Chapter 1 Solutions

Arthur Chen

March 29, 2024

Exercise 1.2

Let W be an $r \times s$ random matrix, and let A and C be $n \times r$ and $n \times s$ matrices of constants. Show that $\boldsymbol{E}(AW + C) = A\boldsymbol{E}(W) + C$. If B is $s \times t$ constant matrix, show that $\boldsymbol{E}(AWB) = A\boldsymbol{E}(W)B$. If $s = 1$, show that $\boldsymbol{Cov}(AW + C) = A\boldsymbol{Cov}(W)A'$.

For $\boldsymbol{E}(AW + C)$, by definition $(AW + C)_{ij} = \sum_{k=1}^{r} A_{ik}W_{kj} + C_{ij}$. Taking expectations, we have

$$\boldsymbol{E}((AW + C)_{ij}) = \sum_{k=1}^{r} A_{ik}\boldsymbol{E}(W_{kj}) + C_{ij}$$

and we see that this equals $A\boldsymbol{E}(W) + C$.

For $\boldsymbol{E}(AWB)$, $\boldsymbol{E}(AWB) = A\boldsymbol{E}(WB)$ by the above.

Exercise 1.3

Show that $\boldsymbol{Cov}(Y)$ is nonnegative definite for any random vector Y.

Let z be an arbitrary fixed vector with the same length as Y. Then

$$\begin{aligned} z'\boldsymbol{Cov}(Y)z &= z'\boldsymbol{E}[(Y-\mu)(Y-\mu)']z \\ &= \boldsymbol{E}[z'(Y-\mu)(Y-\mu)'z] \\ &= \boldsymbol{E}[||z'(Y-\mu)'||^2] \end{aligned}$$

and since $||z'(Y-\mu)'||^2 \geq 0$, its expectation is also ≥ 0.

Exercise 1.4

Let M be the ppo onto $C(X)$. Show that $(I - M)$ is the ppo onto $C(X)^{\perp}$.

$I - M$ is symmetric and $(I-M)(I-M) = I - 2M + M = I - M$ idempotent and thus is a ppo. It projects onto $C(X)^{\perp}$ because for $x \in C(X)$, $(I - M)x = x - x = 0$ and for $x \in C(X)^{\perp}$, $(I - M)x = x$. For the trace, $\boldsymbol{tr}(I - M + M) = n$ implies that $\boldsymbol{tr}(I - M) + \boldsymbol{tr}(M) = n$. Since $\boldsymbol{tr}(M) = r(X)$, we have that $\boldsymbol{tr}(I - M) = n - r(X)$.

Exercise 1.6

For a linear model $Y = X\beta + e$, $\boldsymbol{E}(e) = 0$, $\boldsymbol{Cov}(E) = \sigma^2 I$, the residuals are

$$\hat{e} = Y - X\hat{\beta} = (I - M)Y$$

where M is the perpendicular projection operator onto $C(X)$. Find the following.

Part a

$$\boldsymbol{E}(\hat{e}) = (I - M)\boldsymbol{E}(Y) = (I - M)X\beta = 0$$

because $X\beta \in C(X)$.

Part b

$$\boldsymbol{Cov}(\hat{e}) = (I - M)\boldsymbol{Cov}(Y)(I - M)' = \sigma^2(I - M)$$

because $I - M$ being a ppo implies $(I - M)(I - M)' = (I - M)$.

Part c

$$\boldsymbol{Cov}(\hat{e}, MY) = \boldsymbol{Cov}(\hat{e}, Y)M' = \boldsymbol{Cov}((I - M)Y, Y)M = \sigma^2(I - M)M = 0$$

Part d

$$\boldsymbol{E}(\hat{e}'\hat{e}) = \boldsymbol{E}[Y'(I - M)Y] = \boldsymbol{E}(e'(I - M)e)$$

where the second equality uses $(I - M)Y = (I - M)X\beta + e = (I - M)e$ and the symmetry of $I - M$. By the distribution of quadratic forms,

$$\boldsymbol{E}(e'(I - M)e) = \boldsymbol{tr}((I - M)(\sigma^2 I)) = \sigma^2(n - r)$$

Part e

To show that $\hat{e}'\hat{e} = Y'Y - Y'MY$, this immediately follows from $\hat{e}'\hat{e} = Y'(I - M)Y$ and distributing.

Part f

Using that $MY = X\hat{\beta}$, rewrite c. and e.

Making the substitution, we get that $\boldsymbol{Cov}(\hat{\beta}, \hat{e}) = 0$ and $\hat{e}'\hat{e} = Y'Y - \hat{\beta}'Y$.

Exercise 1.7

Given that $Y \sim N(\mu, V)$ and V is nonsingular, show that the density of Y is

$$f(y) = (2\pi)^{-n/2}[\det(V)]^{-1/2} \exp(-(y-\mu)'V^{-1}(y-\mu)/2)$$

We can write V as $V = AA'$, where A is nonsingular. By the definition of normals, we can rewrite Y as $Y = G(Z) = \mu + AZ$, where Z, where Z is a standard normal. Thus $G^{-1}(Y) = A^{-1}(Y-\mu)$, $dG^{-1}(Y) = A^{-1}$, and

$$\det(dG^{-1}(Y)) = \det(A^{-1}) = \det(A)^{-1} = \det(V)^{-1/2}$$

since $AA' = V$ implies $\det(A) = \det(V)^{1/2}$. For the density of Z,

$$\begin{aligned} f_Z(z) &= f_Z(G^{-1}(y)) \\ &= (2\pi)^{-n/2} \exp(-(Y-\mu)'A'^{-1}A^{-1}(Y-\mu)/2) \\ &= (2\pi)^{-n/2} \exp(-(Y-\mu)'V^{-1}(Y-\mu)/2) \end{aligned}$$

since $AA' = V$ implies $A'^{-1}A^{-1} = V^{-1}$. Thus the final density of Y is

$$\begin{aligned} f_Y(y) &= f_Z(G^{-1}(y))|\det(dG^{-1}(Y))| \\ &= (2\pi)^{-n/2} \det(V)^{-1/2} \exp(-(Y-\mu)'V^{-1}(Y-\mu)/2) \end{aligned}$$

as desired.

Exercise 1.8

Show that if $Y \sim N(\mu, V)$ and B is a fixed $n \times r$ matrix, then $BY \sim N(B\mu, BVB')$.

Let Z be an r-dimensional standard normal. Define $Y = AZ + \mu$ where $AA' = V$. Then $BY = BAZ + B\mu$ is a normally distributed random variable with distribution $N(B\mu, BAA'B') = N(B\mu, BVB')$.

Exercise 1.11

Prove that if $Y \sim N(\mu, V)$ and $VAVBV = 0$, $VAVB\mu = 0$, $VBVA\mu = 0$, and the conditions from Theorem 1.3.6 hold for $Y'AY$ and $Y'BY$, then $Y'AY$ and $Y'BY$ are independent.

Let $V = QQ'$ and rewrite $Y = \mu + QZ$, where $Z \sim N(0, I)$. We now show that the following variables are independent:

$$\begin{bmatrix} Q'AQZ \\ \mu'AQZ \end{bmatrix} \perp\!\!\!\perp \begin{bmatrix} Q'BQZ \\ \mu'BQZ \end{bmatrix} \tag{1}$$

Since these variables are all normal, showing uncorrelatedness shows independence. We first show this for one of the terms.

$$\begin{aligned}
\boldsymbol{Cov}(Q'AQZ, Q'BQZ) &= \boldsymbol{E}(Q'AQZZ'Q'B'Q) - \boldsymbol{E}(Q'AQZ)E(Z'Q'B'Q) \quad (2)\\
&= Q'AQQ'B'Q = Q'AVBQ \quad (3)
\end{aligned}$$

because $Z \sim N(0, I)$ and $QQ' = V$. We have from the same argument as the proof of Theorem 1.3.6 in the book that $Q = Q_1Q_2$, where Q_1 has orthonormal columns and Q_2 is nonsingular. Thus by the same argument,

$$Q_2^{-1}Q_1'V = Q'$$

Applying this result to Equation 2, we get that

$$\boldsymbol{Cov}(Q'AQZ, Q'BQZ) = Q_2^{-1}Q_1'VAVBV'Q_1Q_2'^{-1} = 0$$

since by assumption $VAVABV = 0$. Similar results hold for the other cross terms.

We then have by the definition of $Y = \mu + QZ$ that

$$\begin{aligned}
Y'AY &= (\mu' + Z'Q')A(QZ + \mu)\\
&= Z'Q'AQZ + Z'Q'A\mu + \mu'AQZ + \mu'A\mu
\end{aligned}$$

The second, third, and fourth terms are trivially functions of variables on the left side of Equation 1. The first term is, too, because

$$\begin{aligned}
Z'Q'AQZ &= Z'Q_2^{-1}Q_1'VAVQ_1Q_2'^{-1}Z\\
&= Z'Q_2^{-1}Q_1'VAVAVQ_1Q_2'^{-1}Z\\
&= Z'Q_2^{-1}Q_1'VAQQ'AVQ_1Q_2'^{-1}Z\\
&= Z'Q'VAQQ'AQZ\\
&= (Q'AQZ)'(Q'AQZ)\\
&= ||Q'AQZ||^2
\end{aligned}$$

is a function of $Q'AQZ$, where the second line follows from the assumption that $VAVAV = VAV$ from Theorem 1.3.6 in the book. A similar argument holds for $Y'BY$. Thus $Y'AY$ is a function of variables on the left side of Equation 1, and $Y'BY$ a function of the right side. Since the two sides of Equation 1 are independent, $Y'AY$ and $Y'BY$ are independent.