Project 10 Structured sparsity in genetics

Team K - Minghang Li

May 04, 2023

Problem 28: Uniqueness of predictions from the lasso

Proof. To prove that $\mathbf{X}\hat{\beta}_1 = \mathbf{X}\hat{\beta}_2$, it's equivalent to prove $\hat{\beta}_1 = \hat{\beta}_2$. We'll prove by contradiction.

Let the solution set be denoted as S, we have $\hat{\beta}_1, \hat{\beta}_2 \in S$.

Since S is convex, we also have $\alpha\hat{\beta}_1 + (1-\alpha)\hat{\beta}_2 \in S$.

So that we have:

$$\frac{1}{2}\left\|\mathbf{y} - \mathbf{X}\left(\alpha\hat{\beta}_1 + (1-\alpha)\hat{\beta}_2\right)\right\| + \lambda\left\|\alpha\hat{\beta}_1 + (1-\alpha)\hat{\beta}_2\right\| = c^* \tag{1}$$

Define the loss function as $f : S \mapsto \|\mathbf{y} - \mathbf{X}\beta\|_2^2$. We know that the lineare loss function is stirctly convex, so that we have:

$$f\left(\alpha\hat{\beta}_1 + (1-\alpha)\hat{\beta}_2\right) < \alpha f(\hat{\beta}_1) + (1-\alpha)f(\hat{\beta}_2) \tag{2}$$

Similarly, since l_1-nrom is convex, we have:

$$\left\|\alpha\hat{\beta}_1 + (1-\alpha)\hat{\beta}_2\right\|_1 \leq \alpha\|\hat{\beta}_1\| + (1-\alpha)\|\hat{\beta}_2\| \tag{3}$$

Combining Equation 2 and Equation 3, we can start the following derivation:

$$\begin{aligned}
\frac{1}{2}f\left(\alpha\hat{\beta}_1 + (1-\alpha)\hat{\beta}_2\right) + \lambda\left\|\alpha\hat{\beta}_1 + (1-\alpha)\hat{\beta}_2\right\| &< \alpha\left(f(\hat{\beta}_1) + \lambda\|\hat{\beta}_1\|\right) + (1-\alpha)\left(f(\hat{\beta}_2) + \lambda\|\hat{\beta}_2\|\right) \\
&= \alpha c^* + (1-\alpha)c^* \\
&= c^*
\end{aligned}$$

Just a re-write:

$$\frac{1}{2}f\left(\alpha\hat{\beta}_1 + (1-\alpha)\hat{\beta}_2\right) + \lambda\left\|\alpha\hat{\beta}_1 + (1-\alpha)\hat{\beta}_2\right\| < c^* \tag{4}$$

Equation 4 contradicts with Equation 1, which suggests that our assumption $\hat{\beta}_1 \neq \hat{\beta}_2$ is wrong.

□

Problem 29: Ridge regression solution

Problem 30: Variable selection under various norms