STAT 640: Homework 4

Due **Wednesday, February 16, 11:59pm MT** on the course Canvas webpage. Please follow the homework guidelines on the syllabus.

Name: Hannah Butler

Problem 1

Suppose that $\boldsymbol{Y} \sim N(\boldsymbol{\mu}, \sigma^2\boldsymbol{\Sigma})$ where

$$\boldsymbol{\Sigma} = (1-\rho)\boldsymbol{I} + \rho n \boldsymbol{J}_n = \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho & \cdots & \rho & 1 \end{bmatrix} \text{ where } \rho > -1/(n-1)$$

In Example 3.4, we show that if $\rho = 0$, then $\overline{Y} = \frac{1}{n}\sum_{i=1}^n Y_i$ is independent of $(Y_1 - \overline{Y}, Y_2 - \overline{Y}, \ldots, Y_n - \overline{Y})$. Prove whether or not they are independent when $\rho \neq 0$.

Answer: Consider the transformation $\boldsymbol{W} = \boldsymbol{DY}$ where

$$\boldsymbol{D} = \frac{1}{n}\begin{bmatrix} 1 & 1 & \ldots & 1 \\ n & 0 & \ldots & 0 \end{bmatrix}$$

resulting in the random vector

$$\boldsymbol{W} = \begin{bmatrix} \bar{Y} \\ Y_1 \end{bmatrix}$$

where the first component is $\bar{Y}$ and the second component is Y_1.

The covariance matrix of $\boldsymbol{W}$ is

$$\begin{aligned}
\boldsymbol{\Sigma_W} &= \sigma^2 \boldsymbol{D\Sigma D}^T \\
&= \frac{\sigma^2}{n^2}\begin{bmatrix} 1 & 1 & \ldots & 1 \\ n & 0 & \ldots & 0 \end{bmatrix}\begin{bmatrix} 1 & \rho & \ldots & \rho \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho & \ldots & \rho & 1 \end{bmatrix}\begin{bmatrix} 1 & n \\ 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \end{bmatrix} \\
&= \frac{\sigma^2}{n^2}\begin{bmatrix} 1+(n-1)\rho & 1+(n-1)\rho & \ldots & 1+(n-1)\rho \\ n & n\rho & \ldots & n\rho \end{bmatrix}\begin{bmatrix} 1 & n \\ 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \end{bmatrix} \\
&= \frac{\sigma^2}{n^2}\begin{bmatrix} n+n(n-1)\rho & n+n(n-1)\rho \\ n+n(n-1)\rho & n^2 \end{bmatrix}
\end{aligned}$$

This result can easily be extended to $\boldsymbol{W}_i = \begin{bmatrix} \bar{Y} \\ Y_i \end{bmatrix}$ for $i = 1, 2, \ldots, n$.

Now consider the transformation $\boldsymbol{U} = \boldsymbol{D}_2\boldsymbol{W}_i$ where

$$\boldsymbol{D}_2 = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

Then we have $\boldsymbol{U} = \begin{bmatrix} \bar{Y} \\ Y_i - \bar{Y} \end{bmatrix}$. Again we can compute the variance matrix for $\boldsymbol{U}$:

$$\begin{aligned}
\boldsymbol{\Sigma_U} &= \frac{\sigma^2}{n^2} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} n + n(n-1)\rho & n + n(n-1)\rho \\ n + n(n-1)\rho & n^2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \\
&= \frac{\sigma^2}{n^2} \begin{bmatrix} n + n(n-1)\rho & n + n(n-1)\rho \\ 0 & n^2 - n - n(n-1)\rho \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \\
&= \frac{\sigma^2}{n^2} \begin{bmatrix} n + n(n-1)\rho & 0 \\ 0 & n^2 - n - n(n-1)\rho \end{bmatrix}
\end{aligned}$$

The off-diagonal elements are zero, and $\boldsymbol{U} = \begin{bmatrix} \bar{Y} \\ Y_i - \bar{Y} \end{bmatrix}$ is multivariate normal, so $\bar{Y}$ is independent of Y_i.

Problem 2

Suppose $\boldsymbol{S} = \begin{bmatrix} 3 & s_2 \\ s_3 & 1 \end{bmatrix}$. For what values of s_2 and s_3 is $\boldsymbol{S}$ a valid variance matrix for a MVN random vector?

Answer: For $\boldsymbol{S}$ to be a valid variance matrix, it must be symmetric ($s_2 = s_3$) and $\boldsymbol{S}$ must be non-negative definite. Therefore, the roots of the quadratic polynomial $(3-\lambda)(1-\lambda) - s_2^2 = \lambda^2 - 4\lambda + (3 - s_2^2)$ must be non-negative. Using the quadratic formula, we have

$$0 \leq 4 \pm \sqrt{16 - 4(3 - s_2)} = 4 \pm 2\sqrt{1 + s_2^2}$$

Then $2 \leq \sqrt{1 + s_2^2}$ so s_2, s_3 must be between (or equal to) -3 and 3.

Problem 3

Let $\boldsymbol{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \sim N\left(\begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 & 2 & 0 \\ 2 & 3 & -1 \\ 0 & -1 & 4 \end{bmatrix}\right)$ and define $\boldsymbol{Y} = \begin{bmatrix} X_1 - X_2 \\ 2X_1 + X_2 - X_3 \end{bmatrix}$.

a. What are the distribution, variance, and expected value of $\boldsymbol{Y}$?

Answer: To get $\boldsymbol{Y}$, we apply the linear transformation

$$\boldsymbol{D} = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 1 & -1 \end{bmatrix}.$$

Since $\boldsymbol{Y}$ is a linear transformation of a multivariate normal random variable, $\boldsymbol{Y}$ will also be multivariate normal.

We can then find the expectation $E[\boldsymbol{Y}]$:

$$\boldsymbol{\mu_Y} = \boldsymbol{D\mu} = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

and variance

$$\begin{aligned} \boldsymbol{\Sigma_Y} = \boldsymbol{D\Sigma D}^T &= \begin{bmatrix} 1 & -1 & 0 \\ 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} 5 & 2 & 0 \\ 2 & 3 & -1 \\ 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 0 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 3 & -1 & 1 \\ 12 & 8 & -5 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 4 & 37 \end{bmatrix} \end{aligned}$$

---

b. Find $\boldsymbol{V} = \mathrm{Var}\left(\begin{bmatrix}\boldsymbol{X}\\ \boldsymbol{Y}\end{bmatrix}\right)$.

Answer: Here we again are applying a linear transformation $\boldsymbol{D}_1$ to $\boldsymbol{X}$ with

$$\boldsymbol{D}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & -1 \end{bmatrix}.$$

The variance $\boldsymbol{V}$ is

$$\begin{aligned}
\boldsymbol{V} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} 5 & 2 & 0 \\ 2 & 3 & -1 \\ 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 5 & 2 & 0 \\ 2 & 3 & -1 \\ 0 & -1 & 4 \\ 3 & -1 & 1 \\ 12 & 8 & -5 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 5 & 2 & 0 & 3 & 12 \\ 2 & 3 & -1 & -1 & 8 \\ 0 & -1 & 4 & 1 & -5 \\ 3 & -1 & 1 & 4 & 4 \\ 12 & 8 & -5 & 4 & 37 \end{bmatrix}
\end{aligned}$$

---

c. What are $\mathsf{rank}(\boldsymbol{V})$ and $\mathsf{nullity}(\boldsymbol{V})$? Can you provide a conceptual explanation for why this makes sense?

Answer: $\mathsf{rank}(\boldsymbol{V}) = 3$, and $\mathsf{nullity}(\boldsymbol{V}) = 2$. This makes conceptual sense since the transformation $\boldsymbol{D}_1$ also has $\mathsf{rank}(\boldsymbol{D}_1) = 3$.

---

d. What is the distribution of $\begin{bmatrix} X_1 \\ X_3 \end{bmatrix} \bigg| X_2 = x_2$?

Answer: First, let $\boldsymbol{Y}_1 = \begin{bmatrix} X_1 \\ X_3 \end{bmatrix}$ and $\boldsymbol{Y}_2 = [X_2]$. Then we can assign $\boldsymbol{\mu}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, $\boldsymbol{\mu}_2 = [-1]$, $\boldsymbol{\Sigma}_{12} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $\boldsymbol{\Sigma}_{11} = \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}$, $\boldsymbol{\Sigma}_{21} = \begin{bmatrix} 2 & -1 \end{bmatrix}$ and $\boldsymbol{\Sigma}_{22} = 3$.

Then we can compute the expectation $\boldsymbol{\mu}$ as

$$\begin{aligned} \boldsymbol{\mu} &= \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(x_2 - \boldsymbol{\mu}_2) \\ &= \begin{bmatrix} 3 \\ 2 \end{bmatrix} + \frac{x_2 - 1}{3} \begin{bmatrix} 2 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 3 + 2(x_2 - 1)/3 \\ 2 + (x_2 - 1)/3 \end{bmatrix} = \begin{bmatrix} (2x_2 + 7)/3 \\ (x_3 + 7)/3 \end{bmatrix} \end{aligned}$$

and the variance

$$\begin{aligned} \boldsymbol{\Sigma} &= \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} \\ &= \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 2 \\ -1 \end{bmatrix} \begin{bmatrix} 2 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 5 - 4/3 & 2/3 \\ 2/3 & 4 - 1/3 \end{bmatrix} \\ &= \begin{bmatrix} 11/3 & 2/3 \\ 2/3 & 11/3 \end{bmatrix} \end{aligned}$$

So $\begin{bmatrix} X_1 \\ X_3 \end{bmatrix} \bigg| X_2 = x_2 \sim N\left(\begin{bmatrix} (2x_2 + 7)/3 \\ (x_3 + 7)/3 \end{bmatrix}, \begin{bmatrix} 11/3 & 2/3 \\ 2/3 & 11/3 \end{bmatrix} \right)$

Problem 4

A colleague asks for your help in generating samples from the following MVN distribution:

$$\boldsymbol{Y} \sim N\left(\begin{bmatrix}5\\10\\15\end{bmatrix}, \begin{bmatrix}5 & 1 & 2\\1 & 5 & 1\\2 & 1 & 5\end{bmatrix}\right)$$

However, they are on computer whose R installation only includes the `base` and `stats` packages (note: `eigen()` is in `base` and `rnorm()` is in `stats`, but `mvrnorm()` is in neither). Do one of the following: (i) Provide a brief algorithm and example code for how your colleague can generate the desired sample, OR (ii) Explain why it is impossible with only these tools.

General Algorithm

1. Eigen-decompose $\boldsymbol{\Sigma}$
2. Set $\boldsymbol{D} = \boldsymbol{Q}\boldsymbol{\Lambda}^{1/2}$ (from eigen decomposition) and $\boldsymbol{c} = (\mu_1, \ldots, \mu_n)$
3. Simulate n-dimensional $\boldsymbol{X} \sim N(\mathbf{0}, \boldsymbol{I})$
4. Apply transformation $\boldsymbol{D}\boldsymbol{X} + \boldsymbol{c}$

Example

```
# Set desired covariance structure
Sigma = cbind(c(5, 1, 2), c(1, 5, 1), c(2, 1, 5))

# Eigen decompose Sigma
Sigma_eig = eigen(Sigma)

# Set D = QL^1/2
D = Sigma_eig$vectors %*% diag(sqrt(Sigma_eig$values))

# Set mean vector
c = c(5, 10, 15)

# simulate N(0,1) random variables
X1 <- rnorm(10000, 0, 1)
X2 <- rnorm(10000, 0, 1)
X3 <- rnorm(10000, 0, 1)
X <- rbind(X1, X2, X3)

# Apply transformation
Y = D %*% X + c

# Check sample means and variances:
colMeans(t(Y))
```

```
## [1]  4.981813  9.996707 15.026465
```

```
cov(t(Y))
```

```
##          [,1]     [,2]     [,3]
## [1,] 5.093175 0.992844 2.004124
## [2,] 0.992844 4.999959 1.030348
## [3,] 2.004124 1.030348 5.035758
```

Problem 5

Prove Proposition 3.11: Suppose

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_1 \\ \boldsymbol{Y}_2 \end{bmatrix} = N\left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \right). \tag{1}$$

If $|\boldsymbol{\Sigma}_{22}| > 0$, then the conditional distribution of $\boldsymbol{Y}_1$ given $\boldsymbol{Y}_2 = \boldsymbol{y}_2$ is MVN with mean and variance:

$$\boldsymbol{Y}_1 | \boldsymbol{Y}_2 = \boldsymbol{y}_2 \sim N\left(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{y}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right). \tag{2}$$

Do this in the following steps.

a. Let $\boldsymbol{X} = (\boldsymbol{Y}_1 - \boldsymbol{\mu}_1) - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{Y}_2 - \boldsymbol{\mu}_2)$. Find $\mathrm{E}[\boldsymbol{X}]$, $\mathrm{Var}(\boldsymbol{X})$, and $\mathrm{Cov}(\boldsymbol{X}, \boldsymbol{Y}_2)$.

Answer: To find the expectation of $\boldsymbol{X}$, we have

$$\begin{aligned} E(\boldsymbol{X}) &= E(\boldsymbol{Y}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{Y}_2 - \boldsymbol{\mu}_2)) \\ &= \boldsymbol{\mu}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{0} \\ &= \mathbf{0}. \end{aligned}$$

For the variance, we have

$$\begin{aligned} \mathrm{Var}(\boldsymbol{X}) &= E(\boldsymbol{X}^2) - E(\boldsymbol{X})^2 \\ &= E(\boldsymbol{X}^2) \\ &= E\left((\boldsymbol{Y}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{Y}_2 - \boldsymbol{\mu}_2))(\boldsymbol{Y}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{Y}_2 - \boldsymbol{\mu}_2))^T\right) \\ &= E\Big(\boldsymbol{Y}_1\boldsymbol{Y}_1^T - \boldsymbol{Y}_1\boldsymbol{\mu}_1^T - \boldsymbol{Y}_1(\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{Y}_2 - \boldsymbol{\mu}_2))^T - \boldsymbol{\mu}_1\boldsymbol{Y}_1^T + \boldsymbol{\mu}_1\boldsymbol{\mu}_1^T + \boldsymbol{\mu}_1(\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{Y}_2 - \boldsymbol{\mu}_2))^T \\ &\quad - (\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{Y}_2 - \boldsymbol{\mu}_2)\boldsymbol{Y}_1^T + (\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{Y}_2 - \boldsymbol{\mu}_2))\boldsymbol{\mu}_1 + (\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{Y}_2 - \boldsymbol{\mu}_2))(\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{Y}_2 - \boldsymbol{\mu}_2))^T\Big) \\ &= \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} - \mathbf{0} + \mathbf{0} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} + \mathbf{0} + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} \\ &= \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}. \end{aligned}$$

The covariance of $\boldsymbol{X}$ and $\boldsymbol{Y}_2$ is

$$\begin{aligned} \mathrm{Cov}(\boldsymbol{X}, \boldsymbol{Y}_2) &= E(\boldsymbol{X}\boldsymbol{Y}_2^T) - E(\boldsymbol{X})E(\boldsymbol{Y}_2)^T \\ &= E(\boldsymbol{X}\boldsymbol{Y}_2^T) \\ &= E\left((\boldsymbol{Y}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{Y}_2 - \boldsymbol{\mu}_2))\boldsymbol{Y}_2^T\right) \\ &= E(\boldsymbol{Y}_1\boldsymbol{Y}_2^T - \boldsymbol{\mu}_1\boldsymbol{Y}_2^T - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}(\boldsymbol{Y}_2 - \boldsymbol{\mu}_2)\boldsymbol{Y}_2^T) \\ &= \boldsymbol{\Sigma}_{12} - \boldsymbol{\Sigma}_{12} \\ &= \mathbf{0}. \end{aligned}$$

---

b. Provide the joint distribution of $\boldsymbol{X}$ and $\boldsymbol{Y}_2$ and explain why they are independent.

Answer: Since $\boldsymbol{X}$ is multivariate normal and $\boldsymbol{Y}_2$ is multivariate normal, the joint distribution for $\begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{Y}_2 \end{bmatrix}$ is

$$N\left(\begin{bmatrix} \mathbf{0} \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \right).$$

Since $\boldsymbol{X}$ is multivariate normal and $\boldsymbol{Y}_2$ is multivariate normal, $\mathrm{Cov}(\boldsymbol{X}, \boldsymbol{Y}_2) = \mathbf{0}$ implies that they are independent.

c. Write $\boldsymbol{Y}_1$ in terms of $\boldsymbol{X}$ and $\boldsymbol{Y}_2$.

Answer: Since $\boldsymbol{X} = \boldsymbol{Y}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{Y}_2 - \boldsymbol{\mu}_2)$, we have

$$\boldsymbol{Y}_1 = \boldsymbol{X} + \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{Y}_2 - \boldsymbol{\mu}_2).$$

d. Show that the mgf of $\boldsymbol{Y}_1|\boldsymbol{Y}_2 = \boldsymbol{y}_2$ has the desired form for the result to hold.

Answer:

$$\begin{aligned}
M_{\boldsymbol{Y}_1|\boldsymbol{Y}_2}(\boldsymbol{t}) &= E(\exp(\boldsymbol{Y}_1^T\boldsymbol{t})|\boldsymbol{Y}_2 = \boldsymbol{y}_2) \\
&= E\left(\exp(\boldsymbol{X}^T\boldsymbol{t})\exp(\boldsymbol{\mu}_1^T\boldsymbol{t})\exp((\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{Y}_2 - \boldsymbol{\mu}_2))^T\boldsymbol{t})|\boldsymbol{Y}_2 = \boldsymbol{y}_2\right) \\
&= \exp((\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{y}_2 - \boldsymbol{\mu}_2))^T\boldsymbol{t})E(\exp(\boldsymbol{X}^T\boldsymbol{t})) && (\boldsymbol{X} \textit{ and } \boldsymbol{Y}_2 \textit{ are independent}) \\
&= \exp((\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{y}_2 - \boldsymbol{\mu}_2))^T\boldsymbol{t})\exp\left(\boldsymbol{0}^T\boldsymbol{t} + \frac{1}{2}\boldsymbol{t}^T(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})\boldsymbol{t}\right) \\
&= \exp\left((\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{y}_2 - \boldsymbol{\mu}_2))^T\boldsymbol{t} + \frac{1}{2}\boldsymbol{t}^T(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})\boldsymbol{t}\right)
\end{aligned}$$

This is the MGF for a random vector with multivariate normal distribution

$$N\left(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{y}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right).$$