Preston Mui — mui@berkeley.edu
Econ-240a: Pset 2

1 Pencil and Paper Problems

1. Let $f(z)$ be the probability density function of Z; that is,

$$f(z) = (2\pi)^{-\frac{K}{2}}|\Sigma|^{-\frac{1}{2}} \exp\{-\frac{1}{2}(z-\theta)'\Sigma^{-1}(z-\theta)\}$$

Because the variance-covariance matrix is $\Sigma = \frac{\sigma^2}{N} I_K$:

$$\begin{aligned}
f(z) &= \prod_{i=1}^{K}\left(2\pi\frac{\sigma^2}{N}\right)^{-\frac{1}{2}} \exp\{-\frac{1}{2\sigma^2/N}(z_i-\theta_i)^2\} \\
f_k(z) &= -\left(\frac{\sigma^2}{N}\right)^{-1}(z_k-\theta_k)\prod_{i=1}^{K}\left(2\pi\frac{\sigma^2}{N}\right)^{-\frac{1}{2}} \exp\{-\frac{1}{2\sigma^2/N}(z_i-\theta_i)^2\} \\
\nabla f(z) &= \sum_{i=1}^{K} -\left(\frac{\sigma^2}{N}\right)^{-1}(z_k-\theta_k)f(z) \\
-\left(\frac{\sigma^2}{N}\right)\nabla f(z) &= \sum_{i=1}^{K}(z_k-\theta_k)f(z) = [1_{1\times K}](z-\theta)f(z)
\end{aligned}$$

Note that $\mathbb{E}\left[g(Z)'(Z-\theta)\right] = \int_{\mathbb{R}^K} g(z)'(z-\theta)f(z)dz$. By integration of parts,

$$\mathbb{E}\left[g(Z)'(Z-\theta)\right] = -g(z)\frac{\sigma^2}{N}f(z)\Big|_{\mathbb{R}^K} + \int_{\mathbb{R}^K}\nabla g(z)\frac{\sigma^2}{N}f(z)dz = \frac{\sigma^2}{N}\mathbb{E}\left[\nabla g(Z)\right]$$

2. Risk can be rewritten as

$$\begin{aligned}
R(\hat{\theta},\theta) &= E\left[\|\hat{\theta}-Z+Z-\theta\|^2\right] \\
&= E\left[\|g(Z)+Z-\theta\|^2\right] \\
&= E\left[\sum_{k=1}^{K}\left(g(Z_k)^2 + 2g(Z_k)(Z_k-\theta_k) + (Z_k-\theta_k)^2\right)\right] \\
&= E\left[\sum_{k=1}^{K}(\hat{\theta}-Z_k)^2\right] + 2E\left[\sum_{k=1}^{K}g(Z_k)(Z_k-\theta_k)\right] + E\left[\sum_{k=1}^{K}(Z_k-\theta_k)^2\right] \\
&= K\frac{\sigma^2}{N} + 2\frac{\sigma^2}{N}E[\nabla g(Z)] + E\left[\sum_{k=1}^{K}(\hat{\theta}-Z_k)^2\right]
\end{aligned}$$

Because $\hat{R}(Z) = K\frac{\sigma^2}{N} + 2\frac{\sigma^2}{N}\nabla g(Z) + \sum_{k=1}^{K}(\hat{\theta}-Z_k)^2$, $E[\hat{R}(Z)] = R(\hat{\theta},\theta)$

3. Observe that

$$
\begin{aligned}
g_k = \hat{\theta}_k - Z_k &= \operatorname{sgn}(Z_k)(|Z_k| - \lambda)_+ - Z_k \\
&= \begin{cases} Z_k - \lambda - Z_k = -\lambda & \text{if } Z_k > \lambda \\ -Z_k & \text{if } -\lambda \leq Z_k \leq \lambda \\ -(-Z_k - \lambda) - Z_k = \lambda & \text{if } Z_k < -\lambda \end{cases} \\
\implies \frac{\partial g_k}{\partial z_k} &= \begin{cases} 0 \text{ if } |Z_k| > \lambda \\ -1 \text{ if } |Z_k| \leq \lambda \end{cases} = -1(|Z_k| \leq \lambda)
\end{aligned}
$$

and that

$$
(\hat{\theta}_k - Z_k)^2 = \begin{cases} \lambda^2 & \text{if } |Z_k| > \lambda \\ Z_k^2 & \text{if } |Z_k| \leq \lambda \end{cases} = \min(Z_k^2, \lambda^2)
$$

Substituting into the SURE,

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

One concrete prediction problem where I might want to use the soft estimator: Say I am trying to predict firm's price changes using data on their marginal costs and some other set of predictors, and I think there is a lot of noise in the measurement of marginal costs (and maybe other variables as well). Then, I may want to use the soft estimator, because the maximum likelihood estimator will overestimate the magnitude of the coefficient, so I will want to pull it back towards 0.

4. If $\hat{\theta}_k(M) = Z_k 1(k \in M)$,

$$
\begin{aligned}
g(Z) = \hat{\theta}_k(M) - Z &= \begin{cases} 0 & \text{if } k \in M \\ -Z_k & \text{if } k \notin M \end{cases} \\
\implies \frac{\partial g(Z)}{\partial Z_k} &= \begin{cases} 0 & \text{if } k \in M \\ -1 & \text{if } k \notin M \end{cases} \\
&= -1\{k \notin M\}
\end{aligned}
$$

So, the $\hat{R}_{SURE}$ is given by

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