ECOM40006/90013 Econometrics 3

Week 3 Extras: Solutions

Question 1: Sums and Conditional Expectations

(a) Let S_n be as defined in the question. Then

$$\begin{aligned}
\mathbb{E}(S_n|X_1) &= \mathbb{E}(X_1 + \cdots + X_n|X_1) \\
&= \mathbb{E}(X_1|X_1) + \mathbb{E}(X_2 + \cdots + X_n|X_1) \\
&= X_1 + \sum_{i=2}^{n} \mathbb{E}(X_i|X_1) \\
&= X_1 + \sum_{i=2}^{n} \mathbb{E}(X_i) && \text{by independence} \\
&= X_1 + \sum_{i=2}^{n} \mu \\
&= X_1 + (n-1)\mu
\end{aligned}$$

(b) First, suppose we have $n \geq 2$ (it's a bit hard to talk about X_{n-1} when $n = 1$, for instance). We can write

$$\begin{aligned}
\mathbb{E}(S_n|X_n, X_{n-1}) &= \mathbb{E}(X_1 + \cdots + X_n|X_n, X_{n-1}) \\
&= \mathbb{E}(X_1 + \cdots + X_{n-2}|X_n, X_{n-1}) + \underbrace{\mathbb{E}(X_{n-1} + X_n|X_n, X_{n-1})}_{=X_{n+1}+X_n} \\
&= \sum_{i=1}^{n-2} \mathbb{E}(X_i|X_n, X_{n-1}) + X_{n-1} + X_n \\
&= \sum_{i=1}^{n-2} \mathbb{E}(X_i) + X_{n-1} + X_n && \text{by independence} \\
&= (n-2)\mu + X_{n-1} + X_n
\end{aligned}$$

(c) We follow the hint in the questions (noting that this arises by symmetry) and observe that $\mathbb{E}(X_1|S_n) = \mathbb{E}(X_j|S_n)$ for $j = 1, \ldots, n$. Then,

$$\begin{aligned}
S_n &= \mathbb{E}(S_n|S_n) \\
&= \mathbb{E}(X_1 + \cdots + X_n|S_n) \\
&= \mathbb{E}(X_1|S_n) + \mathbb{E}(X_2|S_n) + \cdots + \mathbb{E}(X_n|S_n) \\
&= n\mathbb{E}(X_1|S_n) && \text{by symmetry} \\
\implies \mathbb{E}(X_1|S_n) &= \frac{1}{n} S_n
\end{aligned}$$

(d) For $m \geq 0$,

$$\begin{aligned}
\mathbb{E}(S_{n+m}|S_n) &= \mathbb{E}(S_n + X_{n+1} + \cdots + X_{n+m}|S_n) \\
&= \mathbb{E}(S_n|S_n) + \mathbb{E}(X_{n+1} + \cdots + X_{n+m}|S_n) \\
&= S_n + \sum_{j=1}^{m} \mathbb{E}(X_{n+j}) && \text{by independence} \\
&= S_n + m\mu
\end{aligned}$$

noting that since our RVs are independent, conditioning an RV X_{n+j}, $j = 1, \ldots, m$ on S_n reduces us down to the unconditional expectation of X_{n+j}, which is simply μ.

(e) Using our results from part (c), we immediately note that for $j = 1, \ldots, n$,

$$\mathbb{E}(X_j|S_{n+m}) = \frac{1}{n+m} S_{n+m}$$

So, we can write

$$\begin{aligned}
\mathbb{E}(S_n|S_{n+m}) &= \sum_{j=1}^{n} \mathbb{E}(X_j|S_{n+m}) \\
&= \sum_{j=1}^{n} \frac{1}{n+m} S_{n+m} \\
&= \frac{n}{n+m} S_{n+m}
\end{aligned}$$

In the case where $m = 0$, we reduce down to the trivial case $\mathbb{E}(S_n|S_n) = S_n$.

Question 2: The Jacobian of Transformation

(a) If we have $f(g(y))$ then the expression we're evaluating is $f(x) = f(g(y))$. First, suppose that $f(x)$ has a maximum at x^*. Then it must be the case that the first order condition is satisfied, i.e.

$$f'(x^*) = 0.$$

Now, consider $f(g(y))$. If we wanted to find the value of y that maximises this we can take the first-order condition. Doing this requires the Chain Rule:

$$\frac{df(g(y))}{dy} = \frac{df(g(y))}{dg(y))} \frac{dg(y)}{dy} = 0$$

or just $f'(g(y))g'(y)$. Let the value of y maximising f be y^* and assume $g'(y^*) \neq 0$. Then

$$f'(g(y^*)) = 0.$$

Now, we know that $f'(x^*) = 0$ so the locations of the maximum can be represented by

$$x^* = g(y^*).$$

(As an aside: note that you can't do this for a probability density.[1])

(b) Let X be a random variable with PDF $f_X(x)$. Now, consider a change of variables $x = g(y)$, where g is a strictly monotone increasing/decreasing function.

Then one can write $y = g^{-1}(x)$. In this sense this allows us to deal with the CDFs as follows:

$$\underbrace{\Pr(Y \leq y)}_{F_Y(y)} = \underbrace{\Pr(g^{-1}(X) \leq y)}_{\text{replace } Y \text{ with } g^{-1}(X)} = \begin{cases} \Pr(X \leq g(y)) & \text{if } g \text{ is monotone increasing} \\ \Pr(X \geq g(y)) & \text{if } g \text{ is monotone decreasing} \end{cases}$$

So applying these inverse functions, we can write $F_Y(y)$ as a function of the original CDF $F_X(x)$. The derivations differ mildly depending on whether g is monotone increasing or decreasing. In the event it is increasing, we can write

$$F_Y(y) = F_X(g(y)).$$

Differentiating both sides using the Chain Rule, we get

$$\begin{aligned} f_Y(y) &= \frac{dF_X(g(y))}{dy} \\ &= \frac{dF_X(g(y))}{dg(y)} \frac{dg(y)}{dy} \\ &= f_X(g(y))|g'(y)|. \end{aligned}$$

Observe that $g'(y) > 0$ as g is strictly monotone increasing. In the case where g is monotone decreasing, we write instead

$$F_Y(y) = 1 - \Pr(X \leq g(y)) = 1 - F_X(g(y)).$$

Differentiating both sides in a similar manner to before, we get

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}(1 - F_X(g(y))) \\ &= -f_X(g(y)) \frac{dg(y)}{dy} \\ &= f_X(g(y))|g'(y)|. \end{aligned}$$

[1] The rough sketch is like this: first observe that $f_Y(y) = f_X(x)|g'(y)|$. Get rid of the absolute value by defining $s \in \{-1, 1\}$ so that you have

$$f_X(g(y))sg'(y).$$

Then differentiating this with respect to y you get the even more ugly expression

$$\begin{aligned} f_Y'(y) &= sf_X(g(y))g''(y) + sg'(y)\frac{df_X(g(y))}{dy} \\ &= sf_X(g(y))g''(y) + sg'(y)\frac{df_X(g(y))}{dg(y)}\frac{dg(y)}{dy} \\ &= sf_X(g(y))g''(y) + sf_X'(g(y))[g'(y)]^2 \end{aligned}$$

and setting that equal to zero gives the first order condition. But the point here is that trying to set $x^* = g(y^*)$ does not work this time for probability densities.

Note that by necessity PDFs must be non-negative. The absolute value here helps avoid any confusion with the signs on everything.

(c) Observe that the function $y = -\log(x)$ is monotonically decreasing. In this case it suffices to just check the 'endpoints.'

For $x = 0$ we have $y = \log(x) \to \infty$. For $x = 1$ we have $y = -\log(1) = 0$ So while we have $x \in [0, 1]$, we have $y \in [0, \infty)$.

(d) Rearranging our expression for y we have

$$-y = \log x \implies x = e^{-y}.$$

Notice that we now have $x = g(y)$. The derivative is $g'(y) = -e^{-y}$ with an absolute value $|g'(y)| = e^{-y}$. Using our formulae from earlier, we can simply make direct substitutions:

$$\begin{aligned} f_Y(y) &= f_X(x)|g'(y)| \\ &= e^{-y}, \qquad y \in [0, \infty), \end{aligned}$$

since $f_X(x) = 1$ for all values of $x \in [0, 1]$. The resulting PDF can be recognised (if you're familiar with it) as the density of an exponential density with parameter $\lambda = 1$. In shorthand we usually refer to it as $y \sim \text{Exp}(1)$.

(e) We have values of $x = 0.1$ and $x = 0.2$. Plugging them into our expression for y gives

$$y = -\log(0.1) = 2.302 \qquad \text{and} \qquad y = -\log(0.2) = 1.609.$$

Putting these together, the corresponding values of y are contained in the interval

$$[1.609, 2.302] \qquad \text{or} \qquad [y_0, y_0 + dy],$$

where $y_0 = 1.609$ and $dy = 0.693 \neq dx$. That is: dy and dx are of different magnitude. In terms of calculating the probability that $y \in [1.609, 2.302]$, we can do this directly:

$$\begin{aligned} \Pr(y \in [1.609, 2.302]) &= \int_{1.609}^{2.302} f_Y(y)\, dy \\ &= \int_{1.609}^{2.302} e^{-y}\, dy \\ &= [-e^{-y}]\Big|_{1.609}^{2.302} \\ &= -e^{2.302} + e^{1.609} \\ &= 0.1. \end{aligned}$$

So with a change in variables, we still obtain the same probability; you can think of the variable transformation as a change in perspective. However, the way that x and y are measured is different, which is evidenced by the fact that dx and dy are of different magnitude.

Overall, this exercise is just to get you to try out some of these results for yourself. As with all things, we can go into much deeper detail on any given point. But for now it suffices to just see that these theorems work and that we can demonstrate them if needed!

(f) Before starting, a few rules on variance:

- Property 1: $\text{Var}(aX) = a^2\text{Var}(X)$
- Property 2: $\text{Var}(aX + bY) = a^2\text{Var}(X) + b^2\text{Var}(Y) + 2ab\text{cov}(X, Y)$
- Property 3: $\text{cov}(X, Y) = \text{corr}(X, Y) \times \text{sd}(X)\text{sd}(Y)$

Now, we can just take this directly.

$$\begin{aligned}
\text{Var}(X_2) &= \text{Var}(\sigma_2[\rho Z_1 + \sqrt{1-\rho^2}Z_2] + \mu_2) \\
&= \text{Var}(\sigma_2[\rho Z_1 + \sqrt{1-\rho^2}Z_2]) && \text{(Variance ignores constants)} \\
&= \sigma_2^2\text{Var}(\rho Z_1 + \sqrt{1-\rho^2}Z_2) && \text{(Property 1)} \\
&= \sigma_2^2[\rho^2\text{Var}(Z_1) + (1-\rho^2)\text{Var}(Z_2) + 2\rho\sqrt{1-\rho^2}\text{cov}(Z_1, Z_2)] && \text{(Property 2)} \\
&= \sigma_2^2[\rho^2\text{Var}(Z_1) + (1-\rho^2)\text{Var}(Z_2)] && (\text{corr}(Z_1, Z_2) = 0) \\
&= \sigma_2^2[\rho_2 + 1 - \rho^2] && (\text{Var}(Z_1) = \text{Var}(Z_2) = 1) \\
&= \sigma_2^2.
\end{aligned}$$

(g) Rearrangement of Z_1 is quite straightforward: from there we get

$$Z_1 = g(X_1, X_2) = \frac{X_1 - \mu_1}{\sigma_1}.$$

For Z_2 it takes a bit more work:

$$\begin{aligned}
X_2 &= \sigma_2[\rho Z_1 + \sqrt{1-\rho^2}Z_2] + \mu_2 \\
\implies \frac{X_2 - \mu_2}{\sigma_2} &= \rho Z_1 + \sqrt{1-\rho^2}Z_2 \\
\implies \frac{X_2 - \mu_2}{\sigma_2} - \rho Z_1 &= \sqrt{1-\rho^2}Z_2 \\
\implies Z_2 = h(X_1, X_2) &= \frac{1}{\sqrt{1-\rho^2}}\left[\frac{X_2 - \mu_2}{\sigma_2} - \rho\frac{X_1 - \mu_1}{\sigma_1}\right],
\end{aligned}$$

where in the last line we substituted in Z_1 from above. Now, to cut down on notation, define

$$a = \frac{X_1 - \mu_1}{\sigma_1}, \qquad b = \frac{X_2 - \mu_2}{\sigma_2}.$$

Then

$$
\begin{aligned}
Z_1^2 + Z_2^2 &= a^2 + \left[\frac{1}{\sqrt{1-\rho^2}}(b - \rho a) \right]^2 \\
&= a^2 + \frac{1}{1-\rho^2}(b^2 - 2\rho ab + \rho^2 a^2) \\
&= \frac{a^2(1-\rho^2)}{1-\rho^2} + \frac{1}{1-\rho^2}(b^2 - 2\rho ab + \rho^2 a^2) \\
&= \frac{1}{1-\rho^2}(b^2 - 2\rho ab + a^2 - \rho^2 a^2 + \rho^2 a^2) \\
&= \frac{1}{1-\rho^2}(a^2 + b^2 - 2\rho ab) \\
&= \frac{1}{1-\rho^2}\left(\frac{(X_1-\mu_1)^2}{\sigma_1^2} + \frac{(X_2-\mu_2)^2}{\sigma_2^2} - 2\rho \frac{(X_1-\mu_1)(X_2-\mu_2)}{\sigma_1\sigma_2} \right),
\end{aligned}
$$

where we expand out a and b in the last line. If you recognise this expression from previous questions, you might be able to see where this question is going. But if not, let's keep forging ahead!

(h) Let's first take all the partial derivatives individually. For $g(X_1, X_2)$ we have

$$
\frac{\partial g}{\partial X_1} = \frac{1}{\sigma_1}, \qquad \text{and} \qquad \frac{\partial g}{\partial X_2} = 0.
$$

For $h(X_1, X_2)$ we can calculate

$$
\frac{\partial h}{\partial X_1} = -\frac{\rho}{\sqrt{1-\rho^2}}\frac{1}{\sigma_1} \qquad \text{and} \qquad \frac{\partial h}{\partial X_2} = \frac{1}{\sqrt{1-\rho^2}}\frac{1}{\sigma^2}.
$$

The Jacobian can be written as

$$
J = \begin{bmatrix} \dfrac{\partial g}{\partial X_1} & \dfrac{\partial h}{\partial X_1} \\ \dfrac{\partial g}{\partial X_2} & \dfrac{\partial h}{\partial X_2} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\sigma_1} & -\dfrac{\rho}{\sqrt{1-\rho^2}}\dfrac{1}{\sigma_1} \\ 0 & \dfrac{1}{\sqrt{1-\rho^2}}\dfrac{1}{\sigma^2} \end{bmatrix}.
$$

The determinant of the Jacobian is

$$
\det(J) = \det \begin{bmatrix} \dfrac{1}{\sigma_1} & -\dfrac{\rho}{\sqrt{1-\rho^2}}\dfrac{1}{\sigma_1} \\ 0 & \dfrac{1}{\sqrt{1-\rho^2}}\dfrac{1}{\sigma^2} \end{bmatrix} = \frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2}}.
$$

Nothing changes if we apply an absolute value to this expression since it is already positive (using the fact that $\sigma_1, \sigma_2 > 0$ and $|\rho| < 1$.

(i) Note that we can now apply the more general version of a transformation of random variables using the Jacobian. In particular, we can write

$$
f_{X_1,X_2}(x_1, x_2) = f_{Z_1,Z_2}(z_1, z_2) \times \text{abs}(\det J).
$$

We can apply this directly. Replacing z_1 with $g(X_1, X_2)$ and z_2 with $h(X_1, X_2)$, we can actually use our derivation from part (g) above to write

$$\begin{aligned} f_{X_1,X_2}(x_1,x_2) &= \frac{1}{2\pi}\exp\left\{-\frac{1}{2}(z_1^2+z_2^2)\right\} \times \text{abs}(\det J) \\ &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{1-\rho^2}\left(\frac{(X_1-\mu_1)^2}{\sigma_1^2}+\frac{(X_2-\mu_2)^2}{\sigma_2^2}-2\rho\frac{(X_1-\mu_1)(X_2-\mu_2)}{\sigma_1\sigma_2}\right)\right\}, \end{aligned}$$

which happens to be the probability density function of a bivariate normal distribution with mean and covariance matrix (respectively)

$$\mu = \begin{bmatrix}\mu_1 \\ \mu_2\end{bmatrix} \qquad \text{and} \qquad \Sigma = \begin{bmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix}.$$

In fact it's the same PDF that we dealt with last week!

Question 3: The Partitioned Matrix Inverse

(a) If A_{11} and A_{22} are invertible, they must be square. Let the dimensions of A_{11} be $n \times n$ and A_{22} be $m \times m$. Then $A_{12} = n \times m$ and $A_{21} = m \times n$.

To verify the elements of A^{-1} are indeed the correct elements, we require that

$$A^{-1}A = \begin{bmatrix}B_{11} & B_{12} \\ B_{21} & B_{22}\end{bmatrix}\begin{bmatrix}A_{11} & A_{12} \\ A_{21} & A_{22}\end{bmatrix} = \begin{bmatrix}B_{11}A_{11}+B_{12}A_{21} & B_{11}A_{12}+B_{12}A_{22} \\ B_{21}A_{11}+B_{22}A_{21} & B_{21}A_{12}+B_{22}A_{22}\end{bmatrix} = \begin{bmatrix}a & b \\ c & d\end{bmatrix} = \begin{bmatrix}I_n & 0_{n\times m} \\ 0_{m\times n} & I_m\end{bmatrix}$$

We check each individual element of the matrix separately to ensure that the desired answer is achieved. So to begin with we check the top left element:

$$\begin{aligned} B_{11}A_{11} + B_{12}A_{21} &= (A_{11}^{-1} + A_{11}^{-1}A_{12}FA_{21}A_{11}^{-1})A_{11} - A_{11}^{-1}A_{12}FA_{21} \\ &= A_{11}^{-1}A_{11} + A_{11}^{-1}A_{12}FA_{21}A_{11}^{-1}A_{11} - A_{11}^{-1}A_{12}FA_{21} \\ &= I_n + A_{11}^{-1}A_{12}FA_{21} - A_{11}^{-1}A_{12}FA_{21} \\ &= I_n \end{aligned}$$

Now the bottom right element is

$$\begin{aligned} B_{21}A_{12} + B_{22}A_{22} &= (-FA_{21}A_{11}^{-1})A_{12} + FA_{22} \\ &= FA_{22} - FA_{21}A_{11}^{-1}A_{12} \\ &= F(A_{22} - A_{21}A_{11}^{-1}A_{12}) \\ &= (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}(A_{22} - A_{21}A_{11}^{-1}A_{12}) \\ &= I_m \end{aligned}$$

The top right and bottom left elements should be zero. So checking the top right element,

$$\begin{aligned}
B_{11}A_{12} + B_{12}A_{22} &= (A_{11}^{-1} + A_{11}^{-1}A_{12}FA_{21}A_{11}^{-1})A_{12} - A_{11}^{-1}A_{12}FA_{22} \\
&= A_{11}^{-1}A_{12} + A_{11}^{-1}A_{12}FA_{21}A_{11}^{-1}A_{12} - A_{11}^{-1}A_{12}FA_{22} \\
&= A_{11}^{-1}A_{12} + A_{11}^{-1}A_{12}(FA_{21}A_{11}^{-1}A_{12} - FA_{22}) \\
&= A_{11}^{-1}A_{12} - A_{11}^{-1}A_{12}(FA_{22} - FA_{21}A_{11}^{-1}A_{12}) \\
&= A_{11}^{-1}A_{12} - A_{11}^{-1}A_{12}(F(A_{22} - A_{21}A_{11}^{-1}A_{12})) \\
&= A_{11}^{-1}A_{12} - A_{11}^{-1}A_{12}((A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}(A_{22} - A_{21}A_{11}^{-1}A_{12})) \\
&= A_{11}^{-1}A_{12} - A_{11}^{-1}A_{12} \\
&= 0_{n\times m}
\end{aligned}$$

Finally, we check the bottom left:

$$\begin{aligned}
B_{21}A_{11} + B_{22}A_{21} &= (-FA_{21}A_{11}^{-1})A_{11} + FA_{21} \\
&= -FA_{21}A_{11}^{-1})A_{11} + FA_{21} \\
&= -FA_{21} + FA_{21} \\
&= 0_{m\times n}
\end{aligned}$$

Therefore

$$A^{-1}A = \begin{bmatrix} I_n & 0_{n\times m} \\ 0_{m\times n} & I_m \end{bmatrix}$$

which is sufficient to show that A^{-1} as claimed is the inverse of A.

(b) In order to verify this, we multiply by $(A_{11} - A_{12}A_{22}^{-1}A_{21})$ on both sides, which should return the identity matrix on both sides. On the LHS we have I_n, but on the RHS we have

$$\begin{aligned}
&(A_{11}^{-1} + A_{11}^{-1}A_{12}FA_{21}A_{11}^{-1})(A_{11} - A_{12}A_{22}^{-1}A_{21}) \\
&= A_{11}^{-1}(A_{11} - A_{12}A_{22}^{-1}A_{21}) + A_{11}^{-1}A_{12}FA_{21}A_{11}^{-1}(A_{11} - A_{12}A_{22}^{-1}A_{21}) \\
&= A_{11}^{-1}A_{11} - A_{11}^{-1}A_{12}A_{22}^{-1}A_{21} + A_{11}^{-1}A_{12}FA_{21}A_{11}^{-1}A_{11} - A_{11}^{-1}A_{12}FA_{21}A_{11}^{-1}A_{12}A_{22}^{-1}A_{21} \\
&= I_n - A_{11}^{-1}A_{12}A_{22}^{-1}A_{21} + A_{11}^{-1}A_{12}F(A_{21} - A_{21}A_{11}^{-1}A_{12}A_{22}^{-1}A_{21}) \\
&= I_n - A_{11}^{-1}A_{12}A_{22}^{-1}A_{21} + A_{11}^{-1}A_{12}F(I_m - A_{21}A_{11}^{-1}A_{12}A_{22}^{-1})A_{21} \\
&= I_n - A_{11}^{-1}A_{12}A_{22}^{-1}A_{21} + A_{11}^{-1}A_{12}F(A_{22} - A_{21}A_{11}^{-1}A_{12})A_{22}^{-1}A_{21} \\
&= I_n - A_{11}^{-1}A_{12}A_{22}^{-1}A_{21} + A_{11}^{-1}A_{12}FF^{-1}A_{22}^{-1}A_{21} \\
&= I_n - A_{11}^{-1}A_{12}A_{22}^{-1}A_{21} + A_{11}^{-1}A_{12}A_{22}^{-1}A_{21} \\
&= I_n
\end{aligned}$$

Therefore

$$(A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} = A_{11}^{-1} + A_{11}^{-1}A_{12}FA_{21}A_{11}^{-1}$$

as required.

(c) Let $A = X'X$ with $X = (\, X_1 \; X_2 \,)$. Without loss of generality, define the dimensions of X_1 as $k \times n$ and X_2 as $k \times m$.

Then $X_1'X_1$ is $n \times n$, $X_2'X_2$ is $m \times m$, $X_1'X_2$ is $n \times m$ and $X_2'X_1$ is $m \times n$.

These dimensions mean that we may use the formula for a block inverse as stated in the first part of this question. Firstly, define the following:

$$\begin{aligned}
A_{11} &= X_1'X_1 \\
A_{12} &= X_1'X_2 \\
A_{21} &= X_2'X_1 \\
A_{22} &= X_2'X_2
\end{aligned}$$

Then the elements of the matrix A^{-1} are, using the same notation as before,

$$\begin{aligned}
B_{11} &= A_{11}^{-1} + A_{11}^{-1}A_{12}FA_{21}A_{11}^{-1} \\
&= (X_1'X_1)^{-1} + (X_1'X_1)^{-1}X_1'X_2(X_2'X_2 - X_2'X_1(X_1'X_1)^{-1}X_1'X_2)^{-1}X_1'X_2(X_1'X_1)^{-1} \\
&= (X_1'X_1 - X_1'X_2(X_2'X_2)^{-1}X_2'X_1)^{-1} && \text{(from part b)}
\end{aligned}$$

$$\begin{aligned}
B_{22} &= F \\
&= (X_2'X_2 - X_2'X_1(X_1'X_1)^{-1}X_1'X_2)^{-1}
\end{aligned}$$

$$\begin{aligned}
B_{21} &= -FA_{21}A_{11}^{-1} \\
&= -(X_2'X_2 - X_2'X_1(X_1'X_1)^{-1}X_1'X_2)^{-1}X_2'X_1(X_1'X_1)^{-1}
\end{aligned}$$

$$\begin{aligned}
B_{12} &= -A_{11}^{-1}A_{12}F \\
&= -(X_1'X_1)^{-1}X_1'X_2(X_2'X_2 - X_2'X_1(X_1'X_1)^{-1}X_1'X_2)^{-1}
\end{aligned}$$