Solutions to the 1st Installment of Midterm

Woocheol Kim

Jan. 31, 1998

Question 1

A. Step 1

Under the assumption that $E\left(\widetilde{X}'\widetilde{y}\right)$and $E\left(\widetilde{X}'\widetilde{X}\right)$are finite, $\beta^* = \left\{E\left(\widetilde{X}'\widetilde{X}\right)^{-1}\right\}E\left(\widetilde{X}'\widetilde{y}\right)$is well-defined. Now, let $\varepsilon \equiv \widetilde{y} - \widetilde{X}\beta^*$. Then,

$$\begin{aligned} E\left(\widetilde{X}'\varepsilon\right) &= E\left\{\widetilde{X}'\left(\widetilde{y} - \widetilde{X}\beta^*\right)\right\} = E\left(\widetilde{X}'\widetilde{y}\right) - E\left\{\widetilde{X}'\widetilde{X}\right\}\beta^* \\ &= E\left(\widetilde{X}'\widetilde{y}\right) - E\left\{\widetilde{X}'\widetilde{X}\right\}\left[E\left(\widetilde{X}'\widetilde{X}\right)^{-1}\right]E\left(\widetilde{X}'\widetilde{y}\right) = 0, \end{aligned}$$

as is required.

B. Step 2

After plugging in $\widetilde{y} = \widetilde{X}\beta\ + \varepsilon$ and using the definition of covariance, we get

$$\begin{aligned} & cov\left(\widetilde{X}',\ \widetilde{y}\right) \\ =\ & cov\left(\widetilde{X}',\ \widetilde{X}\beta^* + \varepsilon\right) = E\left[\widetilde{X}'\left(\widetilde{X}\beta^* + \varepsilon\right)\right] - E\left(\widetilde{X}'\right)E\left(\widetilde{X}\beta^* + \varepsilon\right) \\ =\ & E\left[\widetilde{X}'\widetilde{X}\right]\beta^* + E\left(\widetilde{X}'\varepsilon\right) - E\left(\widetilde{X}'\right)E\left(\widetilde{X}\right)\beta^* - E\left(\widetilde{X}'\right)E\left(\varepsilon\right) \\ =\ & \left\{E\left[\widetilde{X}'\widetilde{X}\right] - E\left(\widetilde{X}'\right)E\left(\widetilde{X}\right)\right\}\beta^* - E\left(\widetilde{X}'\right)E\left(\varepsilon\right) \\ =\ & cov\left(\widetilde{X}',\ \widetilde{X}'\right)\beta^* - E\left(\widetilde{X}'\right)E\left(\varepsilon\right), \end{aligned}$$

where the 3rd equality comes from the result of Step 1. Thus,

$$\begin{aligned} & \left[cov\left(\widetilde{X}',\ \widetilde{X}'\right)\right]^{-1}\left[cov\left(\widetilde{X}',\ \widetilde{y}\right) + E\left(\varepsilon\right)E\left(\widetilde{X}'\right)\right] \\ =\ & \left[cov\left(\widetilde{X}',\ \widetilde{X}'\right)\right]^{-1} cov\left(\widetilde{X}',\ \widetilde{X}'\right)\beta^* = \beta^*. \end{aligned}$$

C. Step 3

When $E(\varepsilon)=0$, or $E\left(\widetilde{X}'\right)=0$, the above result implies

$$\begin{aligned}\beta^* &= \left\{E\left(\widetilde{X}'\widetilde{X}\right)^{-1}\right\}E\left(\widetilde{X}'\widetilde{y}\right)\\ &= \left[cov\left(\widetilde{X}',\ \widetilde{X}'\right)\right]^{-1}\left[cov\left(\widetilde{X}',\ \widetilde{y}\right)+E(\varepsilon)E\left(\widetilde{X}'\right)\right]\\ &= \left[cov\left(\widetilde{X}',\ \widetilde{X}'\right)\right]^{-1}cov\left(\widetilde{X}',\ \widetilde{y}\right).\end{aligned}$$

D. Step 4

The general form of the joint normal density is

$$f(x_1,x_2)=(2\pi)^{-n/2}|\Sigma|^{-1}\exp\left[-\frac{(X-\mu)'\Sigma^{-1}(X-\mu)}{2}\right].$$

Let $\mu=\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix}$ and $\Sigma=\begin{bmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{bmatrix}$be a partition according to (x_1,x_2).

Note that

$$|\Sigma|=|\Sigma_{22}|\left|\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right| \ \text{: Greene 2-72.} \tag{1}$$

Also, we can easily verify that

$$\begin{bmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{bmatrix}^{-1}=\begin{bmatrix}\Sigma_{11.2}^{-1} & -\Sigma_{11.2}^{-1}B\\ -B'\Sigma_{11.2}^{-1} & \Sigma_{22}^{-1}+B'\Sigma_{11.2}^{-1}B\end{bmatrix} \text{ with } B=\ \Sigma_{12}\Sigma_{22}^{-1}, \tag{2}$$

from postmultiplying Σ by the inverse matrix.(The verfication can be found at the end of this solution.)

Insert 1 and 2into the joint normal density, and it will give

$$\begin{aligned}&f(x_1,x_2)\\ =\ &(2\pi)^{-n/2}|\Sigma_{22}|^{-1}\left|\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right|^{-1}\\ &\times\exp\left\{-\frac{1}{2}\begin{bmatrix}(x_1-\mu_1)' & (x_2-\mu_2)'\end{bmatrix}\begin{bmatrix}\Sigma_{11.2}^{-1} & -\Sigma_{11.2}^{-1}B\\ -B'\Sigma_{11.2}^{-1} & \Sigma_{22}^{-1}+B'\Sigma_{11.2}^{-1}B\end{bmatrix}\begin{bmatrix}x_1-\mu_1\\ x_2-\mu_2\end{bmatrix}\right\}.\end{aligned}$$

First, by definition, the second determinant $\left|\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right|=|\Sigma_{11.2}|$. Next, calculate the product of three matrices inside the exp. function. It is equal to

$$(x_1-\mu_1)'\Sigma_{11.2}^{-1}(x_1-\mu_1)-(x_1-\mu_1)'\Sigma_{11.2}^{-1}B(x_2-\mu_2)$$
$$-(x_2-\mu_2)'B'\Sigma_{11.2}^{-1}(x_1-\mu_1)+(x_2-\mu_2)'\left[\Sigma_{22}^{-1}+B'\Sigma_{11.2}^{-1}B\right](x_2-\mu_2).$$

If we extract the forth term, $A=(x_2-\mu_2)'\Sigma_{22}^{-1}(x_2-\mu_2)$ that consists of the second subgroup of random variables, x_2, the remaining term can be rewritten as

$$\begin{aligned} B &= (x_1-\mu_1)'\Sigma_{11.2}^{-1}(x_1-\mu_1)-(x_1-\mu_1)'\Sigma_{11.2}^{-1}B(x_2-\mu_2) \\ &\quad -(x_2-\mu_2)'B'\Sigma_{11.2}^{-1}(x_1-\mu_1)+(x_2-\mu_2)'B'\Sigma_{11.2}^{-1}B(x_2-\mu_2) \\ &= [x_1-(\mu_1+B(x_2-\mu_2))]'\Sigma_{11.2}^{-1}[x_1-(\mu_1+B(x_2-\mu_2))]. \end{aligned}$$

Therefore, the joint density can be decomposed into $f(x_1,x_2)\equiv g(x_2)h(x_1;x_2)$, where

$$\begin{aligned} g(x_2) &= (2\pi)^{-n_2/2}|\Sigma_{22}|^{-1}\exp\left[-\frac{1}{2}(x_2-\mu_2)'\Sigma_{22}^{-1}(x_2-\mu_2)\right] \\ h(x_1;x_2) &= (2\pi)^{-n_1/2}|\Sigma_{11.2}|^{-1} \\ &\quad \times\exp\left\{-\frac{1}{2}[x_1-(\mu_1+B(x_2-\mu_2))]'\Sigma_{11.2}^{-1}[x_1-(\mu_1+B(x_2-\mu_2))]\right\}.. \end{aligned}$$

It is obvious that $g(x_2)$ is the normal density of $x_2$with mean μ_2 and variance Σ_{22}, and $h(x_1;x_2)$ is the normal density of $x_1$with mean $\mu_1+B(x_2-\mu_2)$ and variance $\Sigma_{11.2}$, when x_2 is a constant. Let $f_2(x_2)$ be the marginal density of x_2. Then,

$$\begin{aligned} f_2(x_2) &\equiv \int f(x_1,x_2)dx_1=\int g(x_2)h(x_1;x_2)dx_1 \\ &= g(x_2)\int h(x_1;x_2)dx_1=g(x_2), \end{aligned}$$

since $h(x_1;x_2)$ is a density. Now, it follows that

$$\mathbf{x}_2\sim\ N(\mu_2,\ \Sigma_{22})\ :\ \text{the marginal distribution.}$$

Also, the conditional density $f_{1.2}(x_1|x_2)\equiv\frac{f(x_1,x_2)}{f_2(x_2)}=\frac{g(x_2)h(x_1;x_2)}{g(x_2)}=h(x_1;x_2)$, that is,

$$\mathbf{x}_1|\mathbf{x}_2\ \sim N(x_1-(\mu_1+B(x_2-\mu_2)),\ \Sigma_{11.2})\quad \text{:the conditional distribution of } x_1\text{,for given } x_2.$$

If we apply the above result on conditional distribution to the case where $(\widetilde{y}, \widetilde{X}')$ has a joint multivariate normal distribution, then, the conditional distribution of $\widetilde{y}$ given $\widetilde{X} = X$ is

$$\widetilde{y}|\widetilde{X} = X \sim N\left(\mu_X,\ \Sigma_X\right)..$$

$$\begin{aligned}
\mu_X &= E\left(\widetilde{y}\right) + cov\left(\widetilde{X},\ \widetilde{y}\right) cov(\widetilde{X}', \widetilde{X}')^{-1} \left[X' - E\left(\widetilde{X}'\right)\right], \\
\Sigma_X &= var(\widetilde{y}) - cov\left(\widetilde{X},\ \widetilde{y}\right) cov(\widetilde{X}', \widetilde{X}')^{-1} cov\left(\widetilde{X}', \widetilde{y}\right).
\end{aligned}$$

Hence, when $E\left(\varepsilon\right) = 0$, by Step 3,

$$\begin{aligned}
& E\left(\widetilde{y}|\widetilde{X} = X\right) \\
= & \ E\left(\widetilde{y}\right) + cov\left(\widetilde{X},\ \widetilde{y}\right) cov(\widetilde{X}', \widetilde{X}')^{-1} \left[X' - E\left(\widetilde{X}'\right)\right] \\
= & \ E\left(\widetilde{y}\right) + \left[X - E\left(\widetilde{X}\right)\right] cov(\widetilde{X}', \widetilde{X}')^{-1} cov\left(\widetilde{X}', \widetilde{y}\right) = E\left(\widetilde{y}\right) + \left[X - E\left(\widetilde{X}\right)\right] \beta^* \\
= & \ X\beta^* + E\left(\widetilde{y}\right) - E\left(\widetilde{X}\right) \beta^* = X\beta^* - E\left(\varepsilon\right) = X\beta^*.
\end{aligned}$$

$$\begin{aligned}
& Var\left(\widetilde{y}|\widetilde{X} = X\right) \\
= & \ var(\widetilde{y}, \widetilde{y}) - cov\left(\widetilde{X},\ \widetilde{y}\right) cov(\widetilde{X}', \widetilde{X}')^{-1} cov\left(\widetilde{X}', \widetilde{y}\right) \\
= & \ var(\widetilde{y}, \widetilde{y}) - cov\left(\widetilde{X},\ \widetilde{y}\right) cov(\widetilde{X}', \widetilde{X}')^{-1} cov(\widetilde{X}', \widetilde{X}') cov(\widetilde{X}', \widetilde{X}')^{-1} cov\left(\widetilde{X}', \widetilde{y}\right) \\
= & \ var(\widetilde{y}, \widetilde{y}) - \left[cov(\widetilde{X}', \widetilde{X}')^{-1} cov\left(\widetilde{X}', \widetilde{y}\right)\right]' cov(\widetilde{X}', \widetilde{X}') \left[cov(\widetilde{X}', \widetilde{X}')^{-1} cov\left(\widetilde{X}', \widetilde{y}\right)\right] \\
= & \ var(\widetilde{y}, \widetilde{y}) - \beta^{*\prime} cov(\widetilde{X}', \widetilde{X}) \beta^*.
\end{aligned}$$

Appendix.

The inverse of a partitioned matrix:

Consider the multiplication,

$$A^* = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} \Sigma_{11.2}^{-1} & -\Sigma_{11.2}^{-1} B \\ -B' \Sigma_{11.2}^{-1} & \Sigma_{22}^{-1} + B' \Sigma_{11.2}^{-1} B \end{bmatrix},$$

with $B = \Sigma_{12} \Sigma_{22}^{-1}$. It suffices to confirm that $A^* = I$, i.e., $A_{11} = I_1,\ A_{22} = I_2,\ A_{12} = O_{12},\ A_{21} = O_{21}$.

$$\begin{aligned}
A_{11} &= \Sigma_{11} \Sigma_{11.2}^{-1} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12} \Sigma_{11.2}^{-1} = \left[\Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}\right] \Sigma_{11.2}^{-1} \\
&= \Sigma_{11.2} \Sigma_{11.2}^{-1} = I_{11}.
\end{aligned}$$

$$
\begin{aligned}
A_{12} &= -\Sigma_{11}\Sigma_{11.2}^{-1}B + \Sigma_{12}\Sigma_{22}^{-1} + \Sigma_{12}B'\Sigma_{11.2}^{-1}B \\
&= -\Sigma_{11}\Sigma_{11.2}^{-1}B + \Sigma_{11.2}\Sigma_{11.2}^{-1}B + \Sigma_{12}B'\Sigma_{11.2}^{-1}B \\
&= -\left[\Sigma_{11} - \Sigma_{11.2} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right]\Sigma_{11.2}^{-1}B \\
&= -\left[\Sigma_{11} - \Sigma_{11.2} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right]\Sigma_{11.2}^{-1}B \\
&= -\left[\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right]\Sigma_{11.2}^{-1}\Sigma_{12}\Sigma_{22}^{-1} \\
&= O_{12}.
\end{aligned}
$$

$$
A_{21} = \Sigma_{21}\Sigma_{11.2}^{-1} - \Sigma_{22}\Sigma_{22}^{-1}\Sigma_{12}\Sigma_{11.2}^{-1} = \Sigma_{21}\Sigma_{11.2}^{-1} - \Sigma_{21}\Sigma_{11.2}^{-1} = O_{21}.
$$

$$
\begin{aligned}
A_{22} &= -\Sigma_{21}\Sigma_{11.2}^{-1}B + \Sigma_{22}\left[\Sigma_{22}^{-1} + \Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11.2}^{-1}B\right] \\
&= -\Sigma_{21}\Sigma_{11.2}^{-1}B + I + \Sigma_{21}\Sigma_{11.2}^{-1}B = I_{22}.
\end{aligned}
$$