $x - z$ close to 0

2.3. **Sparsity properties.** Based on the ADMM solution.

Since l_1 norm introduce sparsity. The sparsity properties of this ADMM solution will be similar with the sparsity properties of solution of normal LASSO.

The sparsity introduced by l_1 norm is in the update step of z_{k+1}.