Assignment 8 of MATP4820

Xinshi Wang 661975305

Problem 1

Consider the Least Absolute Deviation (LAD) problem:

$$\underset{\mathbf{x}\in\mathbb{R}^n}{\text{minimize}}\, \|\mathbf{A}\mathbf{x}-\mathbf{b}\|_1 \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$ are given. The LAD problem is equivalent to

$$\underset{\mathbf{x}\in\mathbb{R}^n,\mathbf{z}\in\mathbb{R}^m}{\text{minimize}}\, \|\mathbf{z}\|_1, \text{s.t. } \mathbf{A}\mathbf{x}-\mathbf{z}-\mathbf{b}=\mathbf{0}. \tag{2}$$

The augmented Lagrangian function of (2) is

$$\mathcal{L}_\rho(\mathbf{x},\mathbf{z},\mathbf{y}) = \|\mathbf{z}\|_1 + \mathbf{y}^\top(\mathbf{A}\mathbf{x}-\mathbf{z}-\mathbf{b}) + \frac{\rho}{2}\|\mathbf{A}\mathbf{x}-\mathbf{z}-\mathbf{b}\|^2,$$

where $\rho > 0$ is the penalty parameter, and $\mathbf{y} \in \mathbb{R}^m$ is the multiplier vector. The iterative update schemes of ADMM on solving (2) is

$$\mathbf{x}^{k+1} = \underset{\mathbf{x}\in\mathbb{R}^n}{\arg\min}\, \mathcal{L}_\rho(\mathbf{x},\mathbf{z}^k,\mathbf{y}^k), \tag{3a}$$

$$\mathbf{z}^{k+1} = \underset{\mathbf{z}\in\mathbb{R}^m}{\arg\min}\, \mathcal{L}_\rho(\mathbf{x}^{k+1},\mathbf{z},\mathbf{y}^k), \tag{3b}$$

$$\mathbf{y}^{k+1} = \mathbf{y}^k + \rho(\mathbf{A}\mathbf{x}^{k+1}-\mathbf{z}^{k+1}-\mathbf{b}) = 0 \tag{3c}$$

$$\tag{3d}$$

1. Assume $\mathbf{A}$ has column full rank, i.e., $\mathbf{A}^\top\mathbf{A}$ is nonsingular. Give the closed-form solution of $\mathbf{x}^{k+1}$ in the update (3a). The solution should involve the inverse of $\mathbf{A}^\top\mathbf{A}$.

$$
\begin{aligned}
x^{(k+1)} &= \underset{x}{argmin}||z||_1 + y^T(Ax - z - b) + \frac{\rho}{2}||Ax - z - b||_2^2 \\
&= \underset{x}{argmin} y^T(Ax - z - b) + \frac{\rho}{2}||Ax - z - b||_2^2 \\
&= \underset{x}{argmin}\frac{\rho}{2}||Ax - z - b + \frac{y}{\rho}||_2^2 \\
grad = 0 &\implies \rho A^T(Ax - z - b) \\
&\implies x^* = (A^TA)^-1A^T(z + b - \frac{y}{\rho})
\end{aligned}
$$

2. Recall that we have derived the proximal mapping of $\lambda\|\cdot\|_1$ for any $\lambda > 0$ in class. Let's denote it as $\mathbf{prox}_{\lambda\|\cdot\|_1}$. Use the proximal mapping to derive a closed form solution of $\mathbf{z}^{k+1}$ in the update (3b). In the solution, λ should be $1/\rho$.

$$
\begin{aligned}
z^{(k+1)} &= \underset{z}{argmin}||z||_1 + y^T(Ax - z - b) + \frac{\rho}{2}||Ax - z - b||_2^2 \\
&= \underset{z}{argmin}||z||_1 + \frac{\rho}{2}||Ax - z - b||_2^2 \\
&= \underset{z}{argmin}\frac{1}{\rho}||z||_1 + \frac{1}{2}||Ax - z - b + \frac{y}{\rho}||_2^2 \\
&= \mathbf{prox}_{\frac{1}{\rho}||\cdot||_1}(Ax - b + \frac{y}{\rho})
\end{aligned}
$$

3. Suppose $(\mathbf{x}^{k+1}, \mathbf{z}^{k+1}, \mathbf{y}^{k+1})$ is the output. At this point, give the violation of primal and dual feasibility.

   The violation of primal feasibility is

$$
||Ax^{k+1} - z^{k+1} - b||_2
$$

   . The violation of dual feasibility is

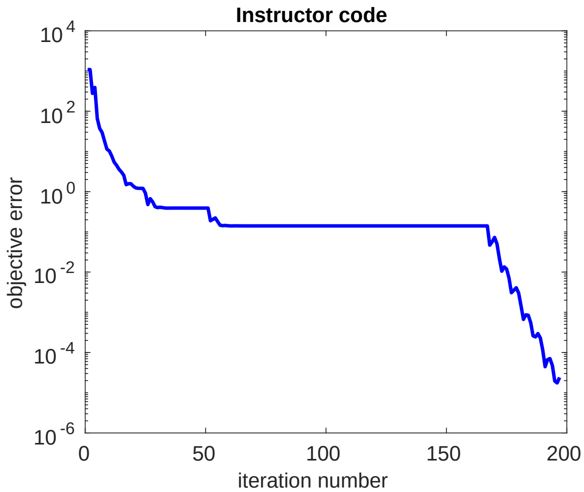
$$
|| - \rho A^T(z^{k+1} - z^k)||_2
$$

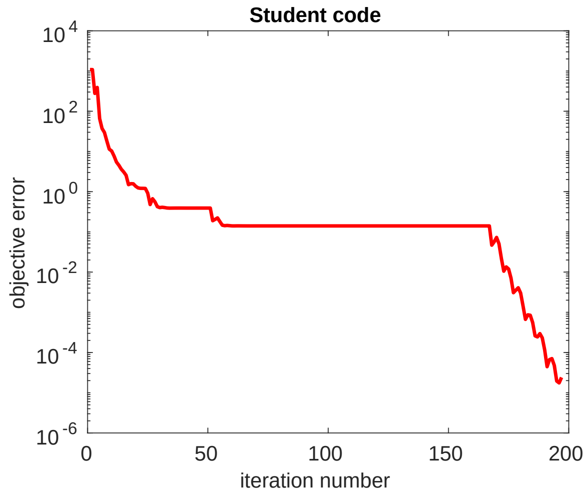
4. Use the instructor's provided file `ADMM_LAD.m` to write a Matlab function `ADMM_LAD` with input $\mathbf{A}$, $\mathbf{b}$, initial vector $\mathbf{x}0$, stopping tolerance `tol`, maximum number of iterations `maxit`, and penalty parameter $\rho > 0$. Also test your function by running the provided test file `test_ADMM_LAD.m` and compare to the instructor's function. Print your code and the results you get.

```
function [x,hist_obj, out] = ADMM_LAD(A,b,rho,tol,maxit,x0)
```

```
% alternating direction method of multipliers for
% min_x ||A*x-b||_1

[m,n] = size(A);

% initialization
x = x0;
z = A*x - b;

hist_obj = norm(z,1);

% initialize the multiplier
y = zeros(m,1);

pr = 0;
dr = 1;

iter = 1;
while max(pr,dr) > tol & iter < maxit
    iter = iter + 1;

    % update x
    x = inv(A'*A)*(A'*(z+b-(y/rho)));

    % update z
    z0 = z;

    z = sign(A*x-b+(y/rho)) .* max(abs(A*x-b+(y/rho)) - 1/rho, 0);
    % update y
    y = y+rho*(A*x-b-z);
    % compute primal and dual residual
    pr = norm(A*x-z-b);
    dr = norm(-rho*A'*(z-z0));

    % compute and save objective value ||A*x-b||_1
    obj = norm(A*x-b,1);
    hist_obj = [hist_obj; obj];
```

```
end
out.iter = iter;
out.pr = pr;
out.dr = dr;
end
```





Student solver: Total running time of instructor code is 0.8514

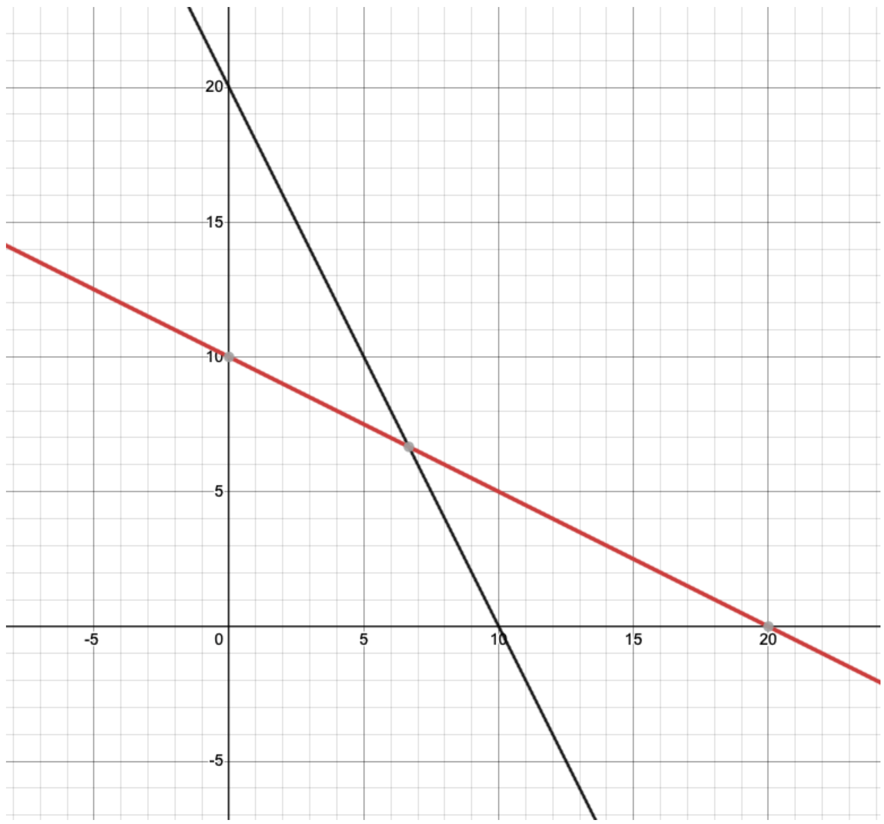
Instructor solver: Total running time of instructor code is 0.8210

Problem 2

Consider the linear program:

$$\begin{aligned} \underset{x_1, x_2}{\text{minimize}} \quad & -10x_1 - 12x_2 \\ \text{s.t.} \quad & x_1 + 2x_2 \leq 20 \\ & 2x_1 + x_2 \leq 20 \\ & x_1 \geq 0, x_2 \geq 0 \end{aligned} \tag{4}$$

1. Plot the feasible region of (4) and find the optimal solution by graph



   At point (0,0), the objective value is 0
   At point (0,10), the objective value is -120
   At point (10,0), the objective value is -100
   At point $(\frac{20}{3}, \frac{20}{3})$, the objective value is -146.67

   Therefore the optimal solution is $(\frac{20}{3}, \frac{20}{3})$ with objective value -146.67

2. In the lecture, we wrote (4) into an equivalent standard LP. Also, we start from a basic feasible solution and perform one step of the simplex method. Continue on the

basic feasible solution obtained in class and find the optimal solution by the simplex method.

Iteration 2:

$\beta = \{1,3\}, \eta = \{2,4\}, B = \begin{bmatrix}1 & 1\\ 2 & 0\end{bmatrix}, N = \begin{bmatrix}2 & 0\\ 1 & 1\end{bmatrix}, C_B = \begin{bmatrix}-10\\ 0\end{bmatrix}, C_N = \begin{bmatrix}-12\\ 0\end{bmatrix}, X_B = \begin{bmatrix}10\\ 10\end{bmatrix}$

$y = B^{-1}C_B = \begin{bmatrix}0\\ -10\end{bmatrix}$ $Z_N = C_N - N^T y = \begin{bmatrix}-2\\ 10\end{bmatrix}$

Select $Z_N(1) < 0$, then $q = \eta(1) = 2$

$W = B^{-1}A_q = \begin{bmatrix}0 & 0.5\\ 1 & -0.5\end{bmatrix}\begin{bmatrix}2\\ 1\end{bmatrix} = \begin{bmatrix}0.5\\ 1.5\end{bmatrix}$

$X_q^+ = \underset{i:w_i>0}{min}\frac{(X_B)_i}{w_i} = \frac{20}{3}, i_0 = 2, \rho = \beta(i_0) = 3$

Update: $X_B^+ = X_B - wx_q^+ = \begin{bmatrix}10\\ 10\end{bmatrix} - \frac{20}{3}\begin{bmatrix}0.5\\ 1.5\end{bmatrix} = \begin{bmatrix}\frac{20}{3}\\ 0\end{bmatrix}$

$X_N^+ = \begin{bmatrix}\frac{20}{3}\\ 0\end{bmatrix}$

$X^+ = \begin{bmatrix}\frac{20}{3}\\ \frac{20}{3}\\ 0\\ 0\end{bmatrix}$

$\beta^+ = \beta \cup \{2\} \setminus \{3\} = \{1,2\}$

Iteration 3:

$\beta = \{1,2\}, \eta = \{3,4\}, B = \begin{bmatrix}1 & 2\\ 2 & 1\end{bmatrix}, N = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, C_B = \begin{bmatrix}-10\\ -12\end{bmatrix}, C_N = \begin{bmatrix}0\\ 0\end{bmatrix}, X_B = \begin{bmatrix}\frac{20}{3}\\ \frac{20}{3}\end{bmatrix}$

$y = B^{-1}C_B = \begin{bmatrix}-\frac{14}{3}\\ -\frac{8}{3}\end{bmatrix}$ $Z_N = C_N - N^T y = \begin{bmatrix}\frac{14}{3}\\ \frac{8}{3}\end{bmatrix}$

All $Z_N(1) > 0$, thus it is optimal