SOLUTIONS - CHAPTER 2

ZHIZHONG PU

Exercise 2. 1. Find $\mathbb{E}[\mathbb{E}[\mathbb{E}[Y|X_1, X_2, X_3]|X_1, X_2]|X_1]$.

Since $\mathbb{E}[\mathbb{E}[Y|X_1, X_2, X_3]|X_1, X_2] = \mathbb{E}[Y|X_1, X_2]$ by LIE, then we have:

$$\mathbb{E}[\mathbb{E}[\mathbb{E}[Y|X_1, X_2, X_3]|X_1, X_2]|X_1] = \mathbb{E}[\mathbb{E}[Y|X_1, X_2]|X_1] = \mathbb{E}[Y|X_1]$$

Exercise 2. 2. If $\mathbb{E}[Y|X] = a + bX$, find $\mathbb{E}[YX]$ as a function of moments of X.

Since $\mathbb{E}[Y|X] = a + bX$, we follow chapter 2.8 to define CEF error:

$$e = Y - \mathbb{E}[Y|X] = Y - a + bX \Rightarrow Y = a + bX + e$$

Then substitute Y into $\mathbb{E}[YX]$:

$$\mathbb{E}[YX] = \mathbb{E}[(a + bX + e)X] = a\mathbb{E}[X] + b\mathbb{E}[X^2] + \mathbb{E}[eX]$$

Note that the last term is 0 by Theorem 2.4.4, then

$$\mathbb{E}[YX] = \mathbb{E}[(a + bX + e)X] = b\mathbb{E}[X^2] + a\mathbb{E}[X]$$

Exercise 2. 3. Prove Theorem 2.4.4 using the law of iterated expectations: for any function $h(x)$ such that $\mathbb{E}|h(X)e| < \infty$ then $\mathbb{E}[h(X)e] = 0$

Proof. Wooldridge (2010) CE.6 Proof

Given any function $h()$, we can write

$$\mathbb{E}[h(X)e] = \mathbb{E}[\mathbb{E}(h(X)e|X)] = \mathbb{E}[h(X)\mathbb{E}(e|X)] = \mathbb{E}[h(X)0] = 0$$

using LIE, take-out-what's-known property, and $\mathbb{E}(e|X) = 0$ from Theorem 2.4.1, respectively. $\square$

Date: Feb 2023.

Exercise 2. 4. See book

By the definition of conditional expectation:

$$\mathbb{E}[Y|X=0] = \sum_{y\in 0,1} yP(Y=y|X=0) = 1P(Y=1|X=0) = 0.8$$

$$\mathbb{E}[Y|X=1] = \sum_{y\in 0,1} yP(Y=y|X=1) = 1P(Y=1|X=1) = 0.6$$

(Note that LIE can be verified: $\mathbb{E}[\mathbb{E}[Y|X]] = 0.7 = \mathbb{E}[Y]$)

By LOTUS:

$$\mathbb{E}[Y^2|X=0] = \sum_{y\in 0,1} y^2P(Y=y|X=0) = 1P(Y=1|X=0) = 0.8$$

$$\mathbb{E}[Y^2|X=1] = \sum_{y\in 0,1} y^2P(Y=y|X=1) = 1P(Y=1|X=1) = 0.6$$

By definition of conditional variance:

$$Var(Y|X=0) = \mathbb{E}[Y^2|X=0] - (\mathbb{E}[Y|X=0])^2 = 0.8 - (0.8)^2 = 0.16$$

$$Var(Y|X=1) = \mathbb{E}[Y^2|X=1] - (\mathbb{E}[Y|X=1])^2 = 0.6 - (0.6)^2 = 0.24$$

Exercise 2. 5. Show that $\sigma^2(X)$ is the best predictor of e^2 given X

According to Theorem 2.7: for any random variable, the best linear predictor given X is its CEF given X. In this case:

$$\begin{aligned}
BLP[e^2|X] = \mathbb{E}[e^2|X] &= \mathbb{E}[Y^2 - 2Y\mathbb{E}[Y|X] + \mathbb{E}[Y|X]^2|X] \\
&= \mathbb{E}[Y^2|X] - 2\mathbb{E}[\mathbb{E}(Y|X)Y|X] + \mathbb{E}[Y|X]^2 \\
&= \mathbb{E}[Y^2|X] - 2\mathbb{E}[Y|X]^2 + \mathbb{E}[Y|X]^2 \\
&= Var(Y|X)
\end{aligned}$$

Exercise 2. 6. Use $Y = m(X) + e$ to show that $Var(Y) = Var[(m(X)] + \sigma^2$.

*Note on notation: $\sigma^2 = \mathbb{E}[e^2]$.

Consider the LHS. By definition of variance:

$$Var(Y) = Var[m(X) + e] = \mathbb{E}[(m(X) + e)^2] - (\mathbb{E}[m(X) + e])^2$$

The first term is (using $\mathbb{E}[m(X)e] = 0$ from Theorem 2.4.4.):

$$\mathbb{E}[(m(X) + e)^2] = \mathbb{E}[m^2(X)] + 2\mathbb{E}[m(X)e] + \mathbb{E}[e^2] = \mathbb{E}[m^2(X)] + \mathbb{E}[e^2]$$

The second term is (using 0 unconditional mean of error term):

$$-(\mathbb{E}[m(X) + e])^2 = -(\mathbb{E}[m(X)] + \mathbb{E}[e])^2 = -(\mathbb{E}[m(X)] + 0)^2 = -\mathbb{E}[m(X)]^2$$

Sum up the two terms:

$$Var(Y) = \mathbb{E}[(m(X)+e)^2]-(\mathbb{E}[m(X)+e])^2 = \mathbb{E}[m^2(X)]+\mathbb{E}[e^2]-\mathbb{E}[m(X)]^2 = Var[(m(X)]+\sigma^2$$

Exercise 2. 7. Show that the conditional variance can be written as $\sigma^2(X) = \mathbb{E}[Y^2|X] - \mathbb{E}[Y|X]^2$

By definition 2.1, we can write

$$\begin{aligned}\sigma^2(X) \equiv Var(Y|X) &= \mathbb{E}[(Y - \mathbb{E}(Y|X))^2|X] \\ &= \mathbb{E}[Y^2 - 2Y\mathbb{E}(Y|X) + \mathbb{E}(Y|X)^2|X] \\ &= \mathbb{E}[Y^2|X] - 2\mathbb{E}[Y\mathbb{E}(Y|X)|X] + \mathbb{E}[\mathbb{E}(Y|X)^2|X]\end{aligned}$$

Note that in the last two terms, $\mathbb{E}(Y|X)$ and $\mathbb{E}(Y|X)^2$ are both functions of X, so we can apply take-out-what's-known property:

$$= \mathbb{E}[Y^2|X] - 2\mathbb{E}(Y|X)\mathbb{E}(Y|X) + \mathbb{E}(Y|X)^2\mathbb{E}[1|X] = \mathbb{E}[Y^2|X] - \mathbb{E}(Y|X)^2$$

Exercise 2. 8.

Exercise 2. 9.

Exercise 2. 10. True or False. If $Y = X\beta + e$, $X \in \mathbb{R}$, and $\mathbb{E}[e|X] = 0$, then $\mathbb{E}[X^2 e] = 0$.

True.

$$\mathbb{E}[X^2 e] = \mathbb{E}[\mathbb{E}[X^2 e|X]] = \mathbb{E}[X^2 \mathbb{E}[e|X]] = \mathbb{E}[0X^2] = 0$$

Exercise 2. 11.

Exercise 2. 12.

Exercise 2. 13.

Exercise 2. 14.

Exercise 2. 15. Consider the intercept-only model $Y = a + e$ with a the best linear predictor. Show that $a = \mathbb{E}[Y]$.

We can write MSE as:

$$\mathbb{E}[e^2] = \mathbb{E}[(Y-a)^2] = \mathbb{E}[Y^2] - 2a\mathbb{E}[Y] + a^2$$

Apply the first order condition:

$$\frac{\partial}{\partial a}\mathbb{E}[e^2] = -2\mathbb{E}[Y] + 2a \equiv 0$$

This shows $a = \mathbb{E}[Y]$.

Exercise 2. 16. I do not recommend this problem for calculations are overly complicated.

$$\beta = E(XX')^{-1}(XY) = \frac{\mathbb{E}(XY)}{\mathbb{E}(X^2)} = \frac{7}{16} \cdot \frac{15}{7} = \frac{15}{16}$$

$$\mathbb{E}[Y|X] = \int_0^1 y \cdot f_{Y|X}(y|x)\, dy = \int_0^1 y \frac{x^2 + y^2}{x^2 + \frac{1}{3}}\, dy = \frac{6x^2 + 3}{12x^2 + 4}$$

Exercise 2. 17.

Exercise 2. 18. Suppose that $X = (1, X_2, X_3)$ where $X_3 = a_1 + a_2 X_2$ is a linear function of X_2.

(a) Show that $\mathbb{E}[XX']$ is not invertible.

$$\mathbb{E}[XX'] = \mathbb{E}\begin{bmatrix} 1 & X_2 & X_3 \\ X_2 & X_2^2 & X_2X_3 \\ X_3 & X_2X_3 & X_3^2 \end{bmatrix}$$

Note that the columns are linearly dependent:

$$c_3 = a_1c_1 + a_2c_2$$

Then the matrix is singular.

(b) Use a linear transformation of X to find an expression for the best linear predictor of Y given X. (Be explicit, do not just use the generalized inverse formula.)

Define a linear transformation of X removing X_3:

$$Z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \Rightarrow ZX = (1, X_2)$$

Note that vector $\mathbb{E}(ZXX'Z')$ is now invertible, so the formula for β now applies:

$$BLP[Y|X] = BLP[Y|ZX] = ZX\beta = (ZX)'\mathbb{E}[ZXX'Z']^{-1}\mathbb{E}[ZXY] = X'Z'\mathbb{E}\begin{bmatrix} 1 & X_2 \\ X_2 & X_2^2 \end{bmatrix}\mathbb{E}\begin{bmatrix} Y \\ X_2Y \end{bmatrix}$$

Exercise 2. 19. Show that the best linear approximation of $m(X) = \mathbb{E}[Y|X]$ has coefficients $\beta = \mathbb{E}[XX']^{-1}\mathbb{E}[Xm(X)] = \mathbb{E}[XX']^{-1}\mathbb{E}[XY]$

First, write MSE:

$$d(\beta) = \mathbb{E}[(m(X) - X)^2] = \mathbb{E}[m(X)^2] - \mathbb{E}[m(X)'X'\beta] - \mathbb{E}[\beta'Xm(X)] + \mathbb{E}[\beta'XX'\beta]$$

Then by first order condition and LIE:

$$\beta = \mathbb{E}[XX']^{-1}\mathbb{E}[Xm(X)] = \mathbb{E}[XX']^{-1}\mathbb{E}[X\mathbb{E}(Y|X)] = \mathbb{E}[XX']^{-1}\mathbb{E}[\mathbb{E}(XY|X)] = \mathbb{E}[XX']^{-1}\mathbb{E}[XY]$$

Exercise 2. 20.

Exercise 2. 21.

Exercise 2. 22.