Economics 672: Econometric Analysis II
Winter 2018

Matias Cattaneo
Department of Economics
University of Michigan
Authored collaboratively by PhD entering class of 2017

January 16, 2018

Outline

Introduction

- This lecture:
  - Ordinary Least Squares

Outline

Assumptions and setup

- $\{(y_i, x_i')' : i = 1, \ldots, n\}$, $y_i \in \mathbb{R}$, $x_i \in \mathbb{R}^d$
  - y is a $n \times 1$ vector of outcomes of interest, X is a $n \times d$ matrix; the i^{th} row of which is x_i'
- Postulate 2 basic assumptions:
  - $\mathbb{E}(y|x) = x\beta_0$, where $\beta_0 \in \mathbb{R}^d$
  - $\mathbb{V}(y|x) = \sigma_0^2 I_n$ (true under homoskedasticity and iid sampling)
  - Later, add a third assumption: $Y|X \sim \mathcal{N}(x\beta_0, \sigma_0^2 I_n)$

With a model of the form $Y = X\beta_0 + \varepsilon$, these conditions become: $\mathbb{E}(\varepsilon|X) = 0$ (A1), $\mathbb{V}(\varepsilon|X) = \sigma_0^2 I_n$ (A2), and $\varepsilon|X \sim \mathcal{N}(0, \sigma_0^2 I_n)$ (A3).

OLS Normal Equations

We want to solve the problem $\min_{\beta \in \mathbb{R}^d} ||y - X\beta||^2 = \min_{\beta \in \mathbb{R}^d} \sum_{i=1}^{n} (y_i - x_i'\beta)^2$. Postulate that there exists at least one solution to this: call it $\hat{\beta}$.

$$
\begin{aligned}
||y - X\beta||^2 &= ||y - X\hat{\beta} + X\hat{\beta} - X\beta||^2 \\
&= (y - X\hat{\beta} + X\hat{\beta} - X\beta)'(y - X\hat{\beta} + X\hat{\beta} - X\beta) \\
&= (y - X\hat{\beta})'(y - X\hat{\beta}) + 2(y - X\hat{\beta})(X\hat{\beta} - X\beta) + (X\hat{\beta} - X\beta)'(X\hat{\beta} - X \\
&= ||y - X\hat{\beta}||^2 + 2(y - X\hat{\beta})'X(\hat{\beta} - \beta) + ||X(\hat{\beta} - \beta)||^2
\end{aligned}
$$

Setting $\hat{\beta}$ such that the middle term $= 0$ implies that we choose $\hat{\beta}$ such that $(y - X\hat{\beta})'X = 0$. This implies the **least squares normal equations**:

$$\hat{\beta} = (X'X)^{-}X'y$$

- $y - X\hat{\beta} = 0$ only in the case of perfect prediction. In general this will not happen; but if y is a vector that can be generated perfectly by the column space of X, then it will hold.

Projection and Annihilator Matrices

- By A1, $\hat{\beta} = (X'X)^{-}X'y \Rightarrow \hat{y} = X(X'X)^{-}X'y$
  - $\hat{y}$ is the closest approximation to y that lives in the space generated by the column vectors of X.
- $P_x = X(X'X)^{-}X'$ is the **projection matrix**. Properties include:
  - X is invariant under P_x: $P_x X = X$
  - Symmetry: $P_x' = P_x$
  - Idempotence: $P_x P_x = P_x$
- $M_x = I - P_x$ projects on the orthogonal complement of X. Properties include:
  - $M_x X = 0$
  - $M_x P_x = 0$
  - $M_x' = M_x$
  - $M_x M_x = M_x$
- By definition of $\hat{\varepsilon}$ and $\hat{y}$, $\hat{\varepsilon} = y - \hat{y} = (I - P_x)y$
  - The residuals, $\hat{\varepsilon}$, are the projection of y onto the orthogonal complement of X.
  - $\hat{\varepsilon}'\hat{y} = 0$ and $X'\hat{\varepsilon} = 0$
  - Any vector v can be written as $v = P_v v + M_v v$

Gauss-Markov Theorem

If the following two assumptions hold:

- $\mathbb{E}(Y|X) = X\beta_0$
- $\mathbb{V}(Y|X) = \sigma_0^2 I_n$

Then $\hat{\beta} = (X'X)^{-}X'y$ is the best linear (conditionally) unbiased estimator of β_0. $\mathbb{V}(\hat{\beta}|X)$ is no larger than that of any other (conditionally) unbiased estimator of β_0.

Outline

Next Class Topics and Readings

- Problem Set #
- Next topic: blah blah blah
- **Readings:**
  1. books
  2. more books