Problem Set 1

Nicholas Wu

Spring 2021

Problems

(2.2) By the law of iterated expectation

$$E[YX] = E[E[YX|X]] = E[XE[Y|X]] = E[aX + bX^2] = aE[X] + bE[X^2]$$

(2.4)

$$E[Y|X = 0] = 0.8$$

$$E[Y^2|X = 0] = 0.8$$

$$V[Y|X = 0] = E[Y^2|X = 0] - (E[Y|X = 0])^2 = 0.8 - 0.8^2 = 0.16$$

$$E[Y|X = 1] = E[Y^2|X = 1] = 0.6$$

$$V[Y|X = 1] = E[Y^2|X = 0] - (E[Y|X = 0])^2 = 0.6 - 0.6^2 = 0.24$$

(2.5)

(a)

$$\mathbb{E}[(h(X) - e^2)^2|X]$$

(b) Predicting e^2 means we want to minimize the MSE between $h(X)$ and e^2.

(c) The FOC for maximization requires

$$\mathbb{E}[(h(X) - e^2)^2|X] = E[h(X)^2 - 2h(X)e^2 + e^4|X]$$

By the law of iterated expectation

$$= h(X)^2 - 2h(X)\sigma(X)^2 + E(e^4|X)$$

The FOC for minimization requires

$$2h(X) - 2\sigma(X)^2 = 0$$

$$h(X) = \sigma^2(X)$$

thus $\sigma^2(X)$ minimizes the MSE and is the best predictor.

(2.6)

$$V(Y) = E((m(X)+e)^2) - (E(m(X)+e))^2 = E(m(X)^2 + 2m(X)e + e^2) - E(m(X))^2 - 2E(m(X))E(e) - E(e)^2$$

$$= E(m(X)^2) + 2E(m(X)e) + E(e^2) - E(m(X))^2 - 2E(m(X))E(e) - E(e)^2$$

$$= E(m(X)^2) - E(m(X))^2 + E(e^2) - E(e)^2$$

$$= V(m(X)) + \sigma^2$$

since $E(e) = E(m(X)e) = 0$.

(2.8)

$$\mathbb{E}[Y|X] = X'\beta$$

$$\mathbb{V}(Y|X) = X'\beta$$

(2.10) True. By the law of iterated expectation:

$$E[X^2 e] = E[E[X^2 e|X]] = E[X^2 E[e|X]] = E[X^2 \cdot 0] = 0$$

(2.11) False. Suppose (X, e) has a joint distribution 0.5 probability of $(1, -1)$ and 0.5 probability of $(-1, -1)$. Then $E[Xe] = 0$ but $E[X^2 e] = -1$.

(2.12) False. Consider the joint distribution of (X, e) with 0.5 probability of $(-1, 0)$, and 0.25 probability each of $(1, -1)$ and $(1, 1)$. Then $E[e|X = 1] = 0$, $E[e|X = -1] = 0$. However, e and X are clearly not independent.

(2.13) False. See the 2.11 counterexample I gave. $E[e|X] = -1$, not 0.

(2.14) False. Conwsider (X, e) with joint distribution: 0.25 probability of $(1, 1)$, 0.25 probability of $(1, -1)$, 0.125 probability of $(-1, \sqrt{2})$, 0.125 probability of $(-1, -\sqrt{2})$, 0.25 probability of $(-1, 0)$. Clearly, X, e are not independent ($e \neq 0$ if $X = 1$, but e can be zero if $X = -1$). However, $E(e|X) = 0$ and $E(e^2|X) = 1$ (we can check this by hand for the two values of X).

(2.15) The best linear predictor minimizes

$$E((y - \alpha)^2)$$

$$= E(y^2 - 2\alpha y + \alpha^2)$$

$$= E(y^2) - 2\alpha E(y) + \alpha^2$$

$$= E(y^2) - E(y)^2 + E(y)^2 - 2\alpha E(y) + \alpha^2$$

$$= V(y) + (E(y) - \alpha)^2$$

Since the variance is always positive and the expression $(E(y) - \alpha)^2$ is also nonnegative, this is minimized when $E(y) = \alpha$.

(2.16) The best linear predictor minimizes

$$E[(Y - \alpha - \beta X)^2]$$

$$= E[Y^2 - 2(\alpha + \beta X)Y + (\alpha + \beta X)^2]$$

$$= E[Y^2] - 2\alpha E[Y] - 2\beta E[XY] + \alpha^2 + 2\alpha\beta E[X] + \beta^2 E[X^2]$$

Dropping terms irrelevant to α, β, we have

$$= -2\alpha E[Y] - 2\beta E[XY] + \alpha^2 + 2\alpha\beta E[X] + \beta^2 E[X^2]$$

The FOCs yield:

$$0 = -2E[Y] + 2\alpha + 2\beta E[X]$$

$$0 = -2E[XY] + 2\alpha E[X] + 2\beta E[X^2]$$

$$0 = -E[Y] + \alpha + \beta E[X]$$

$$\alpha = E[Y] - \beta E[X]$$

Plugging in for α,

$$0 = -E[XY] + (E[Y] - \beta E[X])E[X] + \beta E[X^2]$$

$$E[XY] - E[Y]E[X] = \beta(E[X^2] - E[X]^2)$$

$$\beta = \frac{E[XY] - E[Y]E[X]}{E[X^2] - E[X]^2} = \frac{(3/8) - (5/8)^2}{(7/15) - (5/8)^2} = \frac{-15}{73}$$

$$\alpha = E[Y] - \beta E[X] = \frac{55}{73}$$

The conditional expectation function is

$$\int_0^1 y \frac{f(x,y)}{f_X(x)} dy = \frac{3}{4}\left(\frac{1 + 2x^2}{1 + 3x^2}\right)$$

Note that $\alpha + \beta x$ is a good approximation to this on the interval $[0, 1]$.

(2.17) Then

$$E[g(X, m, s)] = E\begin{pmatrix} X - m \\ (X - m)^2 - s \end{pmatrix}$$

$$= \begin{pmatrix} E[X - m] \\ E[(X - m)^2 - s] \end{pmatrix}$$

$$= \begin{pmatrix} E[X] - m \\ E[(X-m)^2] - s \end{pmatrix}$$

This is only 0 if both vector entries are 0. This is true iff $E[X] - m = 0$, which implies $m = \mu$. Further, $E[(X-\mu)^2] - s = 0$ is true iff $E[(X-\mu)^2] = s$, or $s = \sigma^2$ Hence, $E[g(X, m, s)] = 0$ iff $m = \mu, s = \sigma^2$.