Econometric Analysis of Cross Section and Panel Data
Problem Set 2

Due date: October 28

Problem 1: Recall the prediction errors $\widetilde{e}_i = (1 - h_{ii})^{-1} \widehat{e}_i$. In vector notation, $\widetilde{\boldsymbol{e}} = \boldsymbol{M}^* \widehat{\boldsymbol{e}}$ where $\boldsymbol{M}^*$ is a diagonal matrix with i^{th} diagonal element $(1 - h_{ii})^{-1}$. Thus $\widetilde{\boldsymbol{e}} = \boldsymbol{M}^* \boldsymbol{M} \boldsymbol{e}$.

(a) Calculate $\mathbb{E}[\widetilde{\boldsymbol{e}} \mid \boldsymbol{X}]$.

(b) Calculate $\text{var}[\widetilde{\boldsymbol{e}} \mid \boldsymbol{X}]$.

(c) Calculate the variance of the i^{th} prediction error $\text{var}\left[\widetilde{e}_i \mid \boldsymbol{X}\right]$ under the assumption of homoskedasticity.

(d) Define the standardized residuals as $\bar{e}_i = (1 - h_{ii})^{-1/2} \widehat{e}_i$ or in vector notation $\overline{\boldsymbol{e}} = (\bar{e}_1, \ldots, \bar{e}_n)' = \boldsymbol{M}^{*1/2} \boldsymbol{M} \boldsymbol{e}$. Show that $\text{var}\left[\bar{e}_i \mid \boldsymbol{X}\right] = \mathbb{E}\left[\bar{e}_i^2 \mid \boldsymbol{X}\right] = \sigma^2$.

Problem 2: Of the variables (Y^*, Y, X) only the pair (Y, X) are observed. In this case we say that Y^* is a **latent variable**. Suppose

$$\begin{aligned} Y^* &= X'\beta + e \\ \mathbb{E}[Xe] &= 0 \\ Y &= Y^* + u \end{aligned}$$

where u is a measurement error satisfying

$$\begin{aligned} \mathbb{E}[Xu] &= 0 \\ \mathbb{E}\left[Y^* u\right] &= 0 \end{aligned}$$

Let $\widehat{\beta}$ denote the OLS coefficient from the regression of Y on X.

(a) Is β the coefficient from the linear projection of Y on X?

(b) Is $\widehat{\beta}$ consistent for β as $n \to \infty$?

(c) Find the asymptotic distribution of $\sqrt{n}(\widehat{\beta} - \beta)$ as $n \to \infty$.

Problem 3: Take the linear regression model with $\mathbb{E}[\boldsymbol{Y} \mid \boldsymbol{X}] = \boldsymbol{X}\boldsymbol{\beta}$. Define the ridge regression estimator

$$\widehat{\beta} = (\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{I}_k\lambda)^{-1} \boldsymbol{X}'\boldsymbol{Y}$$

where $\lambda > 0$ is a fixed constant.

(a) Find $E[\widehat{\boldsymbol{\beta}} \mid \boldsymbol{X}]$. Is $\widehat{\beta}$ biased for β in limited samples?

(b) Is $\widehat{\beta}$ consistent for β as $n \to \infty$?