Midterm Exam

Gov 51, Spring 2024

March, 2024

1 Matching Estimator

One of the key identifying assumption of matching is:

$$\underbrace{\mathbb{E}(Y_i(0)|T_i = 1)}_{\text{unobserved}} = \mathbb{E}(Y_i(0)|T_i = 0, i \in \mathcal{M})$$

Can you explain this equality in several sentences? (5 pts)

Answer:

On average, outcomes for those treated would have been the same to those in the matched set (not treated), if they were not treated.

2 Instrumental Variable

Can you explain exclusion restriction in words? (no mathematical expression needed). If it is helpful, draw a causal chart as we've seen in the slides. (5 pts)

Answer:

Instrument can only affect outcome through treatment.

3 OLS and Lasso

3.1 Loss function

We know that OLS regression and Lasso have different loss functions. Write down the two loss functions. (5 pts)

Answer:

$$\widehat{\beta}_{\text{OLS}} = \arg\min_{\widetilde{\beta}} \sum_{i=1}^{N} \left(Y_i - \widetilde{\beta} X_i\right)^2$$

$$\widehat{\beta}_{\text{Lasso}} = \arg\min_{\widetilde{\beta}} \sum_{i=1}^{N} \left(Y_i - \widetilde{\beta} X_i\right)^2 + \lambda|\tilde{\beta}|$$

3.2 Interpret the difference

1. When should we choose OLS over Lasso? And when do we prefer Lasso to OLS? (5 pts)

Answer:

- Loss function tells us how wrong we are.
- Lasso: Penalization and shrinkage.
- OLS - for unbiased estimation of β. LASSO for variable selection.

3.3 Toy example

Please consider the toy dataset below.

Unit	Y	X
1	-5	1
2	-10	2
3	-15	3
4	-20	4
5	-25	5

3.3.1 OLS

Assume a linear model: $Y = \beta X + 0$. Calculate β_{OLS} for this dataset. (5 pts)

Answer:

$$\begin{aligned}
\beta_{\text{OLS}} &= \frac{\sum_i X_i Y_i}{\sum_i X_i X_i} \\
&= \frac{-5 \cdot 1 - 10 \cdot 2 - 15 \cdot 3 - 20 \cdot 4 - 25 \cdot 5}{1 + 4 + 9 + 16 + 25} \\
&= \frac{-275}{55} \\
&= -5
\end{aligned}$$

3.3.2 Lasso

Suppose $\lambda = 110$, can you calculate β_{Lasso} for this dataset? (5 pts)

Answer:

$$\lambda = 110$$

This makes $\frac{\lambda}{\sum_i^N X_i X_i} = \frac{110}{55} = 2$

$$\begin{aligned}\beta_{\text{Lasso}} &= \left(\frac{\sum_i X_i Y_i}{\sum_i X_i X_i} + \frac{\lambda}{\sum_i X_i X_i}\right) \cdot \mathbf{1}\left(|\beta_{\text{OLS}}| > 1\right) \\ &= (-5 + 2) * 1 \\ &= -3\end{aligned}$$

3.3.3 Comparison

Substitute β_{Lasso} and β_{OLS} to the loss function of Lasso, and show that β_{Lasso} is the better value to minimize the loss function. (5 pts)

Answer:

- When $\hat{\beta} = -3$, the loss function is:

  $\frac{1}{2}\left((-5-(-3)*1)^2 + (-10-(-3)*2)^2 + (-15-(-3)*3)^2 + (-20-(-3)*4)^2 + (-25-(-3)*5)^2\right) + 110*3 = 440$

- When $\hat{\beta} = -5$, the loss function is:

$$\frac{1}{2}(0) + 110 * 5 = 550$$

4 Difference in difference and OLS

4.1 Four scenarios

Assume a linear regression to run DID design:

$$Y_i = \beta_0 + \beta_1 \cdot \text{period} + \beta_2 \cdot \text{treatment} + \beta_3 \cdot \text{period} \cdot \text{treatment} + \epsilon_i$$

Fill in the table below with the aforementioned linear regression specification using the values for treatment and period provided. In order to get full credit, you will need to provide four answers. (8 pts)

	Period= 0	Period= 1
Treatment = 0		
Treatment = 1		

Answer:

	Period= 0	Period= 1
Treatment = 0	$Y_i = \beta_0 + \epsilon_i$	$Y_i = \beta_0 + \beta_1 + \epsilon_i$
Treatment = 1	$Y_i = \beta_0 + \beta_2 + \epsilon_i$	$Y_i = \beta_0 + \beta_1 + \beta_2 + \beta_3 + \epsilon_i$

4.2 DID estimator

From the table above, what is the value of DID estimator? Please show the process that you derive the result. (7 pts)

Answer:

The DID estimator would be:

$$\underbrace{(\beta_0 + \beta_1 + \beta_2 + \beta_3 + \epsilon_i)}_{\text{Period 1, treated}} - \underbrace{(\beta_0 + \beta_2 + \epsilon_i)}_{\text{Period 0, treated}}$$
$$-$$
$$(\underbrace{(\beta_0 + \beta_1 + \epsilon_i)}_{\text{Period 1, control}} - \underbrace{(\beta_0 + \epsilon_i)}_{\text{Period 0, control}})$$

Which yields: β_3

5 Bonus Question

Please show the following equation to be true:

$$\hat{\beta} = \frac{\sum_{i=1}^{N} X_i(Y_i - \bar{Y})}{\sum_{i=1}^{N} X_i(X_i - \bar{X})} = \frac{\sum_{i=1}^{N}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{N}(X_i - \bar{X})^2}$$

Hint: spread out both numerator and denominator of both forms, and take a closer look at the difference. (5 pts)

answer

The key is to use below two qualities:

$$\mathbb{E}(\bar{X}(Y - \bar{Y})) = \bar{X}\mathbb{E}(Y - \bar{Y}) = 0$$
$$\mathbb{E}(\bar{X}(X - \bar{X})) = \bar{X}\mathbb{E}(X - \bar{X}) = 0$$

Answer:

Let's take a look at numerator:

$$\begin{aligned}
&\mathbb{E}\left((X - \bar{X})(Y - \bar{Y})\right) - \mathbb{E}\left(X(Y - \bar{Y})\right) \\
&= \mathbb{E}(\bar{X}(Y - \bar{Y})) \\
&= \bar{X}\mathbb{E}(Y - \bar{Y}) \\
&= 0
\end{aligned}$$

Let's take a look at denominator:

$$\begin{aligned}
&\mathbb{E}\left((X-\bar{X})(X-\bar{X})\right)-\mathbb{E}\left(X(X-\bar{X})\right)\\
&=\mathbb{E}(\bar{X}(X-\bar{X}))\\
&=\bar{X}\mathbb{E}(X-\bar{X})\\
&=0
\end{aligned}$$