Advanced Optimization
Homework 4

Winter 1400
Due date: Dey 22th

Analytical Questions

1. Consider the following two optimization problems. Derive the dual problem and find the optimal dual and primal variables. Then plot the cost vs constraint (as shown in lecture 21 page 4 and also figure 5.1.2 of the nonlinear programming book) and match the obtained Lagrange multipliers with the plots.

   (a) minimize x_1
   subject to $|x_1| + |x_2| \leq 1, \quad x \in \mathbb{R}^2$

   (b) minimize x_1
   subject to $|x_1| + |x_2| \leq 1, \quad x \in \{x \big| \, |x_1| \leq 1, |x_2| \leq 1\}$

Computer Questions

1. Use both the augmented Lagrange Quadratic penalty function and the Multiplier methods to solve the following optimization problem using $x_0 = [0.1 \ 0.2 \ 0.7]^T$ as your starting point. Compare their convergence rate and justify the difference.

$$\min_x x_1^2 + 2x_2^2 + 3x_3^2$$
$$\text{subject to } x_1 + x_2 + x_3 = 1$$

2. Linear SVM with soft margin uses the following cost function to find the maximum margin separating hyperplane.

$$\phi(W, b, \xi) = \min_{W,b,\xi} \frac{1}{2} \|W\| + C \sum_{i=1}^{n} \xi_i$$
$$\text{subject to } \begin{cases} y_i \left(W^T x_i + b\right) \geq 1 - \xi_i \\ \xi_i \geq 0 \end{cases} \quad for \ i = 1, 2, 3, ..., n$$

   Optimize a soft margin linear SVM to classify the MNIST data using **interior-point method**, considering what follows:

   - To reduce the amount of computation required, use only the train and test data associated with the digits of 0 and 1.
   - During the training procedure, only use the train data.
   - As implied from the equations, parameters of the problem are $\theta = \{W, b, \xi\}$, and the corresponding dimensions are $\theta = \{W : 784 \times 1, b : 1 \times 1, \xi_i : 1 \times 1\}$.
   - consider zero as the initial value for all parameters except for $\epsilon_0 = 200$ and $C = 1$.
   - update the ϵ every 5 steps with a proper coefficient.

- The CCR criteria for test data is shown below:

$$y_q^{'} = \begin{cases} +1 & w^T x_q + b \geq 0 \\ -1 & w^T x_q + b < 0 \end{cases}$$

$$CCR = \frac{\sum_{q=1}^{Q} \|y_q - y_q^{'}\|}{2Q}$$

- The report must contain the following elements:
  - A brief report of the modeling procedure.
  - Diagram of cost function value per iteration for the train data.
  - Value of CCR criteria on test data.