Homework 12
36-705
Due: Thursday November 19th by 3pm.

1. Consider the infinite Gaussian sequence model

$$Y_i = \theta_i + \frac{\sigma}{\sqrt{n}}\epsilon_i \qquad i = 1, 2, \ldots,$$

where $\epsilon_1, \epsilon_2, \ldots \sim N(0, 1)$.

(a) Let $\widehat{\theta} = (Y_1, Y_2, \ldots,)$ be the maximum likelihood estimator. Show that this estimator is not consistent i.e. $P(||\widehat{\theta} - \theta|| < \delta)$ does not tend to 1 for every δ.

(b) Let $\widehat{\theta} = (\widehat{\theta}_1, \widehat{\theta}_2, \ldots)$ where

$$\widehat{\theta}_i = \begin{cases} Y_i & i \leq k_n \\ 0 & i > k_n. \end{cases}$$

Show that if $k_n \to \infty$ as $n \to \infty$ and $k_n = o(n)$ then

$$\sum_{i=1}^{\infty} (\widehat{\theta}_i - \theta_i)^2 \xrightarrow{P} 0.$$

2. We showed that the hard-thresholding algorithm for the Gaussian sequence model satisfies

$$R(\widehat{\theta}, \theta^*) \leq CR\sigma\sqrt{\frac{\log d}{n}},$$

for estimating a vector θ^* that is ℓ_1-sparse, i.e. satisfies $\sum_{i=1}^d |\theta_i| \leq R$ (where $C > 0$ is some constant). Suppose that, θ^* is instead ℓ_q sparse for some $q \in (0, 1]$, i.e.

$$\sum_{i=1}^d |\theta_i|^q \leq R_q.$$

Then show that for some constant $C > 0$ the hard-thresholding estimator has risk:

$$R(\widehat{\theta}, \theta^*) \leq CR_q \left(\frac{\sigma^2 \log d}{n}\right)^{1-q/2}.$$

Notice that once again the estimator can have risk $\to 0$ even when $d \gg n$. Furthermore, notice that the same hard-thresholding estimator works for a variety of different notions of sparsity.

3. Let

$$Y_i = \beta X_i + \epsilon_i$$

where $\mathbb{E}[\epsilon_i] = 0$ and $\text{Var}[\epsilon_i] = \sigma^2$.

(a) Find the least squares estimator $\widehat{\beta}$.

(b) Find the limiting distribution of $\sqrt{n}(\widehat{\beta} - \beta)$ (not conditional on the X_i's).

4. Suppose that

$$Y_i = \beta X_i + \epsilon_i,$$

where $\epsilon_i \sim N(0, \sigma^2)$, and X_i are i.i.d., have finite second moment, and $X_i \in \mathbb{R}$. Rather than observe X_i, however, we observe a noisy version W_i where,

$$W_i = X_i + \delta_i,$$

and $\delta_i \sim N(0, \tau^2)$, independently of everything else. So we observe n i.i.d samples of the form $\{(Y_1, W_1), \ldots, (Y_n, W_n)\}$. Consider the least squares estimator:

$$\widehat{\beta} = \frac{\sum_{i=1}^n Y_i W_i}{\sum_{i=1}^n W_i^2},$$

and show that the estimator is inconsistent, i.e. show that $\widehat{\beta} \xrightarrow{P} a\beta$ where $a \neq 1$. Find a. Suppose that τ was known to you. Construct a consistent estimator of β.