Løsninger til problem sæt 0– BA-BMECV1031U Økonometri

Fundamentals of mathematical statistics

1 (a) This is just a special case of what is presented in the textbook, with $n = 4$: $E(\bar{Y}) = \mu$ and $Var(\bar{Y}) = \sigma^2/4$.

(b) $E(W) = E(Y_1)/8 + E(Y_2)/8 + E(Y_3)/4 + E(Y_4)/2 = \mu[(1/8) + (1/8) + (1/4) + (1/2)] = \mu(1 + 1 + 2 + 4)/8 = \mu$, which shows that W is unbiased. Because the Y_i are independent,
$Var(W) = Var(Y_1)/64 + Var(Y_2)/64 + Var(Y_3)/16 + Var(Y_4)/4$
$= \sigma^2[(1/64) + (1/64) + (4/64) + (16/64)] = \sigma^2(22/64) = \sigma^2(11/32)$.

(c) Because $11/32 > 8/32 = 1/4$, $Var(W) > Var(\bar{Y})$ for any $\sigma^2 > 0$, so $\bar{Y}$ is preferred to W because each is unbiased.

2 (a) $E(W_a) = a_1E(Y_1) + a_2E(Y_2) + + a_nE(Y_n) = (a_1 + a_2 + + a_n)\mu$. Therefore, we must have $a_1 + a_2 + + a_n = 1$ for unbiasedness.

(b) $Var(W_a) = a_1^2\,Var(Y_1) + a_2^2\,Var(Y_2) + \ldots + a_n^2\,Var(Y_n) = (a_1^2 + a_2^2 + \ldots + a_n^2)\sigma^2$.

(c) From the hint, when $a_1 + a_2 + + a_n = 1$ – the condition needed for unbiasedness of W_a – we have $1/n \leq a_1^2 + a_2^2 + \ldots + a_n^2$. But then $Var(\bar{Y}) = \sigma^2/n \leq \sigma^2(a_1^2 + a_2^2 + \ldots + a_n^2) = Var(W_a)$.

3 Since $E(Y) = a + bE(X)$,
$E[(Y-E(Y))^2] = E\{[a+bX-a-bE(X)]^2\}$
$= E\{b^2[X-E(X)]^2\}$
$= b^2E\{[X-E(X)]^2\}$
$= b^2Var(X)$

4 (a) $E(W_1) = [(n-1)/n]E(\bar{Y}) = [(n-1)/n]\mu$, and so $Bias(W_1) = [(n-1)/n]\mu - \mu = -\mu/n$.
Similarly, $E(W_2) = E(\bar{Y})/2 = \mu/2$, and so $Bias(W_2) = \mu/2 - \mu = -\mu/2$. The bias in W_1 tends to zero as $n \to \infty$, while the bias in W_2 is $-\mu/2$ for all n. This is an important difference.

(b) $plim(W_1) = plim[(n-1)/n] \cdot plim(\bar{Y}) = 1 \cdot \mu = \mu$. $plim(W_2) = plim(\bar{Y})/2 = \mu/2$. Because $plim(W_1) = \mu$ and $plim(W_2) = \mu/2$, W_1 is consistent whereas W_2 is inconsistent.
(c) $Var(W_1) = [(n-1)/n]^2Var(\bar{Y}) = [(n-1)^2/n^3]\sigma^2$ and $Var(W_2) = Var(\bar{Y})/4 = \sigma^2/(4n)$.

(d) Because $\bar{Y}$ is unbiased, its mean squared error is simply its variance. On the other hand, $MSE(W_1) = Var(W_1) + [Bias(W_1)]^2 = [(n-1)^2/n^3]\sigma^2 + \mu^2/n^2$. When $\mu = 0$, $MSE(W_1) = Var(W_1) = [(n-1)^2/n^3]\sigma^2 < \sigma^2/n = Var(\bar{Y})$ because $(n-1)/n < 1$. Therefore, $MSE(W_1)$ is smaller than $Var(\bar{Y})$ for μ close to zero. For large n, the difference between the two estimators is trivial.

5 (a) Using the hint, $E(Z|X) = E(Y/X|X) = E(Y|X)/X = \theta X/X = \theta$. It follows by Property CE.4, the law of iterated expectations, that $E(Z) = E(\theta) = \theta$.

(b) This follows from part (a) and the fact that the sample average is unbiased for the population average: write

$$W_1 = n^{-1}\sum_{i=1}^{n}(Y_i / X_i) = n^{-1}\sum_{i=1}^{n} Z_i,$$

where $Z_i = Y_i/X_i$. From part (a), $E(Z_i) = \theta$ for all i.

(c) In general, the average of the ratios, Y_i/X_i, is not the ratio of averages, $W_2 = \bar{Y} / \bar{X}$. Nevertheless, W_2 is also unbiased, as a simple application of the law of iterated expectations shows. First, $E(Y_i|X_1,\ldots,X_n) = E(Y_i|X_i)$ under random sampling because the observations are independent. Therefore, $E(Y_i|X_1,\ldots,X_n) = \theta X_i$ and so

$$E(\bar{Y} \mid X_1,\ldots,X_n) = n^{-1}\sum_{i=1}^{n} E(Y_i \mid X_1,\ldots,X_n) = n^{-1}\sum_{i=1}^{n}\theta X_i$$
$$= \theta n^{-1}\sum_{i=1}^{n} X_i = \theta \bar{X}.$$

Therefore, $E(W_2 \mid X_1,\ldots,X_n) = E(\bar{Y} / \bar{X} \mid X_1,\ldots,X_n) = \theta \bar{X} / \bar{X} = \theta,$ which means that W_2 is actually unbiased conditional on $(X_1,\ldots,X_n)$, and therefore also unconditionally unbiased.

6 (a) While the expected value of the numerator of G is $E(\bar{Y}) = \theta$, and the expected value of the denominator is $E(1 - \bar{Y}) = 1 - \theta$, the expected value of the ratio is not the ratio of the expected value.

(b) By Property PLIM.2(iii), the plim of the ratio is the ratio of the plims (provided the plim of the denominator is not zero): $\text{plim}(G) = \text{plim}[\bar{Y}/(1 - \bar{Y})] = \text{plim}(\bar{Y})/[1 - \text{plim}(\bar{Y})] = \theta/(1 - \theta) = \gamma$.

Matrix algebra

7 **A** is a square matrix, **B** is a diagonal matrix and **C** is an identity matrix.

8 (a) $\mathbf{AB} = \begin{pmatrix} 2 & -1 & 7 \\ -4 & 5 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & 6 \\ 1 & 8 & 0 \\ 3 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 20 & -6 & 12 \\ 5 & 36 & -24 \end{pmatrix}$

(b) **BA** does not exist because **B** is 3×3 and **A** is 2×3.

9

$$\boldsymbol{AB} = \begin{pmatrix} 2 & 4 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 4 \\ -3 & -2 \end{pmatrix}$$

$$\boldsymbol{BA} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} -2 & 4 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix}$$

10 An element (i,j) of **AB: sum_k=1^n** $\mathbf{a}_{ik}\,\mathbf{b}_{ki}$ (or **BA)** is only non zero if i=j. Only the element k=i of the sum can be nonzero. Thus it is a diagonal matrix with element (i,i): $\mathbf{a}_{ii}\,\mathbf{b}_{ii.}$

11
A symmetric matrix is quadratic and equals it self when transposed, **A**'=**A**

$$\boldsymbol{A} = \begin{pmatrix} 1 & 3 & 2 \\ 3 & 5 & 0 \\ 2 & 0 & 4 \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} 3 & 1 \\ 4 & 5 \end{pmatrix}$$

A is symmetric in the example, but **B** is not since **B'**≠**B**

12 Using the basic rules for transpose, **(X'X)'=(X')(X')'=X'X** which is what we wanted to show.

13 (a) **A'A** is symmetric (m x m). **AA'** is symmetric (n x n). Note that $\mathbf{a}_{ij} = \mathbf{a'}_{ji.}$

(AA')$_{ii}$**=sum_(k=1)^m** $\mathbf{a}_{ik}\,\mathbf{a'}_{ki}$ **= sum_(k=1)^m** $\mathbf{a}_{ik}$^2

(A'A)$_{ii}$**=sum_(i=1)^n** $\mathbf{a'}_{ki}\,\mathbf{a}_{ik}$ **= sum_(i=1)^n** $\mathbf{a}_{ik}$^2

And thus

tr(AA')= sum_(i=1)^n sum_(k=1)^m $\mathbf{a}_{ik}$^2

tr(A'A)= sum_(k=1)^m sum_(i=1)^n $\mathbf{a}_{ik}$^2

which are the same.

(b) just plug in values of **A**

14 (a) The $n \times n$ matrix **C** is the inverse of **AB** if and only if **C**(**AB**) = $\mathbf{I}_n$ and (**AB**)**C** = $\mathbf{I}_n$. We verify both of these equalities for $\mathbf{C} = \mathbf{B}^{-1}\mathbf{A}^{-1}$. First, $(\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{AB}) = \mathbf{B}^{-1}(\mathbf{A}^{-1}\mathbf{A})\mathbf{B} = \mathbf{B}^{-1}\mathbf{I}_n\mathbf{B} = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I}_n$. Similarly, $(\mathbf{AB})(\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{A}(\mathbf{BB}^{-1})\mathbf{A}^{-1} = \mathbf{AI}_n\mathbf{A}^{-1} = \mathbf{AA}^{-1} = \mathbf{I}_n$.

(b) $(\mathbf{ABC})^{-1} = (\mathbf{BC})^{-1}\mathbf{A}^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}$.

15 We must show that, for any $n \times 1$ vector **x**, $\boldsymbol{x} \neq \mathbf{0}$, $\mathbf{x}'(\mathbf{P}'\mathbf{AP})\,\mathbf{x} > 0$. But we can write this quadratic form as $(\mathbf{P}\,\mathbf{x})'\mathbf{A}(\mathbf{P}\,\mathbf{x}) = \mathbf{z}'\mathbf{Az}$ where $\mathbf{z} \equiv \mathbf{Px}$. Because **A** is positive definite by assumption, $\mathbf{z}'\mathbf{Az} > 0$ for $\mathbf{z} \neq \mathbf{0}$. So, all we have to show is that $\mathbf{x} \neq \mathbf{0}$ implies that $\mathbf{z} \neq \mathbf{0}$. We do this by showing the contrapositive, that is, if $\mathbf{z} = \mathbf{0}$ then $\mathbf{x} = \mathbf{0}$. If $\mathbf{Px} = \mathbf{0}$ then, because $\mathbf{P}^{-1}$ exists, we have $\mathbf{P}^{-1}\mathbf{Px} = \mathbf{0}$ or $\mathbf{x} = \mathbf{0}$, which completes the proof.

16 Note first that Var**(y+b)=E[(y+b-Ey-b)(y+b-Ey-b)']**=var**(y).** Therefore, Var**(Ay+b)=var(Ay).** Var**(Ay)=E[(Ay-AEy)(Ay-AEy)'] =E[A(y-Ey)(A(y-Ey))']= E[A(y-Ey)(y-Ey)'A']=A[**var**(y)]A'.**

OLS with Matrix Notation

17 OLS minimize sum of the squared residuals:

$$SSR = \hat{u}'\hat{u} \leftrightarrow \left(y - X\hat{\beta}\right)'\left(y - X\hat{\beta}\right) = y'y - 2\hat{\beta}'X'y + \hat{\beta}'X'X\hat{\beta}$$

Partial derivate with respect to beta

$$\frac{\partial SSR}{\partial \hat{\beta}} = 0 \leftrightarrow -2X'y + 2\hat{\beta}X'X = 0 \leftrightarrow$$

$$2X'y = 2\hat{\beta}X'X$$

$$X'y = \hat{\beta}X'X$$

$$\hat{\beta} = (X'X)^{-1}X'y$$

18

$\boldsymbol{\hat{\beta}=(X'X)^{-1}X'(X\beta+u)}$

$\boldsymbol{\hat{\beta}=(X'X)^{-1}X'X\beta+(X'X)^{-1}X'u}$

$\boldsymbol{\hat{\beta}=\beta+(X'X)^{-1}X'u}$

Conditioning on X

$\boldsymbol{E(\hat{\beta}|X)=\beta+(X'X)^{-1}X'E(u|X)}$

$\boldsymbol{E(\hat{\beta}|X)=\beta}$