Homework 4

1. Please download training (`train.txt`) and test data sets (`test.txt`) from my dropbox link

   `https://www.dropbox.com/sh/6kfj7zw0t5sm8lh/AAAov8-xtdZRZpFfCFDEOPgZa?dl=0`

   that contains 50 predictors and 1 response variable (in the last column).

   You may use the following code after changing the file path to import the data.

```
> train <- matrix(scan("~/Downloads/train.txt"), 500, 51)
> test  <- matrix(scan("~/Downloads/test.txt"), 500, 51)
>
> x <- train[,-51]
> y <- train[,51]
>
> x.test <- test[,-51]
> y.test <- test[,51]
```

   (a) Write your own code for the pathwise decent algorithm for Elastic-Net-penlaized linear regression to select the best λ that minimizes the test prediction error, via grid search.

   (b) Report the non-zero regression coefficients estimates you obtained, i.e., $\hat{\beta}_j \forall j \in \{\hat{\beta}_j \neq 0\}$.

2. For group LASSO problem, please answer the following questions.

   (a) Suppose $f(\boldsymbol{\beta}) = \sqrt{\beta_1^2 + \beta_2^2}$ where $\boldsymbol{\beta} = (\beta_1, \beta_2)$. Graphically justify that the subdifferential of $f(\boldsymbol{\beta})$ at $\boldsymbol{\beta} = \mathbf{0}$ is "Any vector with $\|\boldsymbol{\beta}\|_2 \leq 1$."

   (b) Under the orthonormality condition $\mathbf{Z}_j^T \mathbf{Z}_j = \mathbf{I}$, show that

$$-\mathbf{Z}_j^T(\mathbf{r}_j - \mathbf{Z}_j\hat{\boldsymbol{\beta}}_j) + \lambda\hat{\mathbf{s}}_j = \mathbf{0}$$

   where

$$\hat{\boldsymbol{\beta}}_j = \left(1 - \frac{\lambda}{\|\mathbf{Z}_j^T\mathbf{r}_j\|_2}\right)_+ \mathbf{Z}_j^T\mathbf{r}_j,$$

   with $\mathbf{r}_j = \mathbf{y} - \sum_{k \neq j} \mathbf{Z}_k\boldsymbol{\beta}_k$ and $\hat{\mathbf{s}}_j$ is the subdifferential of $\|\boldsymbol{\beta}_j\|_2$ given on page 11 of Lecture 6.