STAT151A Homework 5: Due March 22nd

Your name here

Reviewing the distribution of $\hat{\beta}$ under different assumptions

This homework question will reference the following assumptions.

Regressor assumptions:

- R1: The $N \times P$ matrix $\boldsymbol{X}$, which has $\boldsymbol{x}_n^\intercal$ in the n–th row, is full column rank
- R2: The regressors $\boldsymbol{x}_n$ are deterministic, and $\frac{1}{N}\sum_{n=1}^N \boldsymbol{x}_n \boldsymbol{x}_n^\intercal \rightarrow \boldsymbol{\Sigma}_{\boldsymbol{X}}$, where $\boldsymbol{\Sigma}_{\boldsymbol{X}}$ is positive definite
- R3: The regressors $\boldsymbol{x}_n$ are IID, with positive definite covariance $\mathrm{Cov}\left(\boldsymbol{x}_n\right)$, and $\frac{1}{N}\sum_{n=1}^N \boldsymbol{x}_n \boldsymbol{x}_n^\intercal \rightarrow \mathbb{E}\left[\boldsymbol{x}_n \boldsymbol{x}_n^\intercal\right]$ in probability.

Model assumptions (for all n):

- M1: There exists a $\boldsymbol{\beta}$ such that $y_n = \boldsymbol{\beta}^\intercal \boldsymbol{x}_n + \varepsilon_n$ for all n
- M2: The residuals ε_n are IID with $\varepsilon_n | \boldsymbol{x}_n \sim \mathcal{N}\left(0, \sigma^2\right)$ for some σ^2
- M3: The residuals ε_n are independent with $\mathbb{E}\left[\varepsilon_n | \boldsymbol{x}_n\right] = 0$ and $\mathbb{E}\left[\varepsilon_n^2 | \boldsymbol{x}_n\right] = \sigma^2$
- M4: The residuals ε_n are independent with $\mathbb{E}\left[\varepsilon_n | \boldsymbol{x}_n\right] = 0$ and $\mathbb{E}\left[\varepsilon_n^2 | \boldsymbol{x}_n\right] = \sigma_n^2$
- M5: The pairs $(\boldsymbol{x}_n, y_n)$ are IID
- M6: For all finite vectors $\boldsymbol{v}$, $\frac{1}{N}\sum_{n=1}^N \mathbb{E}\left[y_n - \boldsymbol{v}^\intercal \boldsymbol{x}_n)^2 \boldsymbol{x}_n \boldsymbol{x}_n^\intercal\right] \rightarrow \boldsymbol{V}(\boldsymbol{v}) < \infty$, where each entry of the limiting matrix is finite. (The limit depends on $\boldsymbol{v}$, but importantly $\boldsymbol{V}(\boldsymbol{v})$ is finite for all finite $\boldsymbol{v}$).

For M2, M3, and M4 with $\boldsymbol{x}_n$ is deterministic, take the conditioning to mean "for that value of $\boldsymbol{x}_n$."

For this homework, you may use the LLN, the CLT, the continuous mapping theorem, and properties of the multivariate normal distribution.

The term "limiting distribution" means the distribution that the quantity approaches as $N \rightarrow \infty$.

Assume R1 for all questions.

1) Find the distribution of $\hat{\beta}$ under M1, M2, and R2.
2) Find the limiting distribution of $\sqrt{N}(\hat{\beta} - \beta)$ under M1, M2, and R2.
3) Find the limiting distribution of $\sqrt{N}(\hat{\beta} - \beta)$ under M1, M2, and R3.
4) Find the limiting distribution of $\sqrt{N}(\hat{\beta} - \beta)$ under M1, M3, and R2.
5) Find the limiting distribution of $\sqrt{N}(\hat{\beta} - \beta)$ under M1, M3, and R3.
6) Find the limiting distribution of $\sqrt{N}(\hat{\beta} - \beta)$ under M1, M4, M6, and R2.
7) Find the limiting distribution of $\sqrt{N}(\hat{\beta} - \beta)$ under M1, M4, M6, and R3.
8) Under M5, M6, and R3, identify a β^* such that $\sqrt{N}(\hat{\beta} - \beta^*)$ converges to a nondegenerate, finite random variable, and find the limiting distribution.
9) In any of the above settings, what is the limiting distribution of $(\hat{\beta} - \beta)$? (The answer is the same no matter which setting you choose.)

Investigating the assumptions

1) What goes wrong if R1 is violated?
2) What goes wrong if $\boldsymbol{\Sigma}_{\boldsymbol{X}}$ is not positive definite in R2?
3) What goes wrong if $\mathbb{E}\left[\boldsymbol{x}_n \boldsymbol{x}_n^\intercal\right]$ is not positive definite in R3?
4) In terms of the limiting distributions, what is the practical difference between assumptions R2 and R3?
5) Assume R1, R2, M1, and M3, except take $\mathbb{E}\left[\varepsilon_n | \boldsymbol{x}_n\right] = \delta \neq 0$. What happens to $\sqrt{N}(\hat{\beta} - \beta)$ as $N \to \infty$?
6) Assume R1, R3, M5, and M6. In general, is it true that $\mathbb{E}\left[y_n - \beta^* \boldsymbol{x}_n\right] = 0$?

Confidence intervals for components of $\hat{\beta}$

Suppose that $\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{V}\right)$ for some $\boldsymbol{\beta}$ and $\boldsymbol{V}$. Let $\Phi(z) := \mathbb{P}\left(\tilde{z} \leq z\right)$ where $\tilde{z} \sim \mathcal{N}\left(0, 1\right)$. Fix a vector $\boldsymbol{a}$ of the same length as $\boldsymbol{\beta}$, and $0 \leq \alpha \leq 1$.

In terms of Φ, $\boldsymbol{V}$, and β, find a scalar b such that

$$\mathbb{P}\left(-b \leq \boldsymbol{a}^\intercal (\hat{\beta} - \beta) \leq b\right) = 1 - \alpha.$$

In particular, what is the result when $\boldsymbol{a} = (0, \ldots, 0, 1, 0, \ldots, 0)$ is a vector with 1 in location k and 0 elsewhere?