Stat 300A Theory of Statistics

Homework 4

Andrea Montanari *Due on October 24, 2018*

- Solutions should be complete and concisely written. Please, use a separate sheet (or set of sheets) for each problem.
- We will be using Gradescope (`https://www.gradescope.com`) for homework submission (you should have received an invitation) - no paper homework will be accepted. Handwritten solutions are still fine though, just make a good quality scan and upload it to Gradescope.
- You are welcome to discuss problems with your colleagues, but should write and submit your own solution.

1: Small noise limit in the bounded normal mean model

Consider the normal mean model in one dimension $\mathsf{P}_\theta = \mathsf{N}(\theta, \sigma^2)$ with $\theta \in \Theta = [-1, 1]$, and σ^2 known and square loss. Denote by $R_{\rm M}(\Theta; \sigma^2)$ the corresponding minimax error. The objective of this homework is to prove that

$$R_{\rm M}(\Theta; \sigma^2) = \sigma^2 + o(\sigma^2)\,. \tag{1}$$

(In other words, we want to show that $\lim_{\sigma \to 0} R_{\rm M}(\Theta; \sigma^2)/\sigma^2 = 1$.)

(*a*) Prove that, for any $\sigma^2 > 0$,

$$R_{\rm M}(\Theta; \sigma^2) \le \sigma^2\,. \tag{2}$$

(*b*) Construct a sequence of priors Q_σ indexed by $\sigma > 0$, and prove that

$$\liminf_{\sigma \to 0} \frac{1}{\sigma^2} R_{\rm B}(\mathsf{Q}_\sigma) \ge 1\,. \tag{3}$$

[Hint: Here are two possible choices for the prior Q_σ (both allow to prove the claim, with sufficient work, and you might be able to come up with better constructions). Choice #1: Take $\mathsf{Q}_\sigma = \mathsf{Unif}([-1, 1])$ the uniform distribution in $[-1, 1]$. Choice #2: Let $v(\sigma) > 0$, be such that $v(\sigma) \to 0$ and $v(\sigma)/\sigma^2 \to \infty$, as $\sigma \to 0$. Than take Q_σ to be the distribution of a $\mathsf{N}(0, v(\sigma))$ random variable conditional to lie in $[-1, 1]$.]

(*c*) Conclude that the results above imply the claim (1).

2: A modified James-Stein estimator

Consider again the normal mean model in d dimensions $\mathsf{P}_{\boldsymbol{\theta}} = \mathsf{N}(\boldsymbol{\theta}, \mathrm{I})$, $\boldsymbol{\theta} \in \Theta = \mathbb{R}^d$, with square loss. We denote by $\mathbf{1} = (1, \dots, 1)^{\mathsf{T}}$ the all-ones vector in d dimensions, and define the modified James-Stein estimator

$$\hat{\boldsymbol{\theta}}(\boldsymbol{x}) = \overline{x}\,\mathbf{1} + \big(1 - h(\|\boldsymbol{x} - \overline{x}\mathbf{1}\|_2^2)\big)\ (\boldsymbol{x} - \overline{x}\mathbf{1})\,. \tag{4}$$

where $\overline{x} = \langle \boldsymbol{x}, \mathbf{1} \rangle / d$ is the average of the entries of vector $\boldsymbol{x}$.

(a) Use SURE to give an expression for the risk of this estimator. Assume all the nice conditions required to apply SURE.

(b) Specialize this expression to $h(u) = C/u$, for C a constant.

(c) Show that, for $d \geq 3$, the constant C can be chosen in such a way that $\hat{\boldsymbol{\theta}}$ dominates the unbiased estimator $\hat{\boldsymbol{\theta}}_{\rm u}(\boldsymbol{x}) = \boldsymbol{x}$.

(d) Provide an empirical Bayes interpretation of $\hat{\boldsymbol{\theta}}$.

3: A regression problem with random designs

Let $\boldsymbol{\theta} \in \Theta \equiv \{\boldsymbol{v} \in \mathbb{R}^d : \;\; \|\boldsymbol{v}\|_2 = 1\}$ and suppose you are given n i.i.d. data points $(y_1, \boldsymbol{x}_1)$, $\dots (y_n, \boldsymbol{x}_n)$, where $\boldsymbol{x}_i \sim \mathsf{N}(0, \mathrm{I}_d)$ and

$$y_i = \boldsymbol{x}_i^{\mathsf{T}} \boldsymbol{\theta} + w_i \, , \tag{5}$$

with $w_i \sim \mathsf{N}(0, \sigma^2)$ independent of $\boldsymbol{x}_i$. We denote collectively the data by $(\boldsymbol{y}, \boldsymbol{X})$. We consider the action space $\mathcal{A} = \mathbb{R}^d$, and the loss function

$$L(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) = \|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\|^2 \, . \tag{6}$$

(a) Show that, in order to construct a minimax estimator $\hat{\boldsymbol{\theta}}$, you can assume without loss of generality that $\hat{\boldsymbol{\theta}}$ takes values in a closed ball $B^d(R)$.

(b) Construct a least favorable prior Q_*.

(c) Write a formal expression for the minimax estimator.

(d) Upper bound the minimax risk by considering the estimator

$$\hat{\boldsymbol{\theta}}(\boldsymbol{y}, \boldsymbol{X}) = \frac{1}{C(n)} \sum_{i=1}^{n} y_i \boldsymbol{x}_i \, . \tag{7}$$

Optimize the resulting bound by tuning the constant $C(n)$.