Question 1: ADMM (30 points)

Solve the following optimization problem using the scaled form of alternating direction method of multipliers (ADMM).

$$\min_x \frac{1}{2}x^T Px + q^T x + \frac{\lambda}{2}\|z\|_2^2 \qquad s.t. \qquad \begin{cases} x = z \\ a \le z \le b \end{cases}$$

Where $P \in \mathbb{R}^{n\times n}$ and, $a, b, x, q \in \mathbb{R}^n$.

Part1. Write the augmented Lagrangian function (the scaled form) and drive the ADMM updates (Show your work).

Part 2. “Question1.mat” dataset contains variables P, q, a and b. Set $\lambda = 0.5$ and $\rho = 1.1$. Plot the value of $\frac{1}{2}x^T Px + q^T x$ per iteration. Plot $\|x - z\|_2$ per iteration. Plot your final value of x.

Question 2: Coordinate Descent (35 points)

Part 1: Solve the following optimization problem using coordinate descent algorithm.

$$\min_w \frac{1}{2}\|y - Xw\|_2^2 + \frac{\lambda}{2}\|w\|_2^2$$

Where y, $w \in \mathbb{R}^n$ and $X \in \mathbb{R}^{n\times n}$. Consider $\lambda = 0.1$. Drive a closed form solution for w.

Part 2: Solve the following optimization problem using coordinate descent algorithm.

$$\min_w \frac{1}{2}\|y - Xw\|_2^2 + \lambda_1|w|_1 + \frac{\lambda_2}{2}\|w\|_2^2$$

Drive a closed form solution for w. Consider $\lambda_1 = 0.05$ and $\lambda_2 = 0.01$.

Part 3. In this part, you implement your own regression algorithm using your solution in part 1 and 2 to predict the performance decay over time of the Gas Turbine (GT) compressor. The range of decay of compressor has been sampled with a uniform grid of precision 0.001. The compressor decay coefficient is in the range of [0.95,1]. The dataset is provided as “Question2.csv”. The last column of the datasets corresponds to the output we want to predict.

- The 13 features are:

Lever position (lp)
Ship speed (v) [knots]
Gas Turbine (GT) shaft torque (GTT) [kN m]
GT rate of revolutions (GTn) [rpm]
Gas Generator rate of revolutions (GGn) [rpm]
Port Propeller Torque (Tp) [kN]
Hight Pressure (HP) Turbine exit temperature (T48) [C]
GT Compressor outlet air temperature (T2) [C]
HP Turbine exit pressure (P48) [bar]
GT Compressor outlet air pressure (P2) [bar]
GT exhaust gas pressure (Pexh) [bar]

Turbine Injection Control (TIC) [%]
Fuel flow (mf) [kg/s]
GT Compressor decay state coefficient

• Use the first 2000 samples to learn w_{opt}. Report your coefficients (w_{opt}) for each method. Plot the function $\|y - Xw\|_2^2$ versus iterations.

• Report your Sum of Absolute Errors on the test set (data samples 2001-2387) for each method.

Question 3: Proximal Gradient Descent (35 points)

Solve the following optimization problem using proximal gradient method.

$$f(\boldsymbol{\theta}) = \min_{\boldsymbol{\theta}\in\mathbb{R}^d} \frac{1}{m}\sum_{i=1}^{m}\left[\log\left(1 + \exp(\boldsymbol{x}_i\boldsymbol{\theta})\right) - y_i\boldsymbol{x}_i\boldsymbol{\theta}\right] + \frac{\lambda_2}{2}\|\boldsymbol{\theta}\|_2^2 + \lambda_1\|\boldsymbol{\theta}\|_1$$

Where $\log(\cdot)$ is the natural logarithm; $\boldsymbol{x}_i \in \mathbb{R}^{1\times d}$ is sample i; $\boldsymbol{\theta} \in \mathbb{R}^d$, $y_i \in \{0,1\}$ is label for sample i and $\lambda > 0$.

Decomposed $f(\boldsymbol{\theta})$ into a convex and differentiable function g and a convex but not differentiable function h.

Part 1: Derive the proximal gradient method updates for the objective function. (Show your work)

Part 2: "Question3.csv" contains 2000 data. There are two classes "0" and "1". The last column represents the labels. Use the first 1000 data samples to learn $\boldsymbol{\theta}$, and the rest to evaluate the accuracy of your classifier. Use your learned $\boldsymbol{\theta}$ to predict the labels on the test set (Predict labels $= \frac{1}{1+\exp(-\boldsymbol{x}\boldsymbol{\theta})}$). Set $\lambda_1 = 10$ and $\lambda_2 = 5$. Build your own classifier using your equations (part 1).

- Plot accuracy per iteration on the training set.
- Report $\|\boldsymbol{\theta}\|_0$.
- Report your classification accuracy on the test set.