Comp6211e Homework 1

Assignment date: Feb 28
Due date: March 13

Theoretical Problems (10 points)

1. (3 points) Consider convex set $C = \{x \in \mathbb{R}^d : \|x\|_1 \leq 1\}$. Given $y \in \mathbb{R}^d$, compute the projection $\text{proj}_C(y)$.

2. (3 points) Compute $\partial\|x\|_2$ at $x = 0$.

3. (4 points) Consider the Lasso method. Given $X \in \mathbb{R}^{n\times d}$ and $y \in \mathbb{R}^n$, we want to find $w \in \mathbb{R}^d$ to solve

$$[w_*, \xi_*] = \arg\min_{w,b,\xi} \left[\|Xw - y\|_2^2 + \lambda \sum_{j=1}^{d} \xi_j \right], \tag{1}$$

$$\text{subject to } \xi_j \geq w_j, \quad \xi_j \geq -w_j \quad (j = 1, \ldots, d). \tag{2}$$

   Lasso produces sparse solutions. Define the support of the solution as

$$S = \{j : w_{*,j} \neq 0\}.$$

   Write down the KKT conditions. Simplify in terms of $S, X_S, X_{\bar{S}}, y, w_S$. Here X_S contains the columns of X in S, $X_{\bar{S}}$ contains the columns of X not in S, and w_S contains the nonzero components of w_*.

Programming Problem (5 points)

We consider ridge regression problem with randomly generated data. The goal is to implement gradient descent and experiment with different strong-convexity settings and different learning rates.

- Use the python template "prob2_template.py", and implement functions marked with '# implement'.
- Submit your code and outputs. Compare to the theoretical convergence rates in class, and discuss your experimental results.