Solution to HW 1

1\. a. $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 8 & -6 \\ 4 & 1 & 7 \end{bmatrix}$

solve for $c_1 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 4 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 0 \\ 8 \\ 1 \end{pmatrix} + c_3 \begin{pmatrix} 1 \\ 2 \\ -6 \\ 7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

$$\Rightarrow \begin{cases} c_1 + c_2 + c_3 = 0 \\ c_1 = -2c_3 \end{cases} \Rightarrow c_2 = c_3 = -\tfrac{1}{2} c_1$$

and this could solve

$$\begin{cases} c_1 + 8c_2 - 6c_3 = 0 \\ 4c_1 + c_2 + 7c_3 = 0 \end{cases}$$

$\therefore$ Not linearly independent

actually, $A \cdot \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

$B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 8 & 0 \\ 4 & 1 & -2 \end{bmatrix}$

solve for $c_1 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 4 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 0 \\ 8 \\ 1 \end{bmatrix} + c_3 \begin{pmatrix} 1 \\ 2 \\ 0 \\ -2 \end{pmatrix} = 0$

first two equations

$\Rightarrow c_2 = c_3 = -\frac{1}{2} c_1$

plug into the 3rd & 4th equations

$\Rightarrow 3c_1 = 0 \Rightarrow c_1 = c_2 = c_3 = 0$

$\Rightarrow$ linearly independent

1.b Solve

$$\left| \begin{bmatrix} 2-\lambda & 1 \\ 2 & 4-\lambda \end{bmatrix} \right| = 0$$

$$\Rightarrow \lambda = 3 \pm \sqrt{3}$$

for $\lambda_1 = 3+\sqrt{3}$ solve v_1 s.t. $\begin{bmatrix} 2-(3+\sqrt{3}) & 1 \\ 2 & 4-(3+\sqrt{3}) \end{bmatrix} v_1 = 0$

$$\Rightarrow v_1 = c \begin{pmatrix} \sqrt{3}-1 \\ 2 \end{pmatrix}$$

similarly, for $\lambda_2 = 3-\sqrt{3} \Rightarrow v_2 = c \begin{pmatrix} -(\sqrt{3}+1) \\ 2 \end{pmatrix}$

~~1.c. $\mathrm{Cov}(X_1+2X_2, 3X_2+X_3)$~~

2. a. linear combination of ~~joint~~ multi- normal r.v.'s is still normal ; what remains to derive is the mean & variance of that normal distribution

$$E[3X_1 + X_2 + X_3]$$

$$= 3E[X_1] + E[X_2] + E[X_3] = 0$$

$$\mathrm{Var}[3X_1 + X_2 + X_3]$$

$$= 9\mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \mathrm{Var}(X_3) + 2\cdot 3\cdot \mathrm{Cov}(X_1, X_2) + 2\mathrm{Cov}(X_2, X_3) + 2\cdot 3\cdot \mathrm{Cov}(X_1, X_3)$$

$$= 25.6$$

$$\Rightarrow 3X_1 + X_2 + X_3 \sim N(0, 25.6)$$

2.b. Define $y_1 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $y_2 = (x_3)$

Then $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \right)$

where $\mu_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\mu_2 = (0)$

$\Sigma_{11} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, $\Sigma_{12} = \begin{pmatrix} 0.6 \\ 0.5 \end{pmatrix}$, $\Sigma_{21} = \Sigma_{12}^T$, $\Sigma_{22} = (1)$

Using the formula

$$y_1 \mid y_2 = b \sim N\left(\mu_1 + \Sigma_{12} \cdot \Sigma_{22}^{-1} (b - \mu_2),\ \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \right)$$

with $b = 3$

$$\Rightarrow y_1 \mid y_2 = 3 \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0.6 \\ 0.5 \end{pmatrix} (1)^{-1} ((3) - (0)),\right.$$

$$\left. \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} - \begin{bmatrix} 0.6 \\ 0.5 \end{bmatrix} [1]^{-1} [0.6 \ \ 0.5] \right)$$

$$\Rightarrow y_1 \mid y_2 = 3 \sim N\left(\begin{pmatrix} 1.8 \\ 1.5 \end{pmatrix}, \begin{pmatrix} 1.64 & -0.30 \\ -0.30 & 1.75 \end{pmatrix} \right)$$

2.c. $\mathrm{Cov}(X_1 + 2X_2,\ 3X_2 + X_3)$

$= \mathrm{Cov}(X_1, 3X_2) + \mathrm{Cov}(2X_2, 3X_2) + \mathrm{Cov}(2X_2, X_3) + \mathrm{Cov}(X_1, X_3)$

$= 3 \cdot 0 + 6 \cdot 2 + 2 \cdot 0.5 + 0.6$

$= 13.6$

3. Any linear combination of multi-normal r.v.'s is still normal with the mean

$$E[Y] = E[\Sigma_{i=1}^{k} a_i X_i]$$

$$= \Sigma_{i=1}^{k} a_i E[X_i]$$

$$= \Sigma_{i=1}^{k} a_i \mu_i$$

and variance

$$Var[Y] = Var[\Sigma_{i=1}^{k} a_i X_i]$$

$$= \Sigma_{i=1}^{k} Var[a_i X_i] + \Sigma_{i=1}^{k} \Sigma_{j \neq i, j=1}^{k} \overset{Cov}{\cancel{Var}}[a_i X_i, a_j X_j]$$

$$= \Sigma_{i=1}^{k} a_i^2 \, Var(X_i) + 2 \Sigma_{i=1}^{k} \Sigma_{j=1, j \neq i}^{k} a_i a_j Cov(X_i, X_j)$$

$$= \Sigma_{i=1}^{k} a_i^2 \Sigma_{ii} + \Sigma_{i=1}^{k} \Sigma_{j=1, j \neq i}^{k} a_i a_j \Sigma_{ij}$$

$$\therefore \sum_{1}^{k} a_i X_i \sim N\left(\sum_{1}^{k} a_i \mu_i + \sum_{1}^{k} a_i^2 \Sigma_{ii} + \Sigma_{i=1}^{k} \Sigma_{j=1, j \neq i}^{k} a_i a_j \Sigma_{ij} \right)$$

4. (a)

$E[\hat{\beta}_W]$

$= E[(X^TV^{-1}X)^{-1}X^TV^{-1}Y]$

$= \cancel{E}(X^TV^{-1}X)^{-1}X^TV^{-1}E[Y]$

$= (X^TV^{-1}X)^{-1}X^TV^{-1}X\beta$

$= \beta$

because $E[Y] = E[X\beta + \varepsilon]$
$= X\beta + E(\varepsilon)$
$= X\beta$

(b) $Cov(\hat{\beta}_W)$

$= Cov((X^TV^{-1}X)^{-1}X^TV^{-1}Y)$

$= (X^TV^{-1}X)^{-1}X^TV^{-1}\ Cov(Y)\ V^{-1}X(X^TV^{-1}X)^{-1}$

$(\because Cov(Y) = Var(X\beta + \varepsilon) = Var(\varepsilon) = \sigma^2V)$

$= (X^TV^{-1}X)^{-1}X^TV^{-1}(\sigma^2V)V^{-1}X(X^TV^{-1}X)^{-1}$

$= \sigma^2(X^TV^{-1}X)^{-1}$

(c) If ε is multi-normal, then we immediately know that $\hat{\beta}_W \sim N(\beta, \sigma^2(X^TV^{-1}X)^{-1})$

Otherwise, only the mean & variance don't give enough information on the exact distribution of $\hat{\beta}_W$

(d) The variance of the average of m_i equally variable observations could be calculated as

$$Var(y_i) = Var\left(\sum_{i=1}^{m_i} X_i / m_i \right)$$

$$= \left(\frac{1}{m_i}\right)^2 \sum_{i=1}^{m_i} \sigma^2$$

$$= \frac{\sigma^2}{m_i}$$

Therefore the covariance matrix $Cov(Y) = \sigma^2 V$

where $V = \begin{bmatrix} \frac{1}{m_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{m_n} \end{bmatrix}$

In other words

- all off-diagnal entries $= 0$ because we assume the observations are independent
- Diagonal terms are inverse-weighted by the number of observations as more observations mean more information on that $y_i \Rightarrow$ less variable