Ec240a, Fall 2019

Professor Bryan Graham

Problem Set 3

Due: November 22nd, 2019

Problem sets are due at 5PM in the GSIs mailbox. You may work in groups, but each student should turn in their own write-up (including a printout of a narrated/commented and executed Jupyter Notebook if applicable). Please also e-mail a copy of any Jupyter Notebook to the GSI (if applicable).

1 K Normal Means: Theory and Simulation

1. Let $\mathbf{Z} \sim \mathcal{N}\left(\theta, \frac{\sigma^2}{N} I_K\right)$. Show that $\mathbb{E}\left[g\left(\mathbf{Z}\right)\left(\mathbf{Z}-\theta\right)'\right] = \frac{\sigma^2}{N}\mathbb{E}\left[\nabla_{\mathbf{Z}} g\left(\mathbf{Z}\right)\right]$ (HINT: Use integration by parts).

2. Consider the estimate of the Risk associated with the weakly differentiable estimate $\hat{\theta}$ due to Stein and introduced in lecture:

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

   Prove that this risk estimate is unbiased under square error loss: $\mathbb{E}_\theta\left[\hat{R}\left(\mathbf{Z}\right)\right] = R\left(\hat{\theta}, \theta\right)$.

3. (Wasserman, 2006) Let $\mathbf{Z} \sim \mathcal{N}\left(\theta, \frac{\sigma^2}{N} I_K\right)$ and consider the following soft threshold estimate of θ :

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

   In words 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 SURE to show that

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

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

4. (Wasserman, 2006) Let $\mathbf{Z} \sim \mathcal{N}\left(\theta, \frac{\sigma^2}{N} I_K\right)$ and $\mathcal{M}$ be the class of ordered subsets

$$\{\emptyset, \{1\}, \{1,2\}, \{1,2,3\}, \ldots, \{1, \ldots, K\}\}$$

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

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

with $|M|$ denoting the cardinality of M and M^c the absolute complement of M in the universe $\mathcal{M}$ (HINT: This problem was done in lecture!).

5. (Efromovich, 1999) Consider the cosine orthonormal system of $[0,1]$ with elements

$$\phi_1(x) = 1,\ \phi_k(x) = \sqrt{2}\cos(\pi(k-1)x),\ k = 2, \ldots.$$

Let $\mathcal{N}(x; \mu, \sigma)$ denote the normal density with mean μ and standard deviation σ at x. Consider the following "bimodal" and "steps" test functions introduced by Efromovich (1999, p. 18):

$$\begin{aligned} m_{\text{B}}(x) &= \frac{1}{2}\mathcal{N}(x; 0.4, 0.12) + \frac{1}{2}\mathcal{N}(x; 0.7, 0.08) \\ m_{\text{S}}(x) &= 0.6 \times \mathbf{1}\left(0 \le x < \frac{1}{3}\right) + 0.9 \times \mathbf{1}\left(\frac{1}{3} \le x < \frac{3}{4}\right) + \frac{204}{120} \times \mathbf{1}\left(\frac{3}{4} \le x \le 1\right). \end{aligned}$$

(a) Let $X_i = i/(N+1)$ for $i = 1, \ldots, N$ and $Y_i = m_{\text{j}}(X_i) + \sigma U_i$ with U_i a standard normal random variable and $j = \text{B}, \text{S}$. For each of the two functions generate 1000 random samples of size $N = 1000$ with $\sigma = 1/2$. Set $K = 20$. For each sample compute $Z_k = \frac{1}{N}\sum_{i=1}^{N} \phi_k(X_i) Y_i$ for $k = 1, \ldots, K$ and construct two estimates of $\mathbf{m}$: (i) the MLE with $\hat{m}_{\text{ML}}(x) = \sum_{k=1}^{K} \phi_k(x) Z_k$ and (ii) Efromovich's with $\hat{m}_{\text{EF}}(x) = \sum_{k=1}^{N} \hat{c}_k \phi_k(x) Z_k$ with $\hat{c}_k$ as defined in lecture. Compute squared error loss for each sample, estimate and test function.

(b) Compute average loss across your 1000 random samples for each test function and estimator and report the results in a Table. For each test function and estimator find the sample with median loss (across the 1000 simulated samples). For this sample plot the true test function, your estimate and a scatter of the data (four figures in all). Briefly discuss your findings.

(c) Using the same simulated samples from parts (a) and (b) compute the soft thresholding estimate of $\mathbf{m}$. For each sample choose λ to minimize the SURE expression from problem 3 above. Compute average loss across your 1000 random samples for each test function and report the results in a table. For each test function find the sample with median loss (across the 1000 simulated samples). For this sample plot the true test function, your estimate and a scatter of the data (two figures in all). Briefly discuss your findings.

2 K Normal Means: Application

In this problem you will use the Hall and Jones (1999, QJE) dataset, available online at Chad Jone's data archive, to model the relationship between log GDP per worker and latitude. Begin with a design matrix consisting of a constant and the first 14 powers of latitude (for 15 basis functions in all). Use Gram-Schmidt orthonormalization to transform this design matrix as described in class. Next compute the conditional mean of log GDP per worker given latitude by each of (i) MLE, (ii) soft-thresholding, (iii) ordered subset selection and (iv) Efromovich's method. Plot each estimated function on a scatter plot (four figures in all). Comment on your results.