ECE 490: Introduction to Optimization **Fall 2018**

Solutions for Homework 5

1.

(a) We start by writing out the Lagrangian

$$L = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + \lambda(x_1 + x_2 + x_3 + x_4 + x_5 - 5)$$

Based on $\nabla_x L = 0$, we have

$$\begin{aligned} 2x_1 + \lambda &= 0 \\ 2x_2 + \lambda &= 0 \\ 2x_3 + \lambda &= 0 \\ 2x_4 + \lambda &= 0 \\ 2x_5 + \lambda &= 0 \end{aligned}$$

In addition, we have

$$\nabla_\lambda L = 0 \to x_1 + x_2 + x_3 + x_4 + x_5 = 5$$

We can solve the above equations and the only solution we get is

$$x_1^* = x_2^* = x_3^* = x_4^* = x_5^* = 1, \ \lambda^* = -2$$

Notice $\nabla_{xx} L(x^*, \lambda^*) = 2I > 0$ where I is the 5×5 identity matrix. Therefore, the above solution is the only local min for the problem.

(b) We can directly write out the Lagrangian as

$$L(x, y, \lambda) = \frac{1}{2}\|x\|^2 + \frac{1}{2}(y - b)^2 + \lambda(a^\mathsf{T} x - y)$$

If we fix λ, the minimization of L with respect to x and y is just a positive definite quadratic minimization problem. Hence we can directly set the derivatives of L to be zero and solve the global min. We fix λ and minimize L with respect to x and y as

$$\begin{aligned} \nabla_x L = 0 &\to x = -\lambda a \\ \nabla_y L = 0 &\to y = b + \lambda \end{aligned}$$

Therefore, the dual function is

$$\begin{aligned} D(\lambda) &= \frac{1}{2}\|a\|^2\lambda^2 + \frac{1}{2}\lambda^2 + \lambda(-b - \lambda - a^T a\lambda) \\ &= -\frac{1}{2}(1 + \|a\|^2)\lambda^2 - b\lambda \end{aligned}$$

2.

(a) ADMM updates x_{k+1} as follows:

$$\begin{aligned} x_{k+1} &= \arg\min_{x} L_\rho(x, y_k, \lambda_k) \\ &= \arg\min_{x:Ax=b} \left\{ \lambda_k^\mathsf{T}(x - y_k) + \frac{\rho}{2}\|x - y_k\|^2 \right\} \\ &= \arg\min_{x:Ax=b} \left\{ \frac{\rho}{2}\|x - y_k + \lambda_k/\rho\|^2 \right\} \end{aligned}$$

Therefore, we have

$$x_{k+1} = \mathbf{proj}_X \left(y_k - \frac{\lambda_k}{\rho} \right)$$

where X is the set $\{x : Ax = b\}$. Similarly, we can show

$$\begin{aligned} y_{k+1} &= S_{1/\rho} \left(x_{k+1} + \frac{\lambda_k}{\rho} \right) \\ \lambda_{k+1} &= \lambda_k + \rho(x_{k+1} - y_{k+1}) \end{aligned}$$

where $S_{1/\rho}$ is the shrinkage operator that shrinks every value between $-1/\rho$ and $1/\rho$ to 0.

(b) By definition, ADMM iterates as

$$\begin{aligned} x_{k+1}^i &= \arg\min_{x^i} \left\{ f_i(x^i) + (\lambda_k^i)^\mathsf{T}(x^i - y_k) + \frac{\rho}{2}\|x^i - y_k\|^2 \right\} \\ y_{k+1} &= \arg\min_{y} \left\{ \sum_{i=1}^n \left(-(\lambda_k^i)^\mathsf{T} y + \frac{\rho}{2}\|x^i - y\|^2 \right) \right\} \\ \lambda_{k+1}^i &= \lambda_k^i + \rho(x_{k+1}^i - y_{k+1}) \end{aligned}$$

Since $f_i(x^i) = \frac{1}{2}(a_i^\mathsf{T} x^i - b_i)^2$, we eventually have

$$\begin{aligned} x_{k+1}^i &= (a_i a_i^\mathsf{T} + \rho I)^{-1}(a_i b_i + \rho y_k - \lambda_k^i) \\ y_{k+1} &= \frac{1}{n} \sum_{i=1}^n (x_{k+1}^i + \lambda_k^i/\rho) \\ \lambda_{k+1}^i &= \lambda_k^i + \rho(x_{k+1}^i - y_{k+1}) \end{aligned}$$