AS.180.633: Econometrics **Spring 2020**

Homework 1: Suggested Solutions

Instructor: *Yingyao Hu* *By:* *Tong Zhou*

2.15

Consider the intercept-only model $y = \alpha + e$ defined as the best linear predictor. Show that $\alpha = \mathbb{E}[y]$.

Proof. Model $y = \alpha + e$ is a special case of the *linear CEF Model* $y = \boldsymbol{x}'\boldsymbol{\beta} + e$, where $\boldsymbol{\beta} = \alpha \in \mathbb{R}$ and $\boldsymbol{x} = 1$. Then by (2.21) on page 37 of Hansen's book, $\boldsymbol{\beta} = \left(\mathbb{E}[\boldsymbol{x}\boldsymbol{x}']\right)^{-1}\mathbb{E}[\boldsymbol{x}y]$ translates into $\alpha = \mathbb{E}[y]$. $\square$

2.16

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

Proof. Following Theorem 2.9 of Hansen's book, we have coefficients of the best linear predictor:

$$\begin{aligned}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} &= \left\{\mathbb{E}\left[\begin{pmatrix}1\\ x\end{pmatrix}\begin{pmatrix}1 & x\end{pmatrix}\right]\right\}^{-1}\mathbb{E}\left[\begin{pmatrix}1\\ x\end{pmatrix}y\right]\\ &= \begin{bmatrix}1 & \mathbb{E}(x)\\ \mathbb{E}(x) & \mathbb{E}(x^2)\end{bmatrix}^{-1}\begin{bmatrix}\mathbb{E}(y)\\ \mathbb{E}(xy)\end{bmatrix}\end{aligned} \tag{1}$$

From the given joint distribution of x and y, we obtain $\mathbb{E}(x) = \mathbb{E}(y) = 5/8$, $\mathbb{E}(x^2) = 7/15$ and $\mathbb{E}(xy) = 3/8$. Plugging those values and computing the inverse of the matrix in Eq. (1), we obtain coefficients:

$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix}55/73\\ -15/73\end{pmatrix},$$

and the best linear predictor

$$\mathcal{P}(y|x) = \frac{55}{73} - \frac{15}{73}x, \quad x \in [0,1]. \tag{2}$$

The conditional mean function

$$m(x) \equiv \mathbb{E}(y|x) = \int_0^1 y f(y|x)\,\mathrm{d}y = \int_0^1 y \frac{f(x,y)}{f(x)}\,\mathrm{d}y,$$

where $f(x,y)$ is given, the p.d.f. of x is $f(x) = \int_0^1 f(x,y)\,\mathrm{d}y = \frac{1}{2}\left(3x^2+1\right), x \in [0,1]$.

Then it is easy to show

$$m(x) = \frac{6x^2+3}{12x^2+4}, \quad x \in [0,1]. \tag{3}$$

Staring at Eq. (2) and Eq. (3), we know $\mathcal{P}(y|x)$ and $m(x)$ do not coincide.

Their discrepancy arises from the fact that the best linear predictor $\mathcal{P}(y|x)$, by construction, places a linear restriction on the relationship between y and x, implying only their first and second moments matter, whereas the conditional mean function $m(y|x)$ does not rely on any specific parametric restriction and therefore makes full use of information on the joint distribution of x and y. □

Remark. For a 2-by-2 matrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, where $a_{ij} \in \mathbb{R}$, you may use the following formula to compute its inverse if it is invertible:

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} = \frac{1}{(a_{11}a_{22} - a_{12}a_{21})} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

2.19

Show (2.46) – (2.47), namely that for

$$d(\boldsymbol{\beta}) = \mathbb{E}\left[(m(\boldsymbol{x}) - \boldsymbol{x}'\boldsymbol{\beta})^2\right]$$

then

definition:

ddd

$$\begin{aligned}\boldsymbol{\beta} &= \underset{\boldsymbol{b}\in\mathbb{R}^k}{\operatorname{argmin}}\ d(\boldsymbol{b}) \\ &= \left(\mathbb{E}\left[\boldsymbol{x}\boldsymbol{x}'\right]\right)^{-1}\mathbb{E}[\boldsymbol{x}m(\boldsymbol{x})] \\ &= \left(\mathbb{E}\left[\boldsymbol{x}\boldsymbol{x}'\right]\right)^{-1}\mathbb{E}\left[\boldsymbol{x}y\right].\end{aligned}$$

Proof. By definition,

$$\begin{aligned}d(\boldsymbol{\beta}) &= \mathbb{E}\left[\left(m(\boldsymbol{x}) - \boldsymbol{x}'\boldsymbol{\beta}\right)^2\right] \\ &= \mathbb{E}\left[m^2(\boldsymbol{x})\right] + \boldsymbol{\beta}'\mathbb{E}\left[\boldsymbol{x}\boldsymbol{x}'\right]\boldsymbol{\beta} - 2\boldsymbol{\beta}'\mathbb{E}\left[\boldsymbol{x}m(\boldsymbol{x})\right].\end{aligned}$$

FOC w.r.t. $\boldsymbol{\beta}$ gives

$$\mathbf{0} = \frac{\partial}{\partial\boldsymbol{\beta}}d(\boldsymbol{\beta}) = 2\mathbb{E}[\boldsymbol{x}\boldsymbol{x}']\boldsymbol{\beta} - 2\mathbb{E}[\boldsymbol{x}m(\boldsymbol{x})].$$

Therefore,

$$\boldsymbol{\beta} = \left(\mathbb{E}[\boldsymbol{x}\boldsymbol{x}']\right)^{-1}\mathbb{E}\left[\boldsymbol{x}m(\boldsymbol{x})\right]. \tag{4}$$

By the *Law of Iterated Expectations*, we know

$$\begin{aligned}\mathbb{E}[\boldsymbol{x}y] &= \mathbb{E}[\mathbb{E}[\boldsymbol{x}y|\boldsymbol{x}]] \\ &= \mathbb{E}[\boldsymbol{x}\underbrace{\mathbb{E}[y|\boldsymbol{x}]}_{\equiv m(\boldsymbol{x})}]. \\ &= \mathbb{E}[\boldsymbol{x}m(\boldsymbol{x})].\end{aligned}$$

Plugging the above equality into Eq. (4) yields

$$\boldsymbol{\beta} = \left(\mathbb{E}[\boldsymbol{x}\boldsymbol{x}']\right)^{-1} \mathbb{E}\left[\boldsymbol{x}y\right]. \tag{5}$$

So, expressions Eq. (4) and Eq. (5) are equivalent. □