Second Midterm Review Sheet, Part II

Ec240a – Second Half, Fall 2016

In preparing for the exam review your lecture notes, assigned course readings, problem sets and prepare answers to the questions on the review sheet(s). You may bring to the exam a single 8.5×11 inch sheet of paper with notes on it. No calculation aides are allowed in the exam (e.g., calculators, phones, laptops etc.). You may work in pencil or pen, however I advise the use of pencil. Please be sure to bring sufficient blue books with you to the exam if possible.

[1] Let $\{X_i\}_{i=1}^N$ be a simple random sample from some population with density function $f(x)$. Consider the following uniform kernel density estimate of $f(x)$:

$$\hat{f}(x) = \frac{1}{2Nh} \sum_{i=1}^{N} \mathbf{1}(x - h \leq X_i \leq x + h).$$

[a] Show that

$$2Nh\hat{f}(x) \sim \text{binomial}(N, p_h(x))$$

for $p_h(x) = h \int \mathbf{1}(-1 \leq u \leq 1) f(x + uh)\,\mathrm{d}u$ and hence that

$$\mathbb{E}\left[\hat{f}(x)\right] = \frac{1}{2} \int \mathbf{1}(-1 \leq u \leq 1) f(x + uh)\,\mathrm{d}u$$

and further that

$$\mathbb{V}\left(\hat{f}(x)\right) = \frac{p_h(x)[1 - p_h(x)]}{4Nh^2}.$$

[b] Use the Central Limit Theorem introduced in class to prove that

$$\sqrt{N}\left(\hat{f}(x) - \mathbb{E}\left[\hat{f}(x)\right]\right) \xrightarrow{D} \mathcal{N}\left(0, \frac{p_h(x)[1 - p_h(x)]}{4h^2}\right).$$

[c] Use a Taylor series expansion to show that the bias of $\hat{f}(x)$ is, up to order h^2,

$$\mathbb{E}\left[\hat{f}(x)\right] - f(x) = \frac{h^2 f''(x)}{3} + O\left(h^2\right).$$

Further show that the variance is

$$\mathbb{V}\left(\hat{f}(x)\right) = \frac{f(x)}{4Nh} + O\left(\frac{1}{N}\right)$$

and consequently mean square error is

$$\mathbb{E}\left[\left(\hat{f}(x) - f(x)\right)^2\right] = \frac{h^3 [f''(x)]^2}{9} + \frac{f(x)}{4Nh} + O\left(\frac{1}{N}\right) + O\left(h^4\right).$$

[d] Derive an approximate mean squared error minimizing choice of h. Discuss how this choice reflects a "bias-variance trade-off". Can you suggest a feasible way of choosing h in practice?

[2] Consider the statistical model

$$\mathbf{Z} \sim N\left(\theta, \frac{\sigma^2}{N} I_K\right)$$

and estimate of $\theta = (\theta_1, \ldots, \theta_K)'$

$$\hat{\theta} = \arg\min_{\lambda} \sum_{k=1}^{K} (Z_k - \theta_k)^2 + \lambda \sum_{k=1}^{K} |\theta_k| .$$

[a] Show that

$$\hat{\theta}_{k,\lambda} = \operatorname{sgn}(Z_k)\,(|Z_k| - \lambda)_+ .$$

Discuss why this estimator is called the soft-threshold estimator. In what settings will this estimator exhibit low risk?

[b] Show that Stein's Unbiased Estimate of Risk (SURE) for this procedure is

$$\hat{R}_{\mathrm{SURE}}(\mathbf{Z}) = \sum_{k=1}^{K} \left\{ \frac{\sigma^2}{N} - \frac{2\sigma^2}{N} \mathbf{1}(|Z_k| \leq \lambda) + \min\left(Z_k^2, \lambda^2\right) \right\}.$$

Describe how to use $\hat{R}_{\mathrm{SURE}}(\mathbf{Z})$ to choose λ in practice.

[c] Now consider the following curved soft-threshold estimator

$$\hat{\theta}_{k,\lambda} = \begin{cases} -(Z_k + \lambda)^2 & Z_k < -\lambda \\ 0 & -\lambda \leq Z_k \leq \lambda \\ -(Z_k + \lambda)^2 & Z_k > \lambda \end{cases}$$

Can you constructed a penalized least squares representation of this estimator?

[d] Calculate $\hat{R}_{\mathrm{SURE}}(\mathbf{Z})$ for the curved soft-threshold estimator.

[e] In what settings will this estimator exhibit low risk?