AMS256 Homework 3

1.

$$\begin{aligned}
\text{Cov}(\boldsymbol{Ax}) &= \text{E}\left[(\boldsymbol{Ax} - E(\boldsymbol{Ax}))(\boldsymbol{Ax} - E(\boldsymbol{Ax}))^T\right] \\
&= \text{E}\left[(\boldsymbol{Ax} - \boldsymbol{A}\text{E}(\boldsymbol{x}))(\boldsymbol{Ax} - \boldsymbol{A}\text{E}(\boldsymbol{x}))^T\right] \\
&= \text{E}\left[\boldsymbol{A}(\boldsymbol{x} - \text{E}(\boldsymbol{x}))(\boldsymbol{x} - \text{E}(\boldsymbol{x}))^T \boldsymbol{A}^T\right] \\
&= \boldsymbol{A}\text{E}\left[(\boldsymbol{x} - \text{E}(\boldsymbol{x}))(\boldsymbol{x} - \text{E}(\boldsymbol{x}))^T\right] \boldsymbol{A}^T \\
&= \boldsymbol{A}\text{Cov}(\boldsymbol{x})\boldsymbol{A}^T
\end{aligned}$$

2. With the similar method used above,

$$\begin{aligned}
\text{Cov}(\boldsymbol{Ax},\ \boldsymbol{By}) &= \text{E}\left[(\boldsymbol{Ax} - \text{E}(\boldsymbol{Ax}))(\boldsymbol{By} - \text{E}(\boldsymbol{By}))^T\right] \\
&= \text{E}\left[(\boldsymbol{Ax} - \boldsymbol{A}\text{E}(\boldsymbol{x}))(\boldsymbol{By} - \boldsymbol{B}\text{E}(\boldsymbol{y}))^T\right] \\
&= \text{E}\left[\boldsymbol{A}(\boldsymbol{x} - \text{E}(\boldsymbol{x}))(\boldsymbol{y} - \text{E}(\boldsymbol{y}))^T \boldsymbol{B}^T\right] \\
&= \boldsymbol{A}\text{E}\left[(\boldsymbol{x} - \text{E}(\boldsymbol{x}))(\boldsymbol{y} - \text{E}(\boldsymbol{y}))^T\right] \boldsymbol{B}^T \\
&= \boldsymbol{A}\text{Cov}(\boldsymbol{x},\ \boldsymbol{y})\boldsymbol{B}^T
\end{aligned}$$

3. With the similar method used above,

$$\begin{aligned}
\text{Cov}(\boldsymbol{x} - \boldsymbol{a},\ \boldsymbol{y} - \boldsymbol{b}) &= \text{E}\left[(\boldsymbol{x} - \boldsymbol{a} - \text{E}(\boldsymbol{x} - \boldsymbol{a}))(\boldsymbol{y} - \boldsymbol{b} - \text{E}(\boldsymbol{y} - \boldsymbol{b}))^T\right] \\
&= \text{E}\left[(\boldsymbol{x} - \text{E}(\boldsymbol{x}))(\boldsymbol{y} - \text{E}(\boldsymbol{y}))^T\right] \\
&= \text{E}\left[(\boldsymbol{x} - \text{E}(\boldsymbol{x}))(\boldsymbol{y} - \text{E}(\boldsymbol{y}))^T\right] \\
&= \text{Cov}(\boldsymbol{x},\ \boldsymbol{y})
\end{aligned}$$

4.

$$\begin{aligned}
x_{i+1} = \rho x_i + a \quad &= \quad \rho(\rho x_{i-1} + a) + a \\
&= \quad \rho^2 x_{i-1} + a\rho + a \\
&= \quad \rho^2(\rho x_{i-2} + a) + a\rho + a \\
&= \quad \rho^3 x_{i-2} + a\rho^2 + a\rho + a \\
&\cdots \\
&= \quad \rho^i x_1 + \sum_{k=1}^{i} a\rho^{k-1}.
\end{aligned}$$

$$\begin{aligned}
cov(x_i, x_j) \quad &= \quad cov(\rho^{i-1}x_1 + c_1, \rho^{j-1}x_1 + c_2) \\
&= \quad \rho^{i+j-2}cov(x_1, x_1) = \rho^{i+j-2}\sigma^2
\end{aligned}$$

where c_1 and c_2 are some constants.

$$[\text{Cov}(\boldsymbol{X})]_{ij} \quad = \quad \sigma^2 \cdot \rho^{i+j-2}.$$

5. First observe
$$\mathrm{var}(\bar{x}) = \frac{1}{n^2}\sum \mathrm{var}(x_i) = \frac{1}{n^2}\sum \sigma_i^2.$$
Now we show $\mathrm{E}\{\frac{\sum(x_i-\bar{x})^2}{n(n-1)}\} = \frac{1}{n^2}\sum\sigma_i^2$.
$$\begin{aligned}\mathrm{E}\{\frac{\sum(x_i-\bar{x})^2}{n(n-1)}\} &= \frac{1}{n(n-1)}\mathrm{E}\{\sum(x_i-\bar{x})^2\} = \frac{1}{n(n-1)}\{\sum \mathrm{E}(x_i^2) - 2\mathrm{E}(x_i\bar{x}) + n\mathrm{E}(\bar{x}^2)\}\\ &= \frac{1}{n^2}\sum\sigma_i^2,\end{aligned}$$
since $\mathrm{E}(x_i^2) = \sigma_i^2 + \mu^2$ and $\mathrm{E}(x_ix_j) = \mu^2$, $i \neq j$.

6. Observe $\boldsymbol{V} = \mathrm{diag}(1,2)$. So, $\hat{\beta}_{GLS} = (\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}\boldsymbol{V}^{-1}\boldsymbol{y} = (x_1^2 + \frac{1}{2}x_2^2)^{-1}(x_1y_1 + \frac{1}{2}x_2y_2)$.
$$\mathrm{Var}(\hat{\beta}_{GLS}) = \mathrm{Var}(x_1^2 + \frac{1}{2}x_2^2)^{-1}(x_1y_1 + \frac{1}{2}x_2y_2) = (x_1^2 + \frac{1}{2}x_2^2)^{-2}(x_1^2 + \frac{1}{4}x_2^2)\sigma^2.$$

7. The mgf of x is $M_x(t) = \mathrm{E}[e^{tx}] = \exp(\frac{t^2\sigma^2}{2})$.
$$\begin{aligned}M_x'(t) &= t\sigma^2\exp(\frac{t^2\sigma^2}{2})\\ M_x''(t) &= t^2\sigma^4\exp(\frac{t^2\sigma^2}{2}) + \sigma^2\exp(\frac{t^2\sigma^2}{2})\\ M_x^{(3)}(t) &= t^3\sigma^6\exp(\frac{t^2\sigma^2}{2}) + 2t\sigma^4\exp(\frac{t^2\sigma^2}{2}) + t\sigma^4\exp(\frac{t^2\sigma^2}{2})\\ M_x^{(4)}(t) &= t^4\sigma^8\exp(\frac{t^2\sigma^2}{2}) + 3t^2\sigma^6\exp(\frac{t^2\sigma^2}{2}) + 2t^2\sigma^6\exp(\frac{t^2\sigma^2}{2})\\ &\quad + 2\sigma^4\exp(\frac{t^2\sigma^2}{2}) + t^2\sigma^6\exp(\frac{t^2\sigma^2}{2}) + \sigma^4\exp(\frac{t^2\sigma^2}{2})\end{aligned}$$
So, $\mu_3 = M_x^3(0) = 0$ and $\mu_4 = M_x^{(4)}(0) = 3\sigma^4$

8. If $z_i \overset{iid}{\sim} \mathrm{N}(0,1)$ then $\boldsymbol{z} \sim \mathrm{N}_n(\boldsymbol{0}, \boldsymbol{I})$. Since $\boldsymbol{y}$ is a linear combination of a multivariate normal, it is also a multivariate normal.
$$\mathrm{E}(\boldsymbol{y}) = \mathrm{E}(\boldsymbol{A}\boldsymbol{z}) + \mathrm{E}(\boldsymbol{\mu}) = \boldsymbol{A}\mathrm{E}(\boldsymbol{z}) + \mathrm{E}(\boldsymbol{\mu}) = 0 + \boldsymbol{\mu} = \boldsymbol{\mu}$$
and
$$\mathrm{Cov}(\boldsymbol{y}) = \mathrm{Cov}(\boldsymbol{A}\boldsymbol{z}) = \boldsymbol{A}\mathrm{Cov}(\boldsymbol{z})\boldsymbol{A}^T + 0 = \boldsymbol{A}\boldsymbol{I}\boldsymbol{A}^T = \boldsymbol{A}\boldsymbol{A}^T = \Sigma.$$
So $\boldsymbol{y} \sim \boldsymbol{N}_n(\boldsymbol{\mu}, \Sigma)$, which means its density is
$$f(\boldsymbol{y}) = (2\pi)^{-n/2}|\Sigma|^{-1/2}\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\right\}.$$

9. Let $\boldsymbol{z}$ be a multivariate standard normal random variable, and by using the transformation $\boldsymbol{x} = \boldsymbol{m} + \boldsymbol{V}\boldsymbol{z}$, where $\boldsymbol{V}$ is a $k \times k$ invertible matrix such that $\Sigma = \boldsymbol{V}\boldsymbol{V}^T$.
$$\begin{aligned}\mathrm{E}[\boldsymbol{x}] &= \mathrm{E}[\boldsymbol{m} + \boldsymbol{V}\boldsymbol{z}] = \boldsymbol{m}] + \boldsymbol{V}\boldsymbol{0} = \boldsymbol{m},\\ \mathrm{Cov}[\boldsymbol{x}] &= \mathrm{Cov}[\boldsymbol{m} + \boldsymbol{V}\boldsymbol{z}] = \boldsymbol{V}\mathrm{Cov}(\boldsymbol{z})\boldsymbol{V}^T = \Sigma.\end{aligned}$$

10. Since $\boldsymbol{y}$ is distributed as $\mathrm{N}_n(\boldsymbol{\mu}, \Sigma)$, then its moment generating function is $M_{\boldsymbol{y}}(\boldsymbol{t}) = \exp(\boldsymbol{t}^T\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{t}^T\Sigma\boldsymbol{t})$. Let $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y}$. The mgf of $\boldsymbol{x}$ is

$$M_{\boldsymbol{x}}(\boldsymbol{t}) = \mathrm{E}[e^{\boldsymbol{t}^T\boldsymbol{z}}] = \mathrm{E}[e^{\boldsymbol{t}^T\boldsymbol{C}\boldsymbol{y}}] = M_{\boldsymbol{y}}(\boldsymbol{t}^T\boldsymbol{C}) = \exp(\boldsymbol{t}^T(\boldsymbol{C}\boldsymbol{\mu}) + \frac{1}{2}\boldsymbol{t}^T(\boldsymbol{C}\Sigma\boldsymbol{C}^T)\boldsymbol{t}).$$

$\Rightarrow \quad \boldsymbol{x} = \boldsymbol{C}\boldsymbol{y} \sim \mathrm{N}_p(\boldsymbol{C}\boldsymbol{\mu}, \boldsymbol{C}\Sigma\boldsymbol{C}^T)$.

11. ($\longrightarrow$) If $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ are independent, then $\mathrm{Cov}(\boldsymbol{y}_1, \boldsymbol{y}_2) = 0$, and

$$\mathrm{Cov}(\boldsymbol{y}) = \begin{bmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{bmatrix}$$

Therefore, $\Sigma_{12} = \Sigma_{21}^T = 0$.

($\longleftarrow$) If $\Sigma_{12} = \Sigma_{21}^T = 0$, then the mgf for $\boldsymbol{y}$ is

$$m_{\boldsymbol{y}}(\boldsymbol{t}) = \exp\{\boldsymbol{t}^T\boldsymbol{\mu} + \boldsymbol{t}^T\Sigma\boldsymbol{t}/2\} = \exp\{\boldsymbol{t}_1^T\boldsymbol{\mu}_1 + \boldsymbol{t}_2^T\boldsymbol{\mu}_2 + \boldsymbol{t}_1^T\Sigma_{11}\boldsymbol{t}_1/2 + +\boldsymbol{t}_2^T\Sigma_{22}\boldsymbol{t}_2/2\}.$$

That is, $m_{\boldsymbol{y}}(\boldsymbol{t}) = m_{\boldsymbol{y}_1}(\boldsymbol{t}_1) \times m_{\boldsymbol{y}_2}(\boldsymbol{t}_2)$ implying independence between $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$.

12. Observe $[\bar{x}, \bar{y}]^T = \frac{1}{n}\sum Z_i$, that is, a linear combination of bivariate normal random variables.

$$\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} \sim N_2(\boldsymbol{\theta}, \Sigma/n)$$

13. (a) Observe $\mathbf{1}\mathbf{1}^T = \boldsymbol{J}_{n\times n}$. So $\Sigma_{ii} = \sigma^2$ and $\Sigma_{ij} = \rho\sigma^2$, $i \neq j$.
    (b) We can express

$$\sum_{i=1}^{n}(Y_i - \bar{Y})^2/[\sigma^2(1-\rho)] = \boldsymbol{y}^T \underbrace{\{\frac{1}{\sigma^2(1-\rho)}(\boldsymbol{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T)\}}_{\boldsymbol{A}} \boldsymbol{y}$$

    We can show $\boldsymbol{A}\Sigma = (\boldsymbol{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T)$ is idempotent and $\mathrm{rank}(\boldsymbol{A}\Sigma) = n-1$. So $\sum_{i=1}^n (Y_i - \bar{Y})^2/[\sigma^2(1-\rho)] \sim \chi^2(n-1)$.
    (c) We can show

$$\bar{Y} = \underbrace{\frac{1}{n}\mathbf{1}^T}_{\boldsymbol{B}}\boldsymbol{y} \sim \mathrm{N}(\theta\mathbf{1}, \frac{\sigma^2(1-\rho-n\rho)}{n}).$$

    We can easily check $\boldsymbol{A}\Sigma\boldsymbol{B} = 0$ so from the result in class, $\bar{Y}$ and $\sum_i (Y_i - \bar{Y})^2$ are independent.

14. (a) $x_2 \sim \mathrm{N}(\mu_2, 1)$ and $x_3 \sim \mathrm{N}(\mu_3, 1)$
    (b) From a result in the lecture (or in Monahan p.116),

$$x_1|x_2, x_3 \sim \mathrm{N}\left(\mu_1 + \frac{\rho}{1-\rho^2}(x_2 - \mu_2) - \frac{\rho^2}{1-\rho^2}(x_3 - \mu_3), 1 - \frac{\rho^2}{1-\rho^2}\right)$$

    If $\rho = 0$, $x_1 \mid x_2, x_3$ is the same as the marginal of x_1, $\mathrm{N}(\mu_1, 1)$.

(c)

$$\begin{aligned}\lambda &= \boldsymbol{Ax} \\ \begin{bmatrix} x_1+x_2+x_3 \\ x_1-x_2-x_3 \end{bmatrix} &= \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\end{aligned}$$

$\lambda_1 = x_1+x_2+x_3$ and $\lambda_2 = x_1-x_2-x_3$ are independently distributed if $\text{Cov}(\lambda_1, \lambda_2) = 0$.

$$\begin{aligned}\text{Cov}(\lambda) &= \text{Cov}(\boldsymbol{Ax}) \\ &= \boldsymbol{A}\text{Cov}(\boldsymbol{x})\boldsymbol{A}^T \\ &= \begin{bmatrix} 3+4\rho & -1-2\rho \\ -1-2\rho & 3 \end{bmatrix}\end{aligned}$$

So $x_1+x_2+x_3$ and $x_1-x_2-x_3$ are uncorrelated when $\rho = -\frac{1}{2}$.

15.

$$\boldsymbol{x} \sim \text{N}_3(\mu, \boldsymbol{\Sigma})$$

where

$$\mu = \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix} \quad \text{and } \boldsymbol{\Sigma} = \begin{bmatrix} 10 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 12 \end{bmatrix}.$$

16. (a) This is a normal. Using a result in the lecture, we can find $x_1 \mid x_2, x_3 \sim \text{N}(\frac{1}{5}(3x_2 - x_3), \frac{17}{5})$.

(b) Define $A = (4, -6, 1)^T$. Then the shifted linear combination $4x_1 - 6x_2 + x_3$ is distributed as $\text{N}(0, A\Sigma A^T)$, and hence the linear combination we originally wanted is distributed as $\text{N}(-18, A\Sigma A^T)$

17. (a) See result 5.15 (page 112 in Monahan's book). Observe that $\boldsymbol{A}$ is idempotent with rank 2. Thus, $\boldsymbol{y}^T\boldsymbol{Ay}/\sigma^2 \sim \chi^2(2, \frac{1}{2\sigma^2}\boldsymbol{m}^T\boldsymbol{Am})$.

(b) See result 5.16 (page 113 in Monahan's book). Observe that

$$\boldsymbol{BVA} = \sigma^2 \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix} \neq \mathbf{0}$$

So $\boldsymbol{y}^T\boldsymbol{Ay}$ and $\boldsymbol{By}$ are not independent.

(c) Let

$$y_1 + y_2 + y_3 = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \boldsymbol{y} = \boldsymbol{Cy}$$

Find that

$$\boldsymbol{CVA} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$$

So $\boldsymbol{y}^T\boldsymbol{Ay}$ and $y_1 + y_2 + y_3$ are independent.

18. (a)

$$\mathbf{AA} = \mathbf{X}(\mathbf{X^T X})^{-}\mathbf{X^T X}(\mathbf{X^T X})^{-}\mathbf{X^T} = \mathbf{X}(\mathbf{X^T X})^{-}\mathbf{X^T} = \mathbf{A}$$

because $(\mathbf{X^T X})^{-}\mathbf{X^T}$ is the generalized inverse of $\mathbf{X}$, and

$$(\mathbf{I} - \mathbf{A})(\mathbf{I} - \mathbf{A}) = (\mathbf{I} - \mathbf{A})$$

so $\boldsymbol{A}$ and $\mathbf{I}-\mathbf{A}$ are idempotent.

The ranks of the matrices are:

$$rank(\boldsymbol{A}) = trace(\boldsymbol{A}) = p$$

$$rank(\boldsymbol{I}-\boldsymbol{A}) = trace(\boldsymbol{I}-\boldsymbol{A}) = n-p.$$

(b)

$$\mathrm{E}(\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y}) = \|\boldsymbol{X}\boldsymbol{b}\|^2 + p\sigma^2$$
$$\mathrm{E}[\boldsymbol{y}^T(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{y}] = (n-p)\sigma^2$$

(c)

$$\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y}/\sigma^2 \sim \chi^2\left(p, \phi = \frac{1}{2\sigma^2}(\boldsymbol{X}\boldsymbol{b})^{\mathbf{T}}(\boldsymbol{X}\boldsymbol{b})\right)$$
$$\boldsymbol{y}^T(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{y}/\sigma^2 \sim \chi^2(n-p)$$

(d) $\mathbf{Ay}$ and $(\mathbf{I}-\mathbf{A})\mathbf{y}$ are independent since

$$\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{I}-\boldsymbol{A}\end{bmatrix}\boldsymbol{y} \sim N_2\left(\begin{bmatrix}\boldsymbol{X}\boldsymbol{b}\\ \mathbf{0}\end{bmatrix}, \begin{bmatrix}\mathrm{A} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}-\mathbf{A}\end{bmatrix}\right),$$

which implies $\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y} = \|\mathbf{Ay}\|^2$ and $\boldsymbol{y}^T(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{y} = \|(\mathbf{I}-\mathbf{A})\mathbf{y}\|^2$ are independent.

(e)

$$\frac{\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y}/p}{\boldsymbol{y}^T(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{y}/(n-p)} \sim F\left(p, n-p, \phi = \frac{1}{2\sigma^2}(\boldsymbol{X}\boldsymbol{b})^{\mathbf{2}}(\boldsymbol{X}\boldsymbol{b})\right)$$