STAT 32950

Robert Winter

HW #4

1. Data

- Trt 1: [6 7], [5 9], [8 6], [4 9], [7 9]
- Trt 2: [3 3], [1 6], [2 3]
- Trt 3: [2 3], [5 1], [3 1], [2 3]

1a. Break up the observations (indexed by j) into mean, trt (indexed by t), & residual components

$$x_{tj} = \bar{x} + (\bar{x}_t - \bar{x}) + (x_{tj} - \bar{x}_t)$$

by constructing data arrays for each component variable.

- First Component:
  - $\bar{x} = \frac{1}{12}(6+5+8+4+7+3+1+2+2+5+3+2) = \frac{1}{12}(48) = 4$
  - $\bar{x}_1 = \frac{1}{5}(6+5+8+4+7) = \frac{1}{5}(30) = 6$
  - $\bar{x}_2 = \frac{1}{3}(3+1+2) = \frac{1}{3}(6) = 2$
  - $\bar{x}_3 = \frac{1}{4}(2+5+3+2) = \frac{1}{4}(12) = 3$
  - Array:

$$\begin{bmatrix} 6 & 5 & 8 & 4 & 7 \\ 3 & 1 & 2 & & \\ 2 & 5 & 3 & 2 & \end{bmatrix} = \begin{bmatrix} 4 & 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & & \\ 4 & 4 & 4 & 4 & \end{bmatrix} + \begin{bmatrix} 6-4 & 6-4 & 6-4 & 6-4 & 6-4 \\ 2-4 & 2-4 & 2-4 & & \\ 3-4 & 3-4 & 3-4 & 3-4 & \end{bmatrix}$$

$$+ \begin{bmatrix} 6-6 & 5-6 & 8-6 & 4-6 & 7-6 \\ 3-2 & 1-2 & 2-2 & & \\ 2-3 & 5-3 & 3-3 & 2-3 & \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 6 & 5 & 8 & 4 & 7 \\ 3 & 1 & 2 & & \\ 2 & 5 & 3 & 2 & \end{bmatrix} = \begin{bmatrix} 4 & 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & & \\ 4 & 4 & 4 & 4 & \end{bmatrix} + \begin{bmatrix} 2 & 2 & 2 & 2 & 2 \\ -2 & -2 & -2 & & \\ -1 & -1 & -1 & -1 & \end{bmatrix} + \begin{bmatrix} 0 & -1 & 2 & -2 & 1 \\ 1 & -1 & 0 & & \\ -1 & 2 & 0 & -1 & \end{bmatrix}$$

- Second Component
  - $\bar{x} = \frac{1}{12}(7+9+6+9+9+3+6+3+3+1+1+3) = \frac{1}{12}(60) = 5$
  - $\bar{x}_1 = \frac{1}{5}(7+9+6+9+9) = \frac{1}{5}(40) = 8$
  - $\bar{x}_2 = \frac{1}{3}(3+6+3) = \frac{1}{3}(12) = 4$
  - $\bar{x}_3 = \frac{1}{4}(3+1+1+3) = \frac{1}{4}(8) = 2$

· Array:

$$\begin{bmatrix} 7 & 9 & 6 & 9 & 9 \\ 3 & 6 & 3 & & \\ 3 & 1 & 1 & 3 & \end{bmatrix} = \begin{bmatrix} 5 & 5 & 5 & 5 & 5 \\ 5 & 5 & 5 & & \\ 5 & 5 & 5 & 5 & \end{bmatrix} + \begin{bmatrix} 8-5 & 8-5 & 8-5 & 8-5 & 8-5 \\ 4-5 & 4-5 & 4-5 & & \\ 2-5 & 2-5 & 2-5 & 2-5 & \end{bmatrix}$$

$$+ \begin{bmatrix} 7-8 & 9-8 & 6-8 & 9-8 & 9-8 \\ 3-4 & 6-4 & 3-4 & & \\ 3-2 & 1-2 & 1-2 & 3-2 & \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 7 & 9 & 6 & 9 & 9 \\ 3 & 6 & 3 & & \\ 3 & 1 & 1 & 3 & \end{bmatrix} = \begin{bmatrix} 5 & 5 & 5 & 5 & 5 \\ 5 & 5 & 5 & & \\ 5 & 5 & 5 & 5 & \end{bmatrix} + \begin{bmatrix} 3 & 3 & 3 & 3 & 3 \\ -1 & -1 & -1 & & \\ -3 & -3 & -3 & -3 & \end{bmatrix} + \begin{bmatrix} -1 & 1 & -2 & 1 & 1 \\ -1 & 2 & -1 & & \\ 1 & -1 & -1 & 1 & \end{bmatrix}$$

1b. Use part (a) to construct the one-way MANOVA table.

- 3 levels of treatment $\Rightarrow g=3 \Rightarrow g-1=2$
- 12 observations $\Rightarrow n=12 \Rightarrow n-g=9$
- $B = \sum_{\ell=1}^{3} n_\ell (\bar{x}_\ell - \bar{x})(\bar{x}_\ell - \bar{x})^T$

$$= n_1(\bar{x}_1-\bar{x})(\bar{x}_1-\bar{x})^T + n_2(\bar{x}_2-\bar{x})(\bar{x}_2-\bar{x})^T + n_3(\bar{x}_3-\bar{x})(\bar{x}_3-\bar{x})^T$$

$$= 5\left(\begin{bmatrix}6\\8\end{bmatrix}-\begin{bmatrix}4\\5\end{bmatrix}\right)\left(\begin{bmatrix}6\\8\end{bmatrix}-\begin{bmatrix}4\\5\end{bmatrix}\right)^T + 3\left(\begin{bmatrix}2\\4\end{bmatrix}-\begin{bmatrix}4\\5\end{bmatrix}\right)\left(\begin{bmatrix}2\\4\end{bmatrix}-\begin{bmatrix}4\\5\end{bmatrix}\right)^T$$

$$+ 4\left(\begin{bmatrix}3\\2\end{bmatrix}-\begin{bmatrix}4\\5\end{bmatrix}\right)\left(\begin{bmatrix}3\\2\end{bmatrix}-\begin{bmatrix}4\\5\end{bmatrix}\right)^T$$

$$= 5\begin{bmatrix}2\\3\end{bmatrix}\begin{bmatrix}2 & 3\end{bmatrix} + 3\begin{bmatrix}-2\\-1\end{bmatrix}\begin{bmatrix}-2 & -1\end{bmatrix} + 4\begin{bmatrix}-1\\-3\end{bmatrix}\begin{bmatrix}-1 & -3\end{bmatrix}$$

$$= 5\begin{bmatrix}4 & 6\\6 & 9\end{bmatrix} + 3\begin{bmatrix}4 & 2\\2 & 1\end{bmatrix} + 4\begin{bmatrix}1 & 3\\3 & 9\end{bmatrix}$$

$$= \begin{bmatrix}20 & 30\\30 & 45\end{bmatrix} + \begin{bmatrix}12 & 6\\6 & 3\end{bmatrix} + \begin{bmatrix}4 & 12\\12 & 36\end{bmatrix}$$

$$= \begin{bmatrix}36 & 48\\48 & 84\end{bmatrix}$$

$$\bullet\; W = \sum_{\ell=1}^{3} \sum_{j=1}^{n_\ell} (x_{\ell j} - \bar{x}_\ell)(x_{\ell j} - \bar{x}_\ell)^T$$

$$= \sum_{j=1}^{5} (x_{1j} - \bar{x}_1)(x_{1j} - \bar{x}_1)^T + \sum_{j=1}^{3} (x_{2j} - \bar{x}_2)(x_{2j} - \bar{x}_2)^T + \sum_{j=1}^{4} (x_{3j} - \bar{x}_3)(x_{3j} - \bar{x}_3)^T$$

$$= \left(\begin{bmatrix}6\\7\end{bmatrix} - \begin{bmatrix}6\\8\end{bmatrix}\right)\left(\begin{bmatrix}6\\7\end{bmatrix} - \begin{bmatrix}6\\8\end{bmatrix}\right)^T + \left(\begin{bmatrix}5\\9\end{bmatrix} - \begin{bmatrix}6\\8\end{bmatrix}\right)\left(\begin{bmatrix}5\\9\end{bmatrix} - \begin{bmatrix}6\\8\end{bmatrix}\right)^T + \left(\begin{bmatrix}8\\6\end{bmatrix} - \begin{bmatrix}6\\8\end{bmatrix}\right)\left(\begin{bmatrix}8\\6\end{bmatrix} - \begin{bmatrix}6\\8\end{bmatrix}\right)^T$$

$$+ \left(\begin{bmatrix}4\\9\end{bmatrix} - \begin{bmatrix}6\\8\end{bmatrix}\right)\left(\begin{bmatrix}4\\9\end{bmatrix} - \begin{bmatrix}6\\8\end{bmatrix}\right)^T + \left(\begin{bmatrix}7\\9\end{bmatrix} - \begin{bmatrix}6\\8\end{bmatrix}\right)\left(\begin{bmatrix}7\\9\end{bmatrix} - \begin{bmatrix}6\\8\end{bmatrix}\right)^T + \left(\begin{bmatrix}3\\3\end{bmatrix} - \begin{bmatrix}2\\4\end{bmatrix}\right)\left(\begin{bmatrix}3\\3\end{bmatrix} - \begin{bmatrix}2\\4\end{bmatrix}\right)^T$$

$$+ \left(\begin{bmatrix}1\\6\end{bmatrix} - \begin{bmatrix}2\\4\end{bmatrix}\right)\left(\begin{bmatrix}1\\6\end{bmatrix} - \begin{bmatrix}2\\4\end{bmatrix}\right)^T + \left(\begin{bmatrix}2\\3\end{bmatrix} - \begin{bmatrix}2\\4\end{bmatrix}\right)\left(\begin{bmatrix}2\\3\end{bmatrix} - \begin{bmatrix}2\\4\end{bmatrix}\right)^T + \left(\begin{bmatrix}2\\3\end{bmatrix} - \begin{bmatrix}3\\2\end{bmatrix}\right)\left(\begin{bmatrix}2\\3\end{bmatrix} - \begin{bmatrix}3\\2\end{bmatrix}\right)^T$$

$$+ \left(\begin{bmatrix}5\\1\end{bmatrix} - \begin{bmatrix}3\\2\end{bmatrix}\right)\left(\begin{bmatrix}5\\1\end{bmatrix} - \begin{bmatrix}3\\2\end{bmatrix}\right)^T + \left(\begin{bmatrix}3\\1\end{bmatrix} - \begin{bmatrix}3\\2\end{bmatrix}\right)\left(\begin{bmatrix}3\\1\end{bmatrix} - \begin{bmatrix}3\\2\end{bmatrix}\right)^T + \left(\begin{bmatrix}2\\3\end{bmatrix} - \begin{bmatrix}3\\2\end{bmatrix}\right)\left(\begin{bmatrix}2\\3\end{bmatrix} - \begin{bmatrix}3\\2\end{bmatrix}\right)^T$$

$$= \begin{bmatrix}0\\-1\end{bmatrix}\begin{bmatrix}0 & -1\end{bmatrix} + \begin{bmatrix}-1\\1\end{bmatrix}\begin{bmatrix}-1 & 1\end{bmatrix} + \begin{bmatrix}2\\-2\end{bmatrix}\begin{bmatrix}2 & -2\end{bmatrix} + \begin{bmatrix}-2\\1\end{bmatrix}\begin{bmatrix}-2 & 1\end{bmatrix} + \begin{bmatrix}1\\1\end{bmatrix}\begin{bmatrix}1 & 1\end{bmatrix}$$

$$+ \begin{bmatrix}1\\-1\end{bmatrix}\begin{bmatrix}1 & -1\end{bmatrix} + \begin{bmatrix}-1\\2\end{bmatrix}\begin{bmatrix}-1 & 2\end{bmatrix} + \begin{bmatrix}0\\-1\end{bmatrix}\begin{bmatrix}0 & -1\end{bmatrix}$$

$$+ \begin{bmatrix}-1\\1\end{bmatrix}\begin{bmatrix}-1 & 1\end{bmatrix} + \begin{bmatrix}2\\-1\end{bmatrix}\begin{bmatrix}2 & -1\end{bmatrix} + \begin{bmatrix}0\\-1\end{bmatrix}\begin{bmatrix}0 & -1\end{bmatrix} + \begin{bmatrix}-1\\1\end{bmatrix}\begin{bmatrix}-1 & 1\end{bmatrix}$$

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

$$+ \begin{bmatrix}1 & -1\\-1 & 1\end{bmatrix} + \begin{bmatrix}1 & -2\\-2 & 4\end{bmatrix} + \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$$

$$+ \begin{bmatrix}1 & -1\\-1 & 1\end{bmatrix} + \begin{bmatrix}4 & -2\\-2 & 1\end{bmatrix} + \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} + \begin{bmatrix}1 & -1\\-1 & 1\end{bmatrix}$$

$$= \begin{bmatrix}10 & -6\\-6 & 8\end{bmatrix} + \begin{bmatrix}2 & -3\\-3 & 6\end{bmatrix} + \begin{bmatrix}6 & -4\\-4 & 4\end{bmatrix}$$

$$= \begin{bmatrix}18 & -13\\-13 & 18\end{bmatrix}$$

• Thus, the MANOVA table is:

Source of Variation	Matrix of Sum of Squares & Cross-Products	df
Treatments	$B = \begin{bmatrix} 36 & 48 \\ 48 & 84 \end{bmatrix}$	2
Residuals	$W = \begin{bmatrix} 18 & -13 \\ -13 & 18 \end{bmatrix}$	9
Total	$B+W = \begin{bmatrix} 54 & 35 \\ 35 & 102 \end{bmatrix}$	11

1c. Evaluate $|W|$, $|B|$, $\Lambda^* = \dfrac{|W|}{|B+W|}$

• $|W| = \det\left(\begin{bmatrix} 18 & -13 \\ -13 & 18 \end{bmatrix}\right)$

$= (18)(18) - (-13)(-13)$

$= 324 - 169$

$\Rightarrow |W| = 155$

• $|B| = \det\left(\begin{bmatrix} 36 & 48 \\ 48 & 84 \end{bmatrix}\right)$

$= (36)(84) - (48)(48)$

$= 3024 - 2304$

$\Rightarrow |B| = 720$

• $|B+W| = \det\left(\begin{bmatrix} 54 & 35 \\ 35 & 102 \end{bmatrix}\right)$

$= (54)(102) - (35)(35)$

$= 5508 - 1225$

$= 4283$

- So, $\Lambda^* = \frac{|W|}{|B+W|}$

$\Rightarrow$ $\Lambda^* = \frac{155}{4283} \approx 0.036$

1d. Use Bartlett's approximation $-\log(\Lambda^*)(n-1-\frac{p+g}{2}) \approx_{H_0} \chi^2_{p(g-1)}$ to find the observed p-value from the data.

- $-\log(\Lambda^*)(n-1-\frac{p+g}{2}) \approx -\log(0.036)(12-1-\frac{2+3}{2}) = -\frac{17}{2}\log(0.036)$
 ≈ 28.211
- $\chi^2_{p(g-1)} = \chi^2_{2(3-1)} = \chi^2_4$
- Let $X \sim \chi^2_4$. Using R,
 $p = \mathbb{P}(X > 28.211)$
 $\Rightarrow$ $p \approx 1.130 \times 10^{-5}$

1e. Write the hypothesis test (H_0, H_1) that the p-value in Part (d) refers to. Describe in words.

- $\begin{cases} H_0: \tau_1 = \tau_2 = \tau_3 = 0 \\ H_1: \tau_t \neq 0 \text{ for some } t \in \{1,2,3\} \end{cases}$

where τ_t is the treatment effect for treatment t, $0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, & we have imposed the constraint $\Sigma_{\ell=1}^{3} n_\ell \tau_\ell = 0$ for identification purposes.

- Intuitively, the null hypothesis is that all three treatments have no effect on individuals' outcomes $x = (x_1, x_2)$ (so their effects — or lack thereof — are no different from one another's).
- The alternative hypothesis is that at least one treatment has a nonzero effect on individuals' outcomes $x = (x_1, x_2)$ (so that at least one treatment's effect is different than the others').

5. $X_1 \in \mathbb{R}^{d_1}$, $X_2 \in \mathbb{R}^{d_2}$, $X_3 \in \mathbb{R}^{d_3}$ $(d_i \geq 2)$ are jointly MVN, w/

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & 0 \\ \Sigma_{31} & 0 & \Sigma_{33} \end{bmatrix},$$

where Σ_{ii} are positive definite.

5a(i) Derive $\mathbb{E}[X_1 \mid X_2 = x_2, X_3 = x_3]$.

- We partition μ & Σ as follows:

$$\mu = \begin{bmatrix} \mu_1 \\ \hdashline \mu_2 \\ \mu_3 \end{bmatrix} \begin{matrix} \leftarrow \mu_A \\ \\ \leftarrow \mu_B \end{matrix} \qquad \Sigma = \left[\begin{array}{c:cc} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \hdashline \Sigma_{21} & \Sigma_{22} & 0 \\ \Sigma_{31} & 0 & \Sigma_{33} \end{array}\right]$$

Σ_A $\quad$ Σ_B $\quad$ Σ_C $\quad$ Σ_D

- Then

$$\mathbb{E}[X_1 \mid X_2 = x_2, X_3 = x_3] = \mu_A + \Sigma_B \Sigma_D^{-1}\left(\begin{bmatrix} x_2 \\ x_3 \end{bmatrix} - \mu_B\right)$$

$$= \mu_1 + \begin{bmatrix} \Sigma_{12} & \Sigma_{13} \end{bmatrix} \begin{bmatrix} \Sigma_{22} & 0 \\ 0 & \Sigma_{33} \end{bmatrix}^{-1} \left(\begin{bmatrix} x_2 \\ x_3 \end{bmatrix} - \begin{bmatrix} \mu_2 \\ \mu_3 \end{bmatrix}\right)$$

- Since Σ_{22}, Σ_{33} are positive definite, they are invertible: Σ_{22}^{-1}, Σ_{33}^{-1} exist. So,

$$\begin{bmatrix} \Sigma_{22} & 0 \\ 0 & \Sigma_{33} \end{bmatrix}^{-1} = \begin{bmatrix} \Sigma_{22}^{-1} + \Sigma_{22}^{-1} 0 (\Sigma_{33} - 0 \Sigma_{22}^{-1} 0)^{-1} 0 \Sigma_{22}^{-1} & -\Sigma_{22}^{-1} 0 (\Sigma_{33} - 0 \Sigma_{22}^{-1} 0)^{-1} \\ -(\Sigma_{33} - 0 \Sigma_{22}^{-1} 0)^{-1} 0 \Sigma_{22}^{-1} & (\Sigma_{33} - 0 \Sigma_{22}^{-1} 0)^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} \Sigma_{22}^{-1} + \Sigma_{22}^{-1} 0 (\Sigma_{33})^{-1} 0 \Sigma_{22}^{-1} & -\Sigma_{22}^{-1} 0 (\Sigma_{33})^{-1} \\ -(\Sigma_{33})^{-1} 0 \Sigma_{22}^{-1} & (\Sigma_{33})^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} \Sigma_{22}^{-1} & 0 \\ 0 & \Sigma_{33}^{-1} \end{bmatrix}$$

- So,

$$\mathbb{E}[X_1 \mid X_2 = x_2, X_3 = x_3] = \mu_1 + \begin{bmatrix} \Sigma_{12} & \Sigma_{13} \end{bmatrix} \begin{bmatrix} \Sigma_{22}^{-1} & 0 \\ 0 & \Sigma_{33}^{-1} \end{bmatrix} \begin{bmatrix} x_2 - \mu_2 \\ x_3 - \mu_3 \end{bmatrix}$$

$$= \mu_1 + [\Sigma_{12} \quad \Sigma_{13}] \begin{bmatrix} \Sigma_{22}^{-1}(x_2 - \mu_2) \\ \Sigma_{33}^{-1}(x_3 - \mu_3) \end{bmatrix}$$

$$\Rightarrow \mathbb{E}[X_1 | X_2 = x_2, X_3 = x_3] = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2) + \Sigma_{13}\Sigma_{33}^{-1}(x_3 - \mu_3)$$

5a(ii) Derive $Var(X_1 | X_2 = x_2, X_3 = x_3)$.

$$Var(X_1 | X_2 = x_2, X_3 = x_3) = \Sigma_A - \Sigma_B \Sigma_D^{-1} \Sigma_C$$

$$= \Sigma_{11} - [\Sigma_{12} \quad \Sigma_{13}] \begin{bmatrix} \Sigma_{22} & 0 \\ 0 & \Sigma_{33} \end{bmatrix}^{-1} \begin{bmatrix} \Sigma_{21} \\ \Sigma_{31} \end{bmatrix}$$

$$= \Sigma_{11} - [\Sigma_{12} \quad \Sigma_{13}] \begin{bmatrix} \Sigma_{22}^{-1} & 0 \\ 0 & \Sigma_{33}^{-1} \end{bmatrix} \begin{bmatrix} \Sigma_{21} \\ \Sigma_{31} \end{bmatrix}$$

$$= \Sigma_{11} - [\Sigma_{12} \quad \Sigma_{13}] \begin{bmatrix} \Sigma_{22}^{-1}\Sigma_{21} + 0\Sigma_{31} \\ 0\Sigma_{21} + \Sigma_{33}^{-1}\Sigma_{31} \end{bmatrix}$$

$$= \Sigma_{11} - [\Sigma_{12} \quad \Sigma_{13}] \begin{bmatrix} \Sigma_{22}^{-1}\Sigma_{21} \\ \Sigma_{33}^{-1}\Sigma_{31} \end{bmatrix}$$

$$\Rightarrow Var(X_1 | X_2 = x_2, X_3 = x_3) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} - \Sigma_{13}\Sigma_{33}^{-1}\Sigma_{31}$$

5b Derive conditional distribution of X_1 given $X_2 + X_3 = x_0$. Start by finding distribution of $X_2 + X_3$.

- Let $Y = X_2 + X_3$.
- Write $X_1 = (X_{11}, \ldots, X_{1d_1})$, $X_2 = (X_{21}, \ldots, X_{2d_2})$, $X_3 = (X_{31}, \ldots, X_{3d_2})$ (since $d_3 = d_2$ necessarily).
- Since (X_1, X_2, X_3) are jointly MVN, any linear combination $a_1 X_{11} + \cdots + a_{d_1} X_{1d_1} + b_1 X_{21} + \cdots + b_{d_2} X_{2d_2} + c_1 X_{31} + \cdots + c_{d_2} X_{3d_2}$ has normal distribution $\forall a_i, b_i, c_i \in \mathbb{R}$. In particular, if $a_i = 0$ $\forall i = 1, \ldots, d_1$, then $b_1 X_{21} + \cdots + b_{d_2} X_{2d_2} + c_1 X_{31} + \cdots + c_{d_2} X_{3d_2}$ has normal distribution.
- $Y = X_2 + X_3 = (X_{21}, \ldots, X_{2d_2}) + (X_{31}, \ldots, X_{3d_2})$
  $= (X_{21} + X_{31}, \ldots, X_{2d_2} + X_{3d_2})$

· So, linear combinations of the components of Y have the form

$$e_1(X_{21}+X_{31}) + \cdots + e_{d_2}(X_{2d_2}+X_{3d_2})$$
$$= e_1 X_{21} + \cdots + e_{d_2} X_{2d_2} + e_1 X_{31} + \cdots + e_{d_2} X_{3d_2}$$

for some $e_1, \ldots, e_{d_2} \in \mathbb{R}$. But we've already shown that all linear combinations of $X_{21}, \ldots, X_{2d_2}, X_{31}, \ldots, X_{3d_2}$ have normal distribution, which means that all linear combinations of the components of Y have normal distribution. Thus, Y has multivariate normal distribution.

· In particular,

$$\begin{aligned} \mathbb{E}[Y] &= \mathbb{E}[X_2+X_3] \\ &= \mathbb{E}[X_2] + \mathbb{E}[X_3] \\ &= \mu_2+\mu_3 \end{aligned}$$

· And,

$$\begin{aligned} \mathrm{Var}(Y) &= \mathrm{Var}(X_2+X_3) \\ &= \mathbb{E}[(X_2+X_3)(X_2+X_3)^T] - \mathbb{E}[X_2+X_3]\mathbb{E}[X_2+X_3]^T \\ &= \mathbb{E}[(X_2+X_3)(X_2^T+X_3^T)] - (\mu_2+\mu_3)(\mu_2+\mu_3)^T \\ &= \mathbb{E}[X_2X_2^T + X_2X_3^T + X_3X_2^T + X_3X_3^T] - (\mu_2+\mu_3)(\mu_2^T+\mu_3^T) \\ &= \mathbb{E}[X_2X_2^T] + \mathbb{E}[X_2X_3^T] + \mathbb{E}[X_3X_2^T] + \mathbb{E}[X_3X_3^T] \\ &\quad - \mu_2\mu_2^T - \mu_2\mu_3^T - \mu_3\mu_2^T - \mu_3\mu_3^T \\ &= (\mathbb{E}[X_2X_2^T]-\mu_2\mu_2^T) + (\mathbb{E}[X_2X_3^T]-\mu_2\mu_3^T) \\ &\quad + (\mathbb{E}[X_3X_2^T]-\mu_3\mu_2^T) + (\mathbb{E}[X_3X_3^T]-\mu_3\mu_3^T) \\ &= \mathrm{Var}(X_2) + \mathrm{Cov}(X_2,X_3) + \mathrm{Cov}(X_3,X_2) + \mathrm{Var}(X_3) \\ &= \Sigma_{22} + \Sigma_{23} + \Sigma_{32} + \Sigma_{33} \\ &= \Sigma_{22} + 0 + 0 + \Sigma_{33} \\ &= \Sigma_{22} + \Sigma_{33} \end{aligned}$$

· Thus,

$$Y = X_2+X_3 \sim N_{d_2}(\mu_2+\mu_3, \Sigma_{22}+\Sigma_{33})$$

- Note that linear combinations of the components of X_1 & Y have the form

$a_1 X_{11} + \cdots + a_{d_1} X_{1d_1} + e_1(X_{21} + X_{31}) + \cdots + e_{d_2}(X_{2d_2} + X_{3d_2})$
$= a_1 X_{11} + \cdots + a_{d_1} X_{1d_1} + e_1 X_{21} + \cdots + e_{d_2} X_{2d_2} + e_1 X_{31} + \cdots + e_{d_2} X_{3d_2}$
for some a_i's, e_j's $\in \mathbb{R}$. But we've already shown that all linear combinations of $X_{11}, \ldots, X_{1d_1}, X_{21}, \ldots, X_{2d_2}, X_{31}, \ldots, X_{3d_2}$ have normal distribution, which means that all linear combinations of the components of X_1 & Y have normal distribution. Thus, (X_1, Y) are jointly multivariate normal.

- $\mathbb{E}\left[\begin{bmatrix} X_1 \\ Y \end{bmatrix}\right] = \begin{bmatrix} \mathbb{E}[X_1] \\ \mathbb{E}[Y] \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 + \mu_3 \end{bmatrix}$

- $\mathrm{Var}\left(\begin{bmatrix} X_1 \\ Y \end{bmatrix}\right) = \begin{bmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1, Y) \\ \mathrm{Cov}(Y, X_1) & \mathrm{Var}(Y) \end{bmatrix}$
$= \begin{bmatrix} \Sigma_{11} & \mathrm{Cov}(X_1, Y) \\ \mathrm{Cov}(Y, X_1) & \Sigma_{22} + \Sigma_{33} \end{bmatrix}$

- $\mathrm{Cov}(X_1, Y) = \mathbb{E}[X_1 Y^T] - \mathbb{E}[X_1]\mathbb{E}[Y]^T$
$= \mathbb{E}[X_1 (X_2 + X_3)^T] - \mu_1 \mathbb{E}[X_2 + X_3]^T$
$= \mathbb{E}[X_1 (X_2^T + X_3^T)] - \mu_1 (\mu_2 + \mu_3)^T$
$= \mathbb{E}[X_1 X_2^T + X_1 X_3^T] - \mu_1 (\mu_2^T + \mu_3^T)$
$= (\mathbb{E}[X_1 X_2^T] - \mu_1 \mu_2^T) + (\mathbb{E}[X_1 X_3^T] - \mu_1 \mu_3^T)$
$= \mathrm{Cov}(X_1, X_2) + \mathrm{Cov}(X_1, X_3)$
$= \Sigma_{12} + \Sigma_{13}$

- $\mathrm{Cov}(Y, X_1) = \mathrm{Cov}(X_1, Y)^T$
$= (\Sigma_{12} + \Sigma_{13})^T$
$= \Sigma_{12}^T + \Sigma_{13}^T$
$= \Sigma_{21} + \Sigma_{31}$

- So $\mathrm{Var}\left(\begin{bmatrix} X_1 \\ Y \end{bmatrix}\right) = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} + \Sigma_{13} \\ \Sigma_{21} + \Sigma_{31} & \Sigma_{22} + \Sigma_{33} \end{bmatrix}$

· That is,

$$\begin{bmatrix} X_1 \\ Y \end{bmatrix} \sim N_{d_1+d_2}\left(\begin{bmatrix} \overset{\mu_u}{\mu_1} \\ \underset{\mu_v}{\mu_2+\mu_3} \end{bmatrix}, \begin{bmatrix} \overset{\Sigma_Q}{\Sigma_{11}} & \overset{\Sigma_R}{\Sigma_{12}+\Sigma_{13}} \\ \underset{\Sigma_S}{\Sigma_{21}+\Sigma_{31}} & \underset{\Sigma_T}{\Sigma_{22}+\Sigma_{33}} \end{bmatrix} \right)$$

· Thus,

$$X_1 \mid Y = x_0 \sim N_{d_1}\left(\mu_u + \Sigma_R \Sigma_T^{-1}(x_0 - \mu_v),\ \Sigma_Q - \Sigma_R \Sigma_T^{-1} \Sigma_S \right)$$

$$\Rightarrow X_1 \mid X_2 + X_3 = x_0 \sim N_{d_1}\Big(\mu_1 + (\Sigma_{12}+\Sigma_{13})(\Sigma_{22}+\Sigma_{33})^{-1}(x_0 - \mu_2 - \mu_3),$$
$$\Sigma_{11} - (\Sigma_{12}+\Sigma_{13})(\Sigma_{22}+\Sigma_{33})^{-1}(\Sigma_{21}+\Sigma_{31}) \Big)$$