Ahmad Zareei
Econometrics 240A
Homework 2 - Part II

Problem 1. Let $\mathbf{Z} \sim \mathcal{N}(\theta, \sigma^2/NI_K)$.. Show that $\mathbb{E}[g(\mathbf{Z})(\mathbf{Z}-\theta)] = \frac{\sigma^2}{N}\mathbb{E}[\nabla_{\mathbf{Z}} g(\mathbf{Z})]$?

Solution. We have

$$E[g(\mathbf{Z})(\mathbf{Z}-\theta)] = \int (\frac{N}{2\pi\sigma^2})^{K/2} g(\mathbf{Z})(\mathbf{Z}-\theta) e^{-\frac{N}{2\sigma^2}(\mathbf{Z}-\theta)'(\mathbf{Z}-\theta)} \mathrm{d}\mathbf{Z} \tag{1}$$

We use integration by parts, and take $u = g(\mathbf{Z})$ and $v' = (\mathbf{Z}-\theta)\exp[-\frac{N}{2\sigma^2}(\mathbf{Z}-\theta)'(\mathbf{Z}-\theta)]$. Note that $u' = \nabla g(\mathbf{Z})$ and $v = \frac{-\sigma^2}{N}\exp[-\frac{N}{2\sigma^2}(\mathbf{Z}-\theta)'(\mathbf{Z}-\theta)]$. Therefore we obtain

$$\begin{aligned} E[g(\mathbf{Z})(\mathbf{Z}-\theta)] &= -g(\mathbf{Z})\frac{\sigma^2}{N}(\frac{N}{2\pi\sigma^2})^{K/2} e^{-\frac{N}{2\sigma^2}(\mathbf{Z}-\theta)'(\mathbf{Z}-\theta)} \\ &+ \frac{\sigma^2}{N}\int (\frac{N}{2\pi\sigma^2})^{K/2} \nabla g(\mathbf{Z}) e^{-\frac{N}{2\sigma^2}(\mathbf{Z}-\theta)'(\mathbf{Z}-\theta)} \mathrm{d}\mathbf{Z} \end{aligned} \tag{2}$$

provided that $g(\mathbf{Z})$ is bounded, then the first term becomes zero at $\mathbb{R}^K$, and we obtian

$$\begin{aligned} E[g(\mathbf{Z})(\mathbf{Z}-\theta)] &= \frac{\sigma^2}{N}\int (\frac{N}{2\pi\sigma^2})^{K/2} \nabla g(\mathbf{Z}) e^{-\frac{N}{2\sigma^2}(\mathbf{Z}-\theta)'(\mathbf{Z}-\theta)} \mathrm{d}\mathbf{Z} \\ &= \mathbb{E}[\nabla g(\mathbf{Z})] \end{aligned} \tag{3}$$

Problem 2. Consider the estimate of the Risk associated with the weakly differentiable estimate of $\hat{\theta}$ introduced in Theorem 2.5.1 of the main text:

$$\hat{R}(\mathbf{Z}) = K\sigma^2 + 2\sigma^2 \sum_{k=1}^{K} \frac{\partial g(\mathbf{Z})}{\partial Z_k} + \sum_{k=1}^{K} \left(\hat{\theta}_k - Z_k\right) \tag{4}$$

Prove that the risk estimate is unbiased under the sqaure error loss $\mathbb{E}_\theta[\hat{R}(\mathbf{Z})] = R(\hat{\theta}, \theta)$?

Solution. Basically, the risk here is $\mathbb{E}[\hat{R}_{SURE}(\mathbf{Z})]$ and we are asked to prove that $\mathbb{E}[\hat{R}_{SURE}(\mathbf{Z})] = \mathbb{E}[||\hat{\theta}-\theta||^2]$. We start from the RHS, we have

$$\begin{aligned} R(\hat{\theta}, \theta) = \mathbb{E}[||\hat{\theta}-\theta||^2] &= \mathbb{E}[||\hat{\theta} - \mathbf{Z} + \mathbf{Z} - \theta||^2] \\ &= \mathbb{E}[\sum_{k=1}^{K}(\hat{\theta}_k - Z_k)^2] + \mathbb{E}[||\mathbf{Z}-\theta||^2] + 2\mathbb{E}[(\hat{\theta}-\mathbf{Z})'(\mathbf{Z}-\theta)] \\ &= \mathbb{E}[\sum_{k=1}^{K}(\hat{\theta}_k - Z_k)^2] + K\sigma^2 + 2\mathbb{E}[(\hat{\theta}-\mathbf{Z})'(\mathbf{Z}-\theta)] \end{aligned}$$

Note that $g(\mathbf{Z}) = \hat{\theta} - \mathbf{Z}$. Using this equation and the Question 1 results, we obtian

$$\begin{aligned}
R(\hat{\theta}, \theta) &= \mathbb{E}[\sum_{k=1}^{K} (\hat{\theta}_k - Z_k)^2] + K\sigma^2 + 2\sum_{k=1}^{K} \mathbb{E}[g_k(\mathbf{Z})(Z_k - \theta_k)] \\
&= \mathbb{E}[\sum_{k=1}^{K} (\hat{\theta}_k - Z_k)^2] + K\sigma^2 + 2\sum_{k=1}^{K} \mathbb{E}[\frac{\partial g_k(\mathbf{Z})}{\partial Z_k}] \qquad \text{Using Question 1 Results} \\
&= \mathbb{E}[\hat{R}_{SURE}(\mathbf{Z})] \qquad (5)
\end{aligned}$$

Where it means that $R(\hat{\theta}, \theta) = \mathbb{E}[\hat{R}_{SURE}(\mathbf{Z})]$.

Problem 3. Let $\mathbf{Z} \sim \mathcal{N}(\theta, \sigma^2/NI_K)$ and consider the following soft threshhold of estimate of θ:

$$\hat{\theta}_k = \text{sgn}(Z_k)\left(|Z_k - \lambda\right)_+, \qquad k = 1, ..., K \qquad (6)$$

where it means that this estimator shrinks the MLE of θ_k toward zero when it is large (in absolute value) and shrinks it exactly to zero when it is small (in absolute value). Use Theorem 2.5.1 to show that

$$\hat{R}_{SURE}(\mathbf{Z}, \lambda) = \frac{K}{N}\sigma^2 - \frac{2\sigma^2}{N}\sum_{k=1}^{K} \mathbf{1}(|Z_k| < \lambda) + \sum_{k=1}^{K} \min(Z_k^2, \lambda^2) \qquad (7)$$

Provides a concrete prediction problem where you would expect the risk properties of the soft threshold estimator to be attractive.

Solution. We start by considering two different cases of $|Z_k| > \lambda$ and $|Z_k| < \lambda$.. We have

- $|Z_k| > \lambda$:

$$\begin{aligned}
\hat{\theta}_k &= \text{sgn}(Z_k)\left(|Z_k| - \lambda\right)_+ \\
&= \text{sgn}(Z_k)Z_k - \text{sgn}(Z_k)\lambda \\
&= Z_k - \text{sgn}(Z_k)\lambda \qquad (8)
\end{aligned}$$

  Therefore

$$(\hat{\theta}_k - Z_k)^2 = \lambda^2 \text{sgn}(Z_k)^2 = \lambda^2 \qquad (9)$$

  and $g(\mathbf{Z})$, then would become as

$$g_k(\mathbf{Z}) = \hat{\theta}_k - Z_k = -\text{sgn}(Z_k)\lambda \qquad (10)$$

- $|Z_k| < \lambda$:

$$\hat{\theta}_k = \text{sgn}(Z_k)\left(|Z_k| - \lambda\right)_+ = 0 \qquad (11)$$

Therefore

$$(\hat{\theta}_k - Z_k)^2 = Z_k^2 \tag{12}$$

and $g(\mathbf{Z})$, then would become as

$$g_k(\mathbf{Z}) = \hat{\theta}_k - Z_k = -Z_k \tag{13}$$

From these two cases, we can conclude that

$$(\hat{\theta}_k - Z_k)^2 = \min(\lambda^2, Z_k^2) \tag{14}$$

$$\frac{\partial g_k(\mathbf{Z})}{\partial Z_k} = -\mathbf{1}\{|Z_k| < \lambda\} \tag{15}$$

Pluggin these two equations into the result of Theorem 2.5.1, we obtain

$$\begin{aligned}
\hat{R}_{SURE}(\mathbf{Z}) &= K\frac{\sigma^2}{N} + 2\frac{\sigma^2}{N}\sum_{k=1}^{K}\frac{\partial g_k(\mathbf{Z})}{\partial Z_k} + \sum_{k=1]}^{K}\left(\hat{\theta}_k - Z_k\right)^2 \\
&= K\frac{\sigma^2}{N} - 2\frac{\sigma^2}{N}\sum_{k=1}^{K}\mathbf{1}\{|Z_k| < \lambda\} + \sum_{k=1]}^{K}\min\{\lambda^2, Z_k^2\}
\end{aligned} \tag{16}$$

Which is what we were looking for. This estimator is very useful when we have a large set of predictors and we want to select variables. Our goal would be to undestant which variables to keep and which ones to exclude. This estimator by selecting variables, reduce the complexity and the overall risk at the end.

Problem 4. Let $\mathbf{Z} \sim \mathcal{N}(\theta, \frac{\sigma^2}{N}I_K)$, and $\mathcal{M}$ be the class of ordered subsets as

$$\{\emptyset, \{1\}, \{1,2\}, \{1,2,3\}, ..., \{1,2,...,K\}\} \tag{17}$$

and consider the estimator $\hat{\theta}(M) = Z_k\mathbf{1}(k \in M)$ for $k = 1, ..., K$. Use Theorem 2.5.1 to show that

$$\hat{R}_{SURE}(\mathbf{Z}, M) = \frac{\sigma^2}{N}|M| + \sum_{k \in M^c}\left(Z_k^2 - \frac{\sigma^2}{N}\right) \tag{18}$$

where $|M|$ denoting the cardinality of M and M^c the absolute complement of M in the universe $\mathcal{M}$.

Solution. With $\hat{\theta}(M) = Z_k\mathbf{1}(k \in M)$, then $g(\mathbf{Z})$ would become as

$$g(\mathbf{Z}) = \begin{cases} 0 & \text{if} k \in M \\ -Z_k & \text{otherwise} \end{cases} \tag{19}$$

Using Theorem 2.5.1, then the $\hat{R}_{SURE}(\mathbf{Z}, \mathcal{M})$ becomes as

$$\begin{aligned}
\hat{R}_{SURE}(\mathbf{Z}) &= K\frac{\sigma^2}{N} + 2\frac{\sigma^2}{N}\sum_{k=1}^{K}\frac{\partial g_k(\mathbf{Z})}{\partial Z_k} + \sum_{k=1]}^{K}\left(\hat{\theta}_k - Z_k\right)^2 \\
&= K\frac{\sigma^2}{N} + 2\frac{\sigma^2}{N}\sum_{k\in M^c}(-1) + \sum_{k\in M^c} Z_k \\
&= \frac{\sigma^2}{N}\left(K - \sum_{k\in M^c}(-1)\right) + \sum_{k\in M^c}\left(Z_k^2 - \frac{\sigma^2}{N}\right) \\
&= \frac{\sigma^2}{N}M + \sum_{k\in M^c}\left(Z_k^2 - \frac{\sigma^2}{N}\right) \qquad (20)
\end{aligned}$$

Where this is the result we were looking for.

<u>Problem 5.</u> In the python Notebook.