STAT 7934: HW 4

Due at 6 pm by email: Wed April 22

Group question

(1) Consider the following regression problem,

$$\min_{\beta \in \mathbb{R}^p} \frac{1}{2}||y - X\beta||_2^2 + \lambda_1 \sum_{g=1}^{G} w_g ||\beta_g||_2,$$

where $g \in \{1, \cdots, G\}$ denotes the index for the G non-overlapping groups the predictors are organized in, $\{w_g\}_{g=1}^G$ weights that account for the varying group sizes and $||\beta_g||_2$ is the Euclidean norm (not squared) for group's g coefficient vector.

This procedure acts like a lasso penalty, but at the group level.

(a) Devise an algorithm to solve this problem.

em Hint: Write down the KKT and connect the result to those for the lasso. This will give you insights on how to write an algorithm.

(b) Generate data for a 1,000 by 100 design matrix X and generate the response y by adding Gaussian noise. Consider a model where the true coefcient vector β has 25 non-zero elements organized in 5 groups of size 5 and 75 zero elements organized in 15 groups of size 5.

(c) Apply your algorithm to estimate β. Discuss how you picked the tuning parameter λ_1.

(d) Now, consider the same coefficient vector β but it is now partitioned into 5 groups of sizes 10, 5, 5, 3, 2 respectively, while the remaining 75 zero coefficents are organized in 5 groups of size 25, 20, 20, 10, and 5 respectively. Apply your algorithm to estimate β. Again discuss how you selected the tuning parameter λ_1.

(e) The above algorithm does not achieve sparsity within a group. To do so, consider now the following modified objective function:

$$\min_{\beta \in \mathbb{R}^p} \frac{1}{2}||y - X\beta||_2^2 + \lambda_1 + \sum_{g=1}^{G} w_g ||\beta_g||_2 + \lambda_2 ||\beta||_1,$$

that is you add a standard lasso penalty.

Devise an algorithm to solve this new problem and apply it to following two scenarios.
Scenario 1: 20 groups of size 5, 5 groups with 4 non-zero elements and 1 zero element, and 15 groups with all zero elements.
Scenario 2: 5 groups of sizes 10, 5, 5, 3, 2 respectively with 8, 4, 4, 2, and non zero coefficients. 5 groups of size 25, 20, 20, 10, and 5 respectively with all zero coefficients.
Again discuss carefully how you selected the λ_1, λ_2 tuning parameters.