Econometrics I - TA section

Eric Schulman

The University of Texas at Austin

October 30, 2020

Chapter 2

- Theoretical model, no data involved.

Chapter 2

- Theoretical model, no data involved.
- What would happen if we observed the population distribution for y and x?
- What do realizations x tell you about y in theory?
- Compute, $E(y|x)$. Assume that it's linear

Chapter 2

- Theoretical model, no data involved.
- What would happen if we observed the population distribution for y and x?
- What do realizations x tell you about y in theory?
- Compute, $E(y|x)$. Assume that it's linear (possibly incorrectly)

2.16 Let x and y have the joint density $f(x, y) = \frac{3}{2}(x^2 + y^2)$ where $x \in [0, 1]$ and $y \in [0, 1]$
Compute the coefficients of the best linear predictor $y = \alpha + \beta + e$. Compute the conditional mean $m(x) = E(y|x)$. Are the best linear predictor and conditional mean different?

2.16 Let x and y have the joint density $f(x,y) = \frac{3}{2}(x^2 + y^2)$ where $x \in [0,1]$ and $y \in [0,1]$
Compute the coefficients of the best linear predictor $y = \alpha + \beta + e$. Compute the conditional mean $m(x) = E(y|x)$.
Are the best linear predictor and conditional mean different?

2.16 Best linear predictor $y = \alpha + \beta x$

$$\beta = \frac{E(xy) - E(x)E(y)}{E(x^2) - E(x)^2}$$

$$\alpha = E(y) - \beta E(x)$$

2.16 Best linear predictor $y = \alpha + \beta x$

$$\beta = \frac{E(xy) - E(x)E(y)}{E(x^2) - E(x)^2}$$

$$\alpha = E(y) - \beta E(x)$$

- $E(xy) = \int_0^1 \int_0^1 xyf(x,y)dxdy = 3/8$
- $E(x) = \int_0^1 \int_0^1 xf(x,y)dxdy = 5/8$
- $E(y) = \int_0^1 \int_0^1 yf(x,y)dxdy = 5/8$
- $E(x^2) = \int_0^1 \int_0^1 x^2 f(x,y)dxdy = 7/15$

2.16 Best linear predictor $y = \alpha + \beta x$

$$\beta = \frac{E(xy) - E(x)E(y)}{E(x^2) - E(x)^2}$$

$$\alpha = E(y) - \beta E(x)$$

- $E(xy) = \int_0^1 \int_0^1 xyf(x,y)dxdy = 3/8$
- $E(x) = \int_0^1 \int_0^1 xf(x,y)dxdy = 5/8$
- $E(y) = \int_0^1 \int_0^1 yf(x,y)dxdy = 5/8$
- $E(x^2) = \int_0^1 \int_0^1 x^2 f(x,y)dxdy = 7/15$

$$\beta = -15/73$$

$$\alpha = 55/73$$

2.16

Now, $E(y|x)$. First calculate the conditional distribution.

$$f(x|y) = \frac{f(x,y)}{f(x)}$$

$$f(x) = \frac{1}{2}(3x^2 + 1)$$

$$f(y|x) = \frac{3(x^2 + y^2)}{(3x^2 + 1)}$$

2.16

Now, $E(y|x)$. First calculate the conditional distribution.

$$f(x|y) = \frac{f(x,y)}{f(x)}$$

$$f(x) = \frac{1}{2}(3x^2 + 1)$$

$$f(y|x) = \frac{3(x^2 + y^2)}{(3x^2 + 1)}$$

Now integrate

$$E(y|x) = \int yf(y|x)dx = \frac{6x^2 + 3}{12x^2 + 4}$$

2.16

Now, $E(y|x)$. First calculate the conditional distribution.

$$f(x|y) = \frac{f(x,y)}{f(x)}$$

$$f(x) = \frac{1}{2}(3x^2 + 1)$$

$$f(y|x) = \frac{3(x^2 + y^2)}{(3x^2 + 1)}$$

Now integrate

$$E(y|x) = \int yf(y|x)dx = \frac{6x^2 + 3}{12x^2 + 4}$$

Are y and x independent?

2.21 Consider the short and long projections

$$y = x\gamma_1 + e$$

$$y = x\beta_1 + x^2\beta_2 + u$$

(a) Under what condition does $\gamma_1 = \beta_1$?

2.21 Consider the short and long projections

$$y = x\gamma_1 + e$$

$$y = x\beta_1 + x^2\beta_2 + u$$

(a) Under what condition does $\gamma_1 = \beta_1$?

$$\gamma = E(xy)/E(x^2)$$

2.21 Consider the short and long projections

$$y = x\gamma_1 + e$$

$$y = x\beta_1 + x^2\beta_2 + u$$

(a) Under what condition does $\gamma_1 = \beta_1$?

$$\gamma = E(xy)/E(x^2)$$

$$\beta = E\left(\begin{bmatrix} x \\ x^2 \end{bmatrix}\begin{bmatrix} x \\ x^2 \end{bmatrix}'\right)^{-1} E\left(\begin{bmatrix} x \\ x^2 \end{bmatrix} y\right) = \begin{bmatrix} E(x^2) & E(x^3) \\ E(x^3) & E(x^4) \end{bmatrix}^{-1} \begin{bmatrix} E(xy) \\ E(x^2y) \end{bmatrix}$$

$$= \frac{1}{E(x^2)E(x^4) - E(x^3)^2} \begin{bmatrix} E(x^4) & -E(x^3) \\ -E(x^3) & E(x^2) \end{bmatrix} \begin{bmatrix} E(xy) \\ E(x^2y) \end{bmatrix}$$

2.21 Consider the short and long projections

$$y = x\gamma_1 + e$$

$$y = x\beta_1 + x^2\beta_2 + u$$

(a) Under what condition does $\gamma_1 = \beta_1$?

$$\gamma = E(xy)/E(x^2)$$

$$\beta = E\left(\begin{bmatrix} x \\ x^2 \end{bmatrix}\begin{bmatrix} x \\ x^2 \end{bmatrix}'\right)^{-1} E\left(\begin{bmatrix} x \\ x^2 \end{bmatrix} y\right) = \begin{bmatrix} E(x^2) & E(x^3) \\ E(x^3) & E(x^4) \end{bmatrix}^{-1} \begin{bmatrix} E(xy) \\ E(x^2y) \end{bmatrix}$$

$$= \frac{1}{E(x^2)E(x^4) - E(x^3)^2} \begin{bmatrix} E(x^4) & -E(x^3) \\ -E(x^3) & E(x^2) \end{bmatrix} \begin{bmatrix} E(xy) \\ E(x^2y) \end{bmatrix}$$

Thus,

$$\beta_1 = \frac{E(xy)E(x^4) - E(x^3)E(x^2y)}{E(x^2)E(x^4) - E(x^3)^2}$$

2.21 Consider the short and long projections

$$y = x\gamma_1 + e$$

$$y = x\beta_1 + x^2\beta_2 + u$$

(a) Under what condition does $\gamma_1 = \beta_1$?

$$\gamma = E(xy)/E(x^2)$$

$$\beta = E\left(\begin{bmatrix} x \\ x^2 \end{bmatrix}\begin{bmatrix} x \\ x^2 \end{bmatrix}'\right)^{-1} E\left(\begin{bmatrix} x \\ x^2 \end{bmatrix} y\right) = \begin{bmatrix} E(x^2) & E(x^3) \\ E(x^3) & E(x^4) \end{bmatrix}^{-1} \begin{bmatrix} E(xy) \\ E(x^2y) \end{bmatrix}$$

$$= \frac{1}{E(x^2)E(x^4) - E(x^3)^2} \begin{bmatrix} E(x^4) & -E(x^3) \\ -E(x^3) & E(x^2) \end{bmatrix} \begin{bmatrix} E(xy) \\ E(x^2y) \end{bmatrix}$$

Thus,

$$\beta_1 = \frac{E(xy)E(x^4) - E(x^3)E(x^2y)}{E(x^2)E(x^4) - E(x^3)^2}$$

2.21

(b) Now suppose the long projection is

$$y = x\theta_1 + x^3\theta_2 + v$$

Is there a similar condition under which $\gamma_1 = \theta_1$?

2.21

(b) Now suppose the long projection is

$$y = x\theta_1 + x^3\theta_2 + v$$

Is there a similar condition under which $\gamma_1 = \theta_1$?

$$\theta = E\left(\begin{bmatrix} x \\ x^3 \end{bmatrix}\begin{bmatrix} x \\ x^3 \end{bmatrix}'\right)^{-1} E\left(\begin{bmatrix} x \\ x^3 \end{bmatrix} y\right) = \begin{bmatrix} E(x^2) & E(x^4) \\ E(x^4) & E(x^6) \end{bmatrix}^{-1} \begin{bmatrix} E(xy) \\ E(x^2y) \end{bmatrix}$$

Thus,

$$\theta_1 = \frac{E(xy)E(x^6) - E(x^4)E(x^2y)}{E(x^2)E(x^6) - E(x^4)^2}$$

2.21

(b) Now suppose the long projection is

$$y = x\theta_1 + x^3\theta_2 + v$$

Is there a similar condition under which $\gamma_1 = \theta_1$?

$$\theta = E\left(\begin{bmatrix} x \\ x^3 \end{bmatrix}\begin{bmatrix} x \\ x^3 \end{bmatrix}'\right)^{-1} E\left(\begin{bmatrix} x \\ x^3 \end{bmatrix} y\right) = \begin{bmatrix} E(x^2) & E(x^4) \\ E(x^4) & E(x^6) \end{bmatrix}^{-1} \begin{bmatrix} E(xy) \\ E(x^2y) \end{bmatrix}$$

Thus,

$$\theta_1 = \frac{E(xy)E(x^6) - E(x^4)E(x^2y)}{E(x^2)E(x^6) - E(x^4)^2}$$

We need $E(x^4) = 0$.

2.21

(b) Now suppose the long projection is

$$y = x\theta_1 + x^3\theta_2 + v$$

Is there a similar condition under which $\gamma_1 = \theta_1$?

$$\theta = E\left(\begin{bmatrix} x \\ x^3 \end{bmatrix}\begin{bmatrix} x \\ x^3 \end{bmatrix}'\right)^{-1} E\left(\begin{bmatrix} x \\ x^3 \end{bmatrix} y\right) = \begin{bmatrix} E(x^2) & E(x^4) \\ E(x^4) & E(x^6) \end{bmatrix}^{-1} \begin{bmatrix} E(xy) \\ E(x^2y) \end{bmatrix}$$

Thus,

$$\theta_1 = \frac{E(xy)E(x^6) - E(x^4)E(x^2y)}{E(x^2)E(x^6) - E(x^4)^2}$$

We need $E(x^4) = 0$. But, $E(x^4) \geq E(x^2)$. So, it's not possible

Chapter 3

- Empirical model, now there is data.
- Use data to approximate the population distribution for y and x.
- Try to "estimate" the conditional expectation $E(y|x)$.

Chapter 3

- Empirical model, now there is data.
- Use data to approximate the population distribution for y and x.
- Try to "estimate" the conditional expectation $E(y|x)$.
- Coming up, we'll show these are good approximation in a certain sense (i.e. consistency)
- We'll evaluate how confident we are about certain statements about the population (i.e. statistical significance/inference)

3.4 Let $\hat{e}$ be the OLS residual from a regression of y on $X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$ Find $X_2'\hat{e}$

3.4 Let $\hat{e}$ be the OLS residual from a regression of y on $X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$ Find $X_2'\hat{e}$

$$\mathbf{min}(Y - X\beta)$$

First order condition (see equation 3.17 in Hansen).

$$X'\hat{e} = 0$$

$$\begin{bmatrix} X_1' \\ X_2' \end{bmatrix} \hat{e} = 0$$

$$\begin{bmatrix} X_1'\hat{e} \\ X_2'\hat{e} \end{bmatrix} = 0$$

3.4 Let $\hat{e}$ be the OLS residual from a regression of y on $X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$ Find $X_2'\hat{e}$

$$\mathbf{min}(Y - X\beta)$$

First order condition (see equation 3.17 in Hansen).

$$X'\hat{e} = 0$$

$$\begin{bmatrix} X_1' \\ X_2' \end{bmatrix} \hat{e} = 0$$

$$\begin{bmatrix} X_1'\hat{e} \\ X_2'\hat{e} \end{bmatrix} = 0$$

Do the dimensions make sense?

3.11 Show that when X contains a constant, $\dfrac{1}{n}\sum_{i=1}^{n}\hat{y}_i = \bar{y}$

3.11 Show that when X contains a constant, $\frac{1}{n}\sum_{i=1}^{n}\hat{y}_i = \bar{y}$

- $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$
- $= \frac{1}{n}\sum_{i=1}^{n} x_i'\hat{\beta} + \hat{e}_i$
- With an intercept $\sum \hat{e}_i = 0$. Thus, we get
- $\frac{1}{n}\sum_{i=1}^{n} x_i'\hat{\beta} = \sum_{i=1}^{n}\hat{y}_i$

3.19 For the intercept-only model $y_i = \beta + e_i$, show that the leave-one-out prediction error is $\tilde{e}_i = \dfrac{n}{n-1}(y_i - \bar{y})$

3.19 For the intercept-only model $y_i = \beta + e_i$, show that the leave-one-out prediction error is $\tilde{e}_i = \dfrac{n}{n-1}(y_i - \bar{y})$

- What is leave-one-out prediction error?
- $\tilde{y}_i = x_i' \hat{\beta}_{-i}$ is the leave-one-out predicted value i.e. we dropped x_i from the regression and computed $\hat{\beta}$. The leave one out prediction error is $\tilde{e}_i = y_i - \tilde{y}_i$

3.19 For the intercept-only model $y_i = \beta + e_i$, show that the leave-one-out prediction error is $\tilde{e}_i = \dfrac{n}{n-1}(y_i - \bar{y})$

- What is leave-one-out prediction error?
- $\tilde{y}_i = x_i'\hat{\beta}_{-i}$ is the leave-one-out predicted value i.e. we dropped x_i from the regression and computed $\hat{\beta}$. The leave one out prediction error is $\tilde{e}_i = y_i - \tilde{y}_i$
- Regression on a constant $\hat{\beta} = \bar{y}$

3.19 For the intercept-only model $y_i = \beta + e_i$, show that the leave-one-out prediction error is $\tilde{e}_i = \dfrac{n}{n-1}(y_i - \bar{y})$

- What is leave-one-out prediction error?
- $\tilde{y}_i = x_i' \hat{\beta}_{-i}$ is the leave-one-out predicted value i.e. we dropped x_i from the regression and computed $\hat{\beta}$. The leave one out prediction error is $\tilde{e}_i = y_i - \tilde{y}_i$
- Regression on a constant $\hat{\beta} = \bar{y}$
- WLOG assume $i = n$, then $\tilde{y}_i = \dfrac{1}{n-1} \sum_{j=1}^{n-1} y_j$
- $\tilde{e}_i = y_i - \tilde{y}_i = \left(\dfrac{n}{n-1} - \dfrac{1}{n-1} \right) y_i - \dfrac{1}{n-1} \sum_{j=1}^{n-1} y_j$
- $= \dfrac{n}{n-1} y_i - \dfrac{1}{n-1} \sum_{j=1}^{n} y_i = \dfrac{n}{n-1} y_i - \dfrac{n}{n-1} \bar{y}$

3.23 The data matrix is (y, X) with $X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$, and consider the transformed regressor matrix $Z = \begin{bmatrix} X_1 & X_2 - X_1 \end{bmatrix}$. Suppose you do a least-squares regression of y on X , and a least-squares regression of y on Z. Let $\hat{\sigma}^2$ and $\tilde{\sigma}^2$ denote the residual variance estimates from the two regressions. Give a formula relating $\hat{\sigma}^2$ and $\tilde{\sigma}^2$?

3.23 The data matrix is (y, X) with $X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$, and consider the transformed regressor matrix $Z = \begin{bmatrix} X_1 & X_2 - X_1 \end{bmatrix}$. Suppose you do a least-squares regression of y on X , and a least-squares regression of y on Z. Let $\hat{\sigma}^2$ and $\tilde{\sigma}^2$ denote the residual variance estimates from the two regressions. Give a formula relating $\hat{\sigma}^2$ and $\tilde{\sigma}^2$? Use the first order conditions (again)

3.23 First order conditions...

3.23 First order conditions...

- $\begin{bmatrix} X_1' \\ X_2' \end{bmatrix} \hat{e} = 0$

3.23 First order conditions...

- $\begin{bmatrix} X_1' \\ X_2' \end{bmatrix} \hat{e} = 0$
- From the first set of restrictions, $X_1'\hat{e} = 0$ and $X_2'\hat{e} = 0$. Thus, $\hat{\beta}_1$ and $\hat{\beta}_2$ solve:

$$X_1'(Y - \hat{\beta}_1 X_1 - \hat{\beta}_2 X_2) = 0$$

$$X_2'(Y - \hat{\beta}_1 X_1 - \hat{\beta}_2 X_2) = 0$$

3.23 First order conditions...

- $\begin{bmatrix} X_1' \\ X_2' \end{bmatrix} \hat{e} = 0$
- From the first set of restrictions, $X_1'\hat{e} = 0$ and $X_2'\hat{e} = 0$. Thus, $\hat{\beta}_1$ and $\hat{\beta}_2$ solve:

$$X_1'(Y - \hat{\beta}_1 X_1 - \hat{\beta}_2 X_2) = 0$$

$$X_2'(Y - \hat{\beta}_1 X_1 - \hat{\beta}_2 X_2) = 0$$

- $\begin{bmatrix} X_1' \\ X_2' - X_1' \end{bmatrix} \tilde{e} = 0$

3.23

- For the second set of restrictions, $X_1'\tilde{e} = 0$.
- The second condition $(X_2 - X_1)'\tilde{e} = 0$, reduces to $X_2'\tilde{e} = 0$
- Thus we get

$$X_1'(Y - \tilde{\beta}_1 X_1 - \tilde{\beta}_2(X_2 - X_1)) = X_1'(Y - (\tilde{\beta}_1 - \tilde{\beta}_2)X_1 - \tilde{\beta}_2 X_2) = 0$$

$$X_2'(Y - \hat{\beta}_1 X_1 - \hat{\beta}_2(X_2 - X_1)) = X_2'(Y - (\tilde{\beta}_1 - \tilde{\beta}_2)X_1 - \tilde{\beta}_2 X_2) = 0$$

3.23

- For the second set of restrictions, $X_1'\tilde{e} = 0$.
- The second condition $(X_2 - X_1)'\tilde{e} = 0$, reduces to $X_2'\tilde{e} = 0$
- Thus we get

$$X_1'(Y - \tilde{\beta}_1 X_1 - \tilde{\beta}_2(X_2 - X_1)) = X_1'(Y - (\tilde{\beta}_1 - \tilde{\beta}_2)X_1 - \tilde{\beta}_2 X_2) = 0$$

$$X_2'(Y - \hat{\beta}_1 X_1 - \hat{\beta}_2(X_2 - X_1)) = X_2'(Y - (\tilde{\beta}_1 - \tilde{\beta}_2)X_1 - \tilde{\beta}_2 X_2) = 0$$

- k equations and k unknowns in both cases so $\hat{\beta}_1 = (\tilde{\beta}_1 - \tilde{\beta}_2)$, $\hat{\beta}_2 = \tilde{\beta}_2$, $\tilde{e} = \hat{e}$ and the variances are the same