Multivariate Statistical Analysis

Homework 2

Lucas Fellmeth, Helen Kafka, Sven Bergmann

02/15/24

Problem 1

(a)

We have $X \in \mathbb{R}^p$, $\mathbb{E}[X] = \mu$, $\mathrm{Cov}(X) = \Sigma$, A is a $p \times p$ constant matrix and $\mathrm{tr}(Avv^\top) = v^\top Av$. Because A is a constant $p \times p$ matrix, A is symmetric. Because of this symmetry, it follows that it has a Cholesky decomposition as $A = C^\top C$.

Let $y = CX$.

Then

$$\begin{aligned}
\mathbb{E}[X^\top AX] &= \mathbb{E}[X^\top C^\top CX] \\
&= \mathbb{E}[(CX)^\top CX] \\
&= \mathbb{E}[y^\top y] \\
&= \sum_i \mathbb{E}[y_i^2] \\
&= \sum_i \mathrm{Var}(y_i) + \mathbb{E}[y_i]^2 \\
&= \mathrm{tr}(\Sigma_y) + \mu_y^\top \mu_y
\end{aligned}$$

where $\Sigma_y = \mathbb{E}[(y - \mathbb{E}[y])(y - \mathbb{E}[y])^\top) = C\Sigma C^\top$ and $\mu_y = C_\mu$.

$$\begin{aligned}
\implies \mathbb{E}[X^\top AX] &= \mathrm{tr}(C\Sigma C^\top) + \mu^\top \underbrace{C^\top C}_{=A} \mu \\
&= \mathrm{tr}(\Sigma C^\top C) + \mu^\top A\mu \\
&= \mathrm{tr}(\Sigma A) + \mu^\top A\mu
\end{aligned}$$

(b)

$X_1, \ldots, X_n$ uncorrelated:

$\mathrm{Cov}(X_i, X_j) = 0$ for $i \neq j$ and $\mathrm{Cov}(X_i, X_i) = \mathrm{Var}(X_i) = \sigma^2$

$$\implies \Sigma = \begin{pmatrix} \sigma^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma^2 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \dots & \vdots \\ \vdots & \vdots & \dots & \ddots & \vdots \\ 0 & 0 & \dots & \dots & \sigma^2 \end{pmatrix}, J = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix} \in \mathbb{R}^{p \times p}.$$

$$A = T - \frac{1}{p}J = \begin{pmatrix} 1-\frac{1}{p} & -\frac{1}{p} & -\frac{1}{p} & \dots & -\frac{1}{p} \\ -\frac{1}{p} & 1-\frac{1}{p} & -\frac{1}{p} & \dots & -\frac{1}{p} \\ -\frac{1}{p} & -\frac{1}{p} & \ddots & \dots & \vdots \\ \vdots & \vdots & \dots & \ddots & \vdots \\ -\frac{1}{p} & -\frac{1}{p} & \dots & \dots & 1-\frac{1}{p} \end{pmatrix} \in \mathbb{R}^{p \times p},$$

$$A\Sigma = \begin{pmatrix} (1-\frac{1}{p})\sigma^2 & -\frac{\sigma^2}{p} & -\frac{\sigma^2}{p} & \dots & -\frac{\sigma^2}{p} \\ -\frac{\sigma^2}{p} & (1-\frac{1}{p})\sigma^2 & -\frac{1}{p} & \dots & -\frac{1}{p} \\ -\frac{\sigma^2}{p} & -\frac{\sigma^2}{p} & \ddots & \dots & \vdots \\ \vdots & \vdots & \dots & \ddots & \vdots \\ -\frac{\sigma^2}{p} & -\frac{\sigma^2}{p} & \dots & \dots & (1-\frac{1}{p})\sigma^2 \end{pmatrix} \in \mathbb{R}^{p \times p}.$$

$$\begin{aligned}
\mathbb{E}[X^\top AX] &\overset{a)}{=} \operatorname{tr}(A\Sigma) + \mu^\top A\mu \\
&= \sum_{i=1}^{p}(1-\frac{1}{p})\sigma^2 + (\mu \dots \mu)^\top \begin{pmatrix} 1-\frac{1}{p} & -\frac{1}{p} & -\frac{1}{p} & \dots & -\frac{1}{p} \\ -\frac{1}{p} & 1-\frac{1}{p} & -\frac{1}{p} & \dots & -\frac{1}{p} \\ -\frac{1}{p} & -\frac{1}{p} & \ddots & \dots & \vdots \\ \vdots & \vdots & \dots & \ddots & \vdots \\ -\frac{1}{p} & -\frac{1}{p} & \dots & \dots & 1-\frac{1}{p} \end{pmatrix} \begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix} \\
&= p \cdot (1-\frac{1}{p}) \cdot \sigma^2 + (\underbrace{\underbrace{\mu \cdot (1-\frac{1}{p}) + (p-1) \cdot (-\frac{1}{p}) \cdot \mu}_{= \frac{p-1}{p} \cdot \mu + \frac{1-p}{p} \cdot \mu}}_{=0} \dots \mu \cdot (1-\frac{1}{p}) + (p-1) \cdot (-\frac{1}{p}) \cdot \mu) \begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix} \\
&= (p-1) \cdot \sigma^2 + (0 \dots 0) \cdot \begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix} \\
&= (p-1) \cdot \sigma^2
\end{aligned}$$

(c)

$\mathrm{Cov}(X_i, X_j) = \rho\sigma^2$, $i \neq j$, $\mathrm{Cov}(X_i, X_i) = \sigma^2$.

$$
\Sigma = \text{Cov}(X) = \begin{pmatrix} \sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \dots & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 & 0 & \dots & \rho\sigma^2 \\ \rho\sigma^2 & 0 & \ddots & \dots & \vdots \\ \vdots & \vdots & \dots & \ddots & \rho\sigma^2 \\ \rho\sigma^2 & \rho\sigma^2 & \dots & \rho\sigma^2 & \sigma^2 \end{pmatrix}
$$

$$
A = T - \frac{1}{p}J = \begin{pmatrix} 1-\frac{1}{p} & -\frac{1}{p} & -\frac{1}{p} & \dots & -\frac{1}{p} \\ -\frac{1}{p} & 1-\frac{1}{p} & -\frac{1}{p} & \dots & -\frac{1}{p} \\ -\frac{1}{p} & -\frac{1}{p} & \ddots & \dots & \vdots \\ \vdots & \vdots & \dots & \ddots & \vdots \\ -\frac{1}{p} & -\frac{1}{p} & \dots & \dots & 1-\frac{1}{p} \end{pmatrix} \in \mathbb{R}^{p\times p}.
$$

$$
\begin{aligned}
\text{tr}(A\Sigma) &= \sum_{i=1}^{p}(1-\frac{1}{p})\cdot\sigma^2 + (p-1)\cdot(-\frac{1}{p})\cdot\rho\sigma^2 \\
&= p\cdot(1-\frac{1}{p})\cdot\sigma^2 + (p-1)\cdot(-\frac{1}{p})\cdot\rho\sigma^2 \\
&= (p-1)\cdot\sigma^2 + (1-p)\rho\cdot\sigma^2 \\
&= (p-1)\cdot\sigma^2 + (\rho - p\cdot\rho)\cdot\sigma^2 \\
&= (p-1+\rho-p\cdot\rho)\cdot\sigma^2 \\
\mathbb{E}[X^\top AX] &= \text{tr}(A\Sigma) + \underbrace{\mu^\top A\mu}_{=0 \text{ (as shown above)}} \\
&= (p-1+\rho-p\cdot\rho)\cdot\sigma^2
\end{aligned}
$$

Problem 2

(a)

$X \sim \mathcal{N}(\mu, \Sigma)$. We have to prove that $Z = \Sigma^{-\frac{1}{2}}(X-\mu) \sim \mathcal{N}(0, I)$.

$$
\begin{aligned}
\mathbb{E}[Z] &= \mathbb{E}[\Sigma^{-\frac{1}{2}}(X-\mu)] = \Sigma^{-\frac{1}{2}}(\mathbb{E}[X]-\mu) = \Sigma^{-\frac{1}{2}}(\mu-\mu) = 0, \\
\text{Var}(Z) &= \text{Var}(\Sigma^{-\frac{1}{2}}(X-\mu)) = (\Sigma^{-\frac{1}{2}})^2\text{Var}(X-\mu) \\
&= \Sigma^{-1}\cdot\text{Var}(X) = \Sigma^{-1}\cdot\Sigma = I. \\
&\implies Z \sim \mathcal{N}(\mu, \Sigma).
\end{aligned}
$$

(b)

$$
\begin{aligned}
Z &= \Sigma^{-\frac{1}{2}}(X-\mu) \\
&= \Sigma^{-\frac{1}{2}}(X-0) \\
&= \Sigma^{-\frac{1}{2}}\cdot X.
\end{aligned}
$$

```r
Sigma <- matrix(data = c(1, -2, 0, -2, 5, 0, 0, 0, 2), nrow = 3, ncol = 3)
print(Sigma)
```

```
##      [,1] [,2] [,3]
## [1,]    1   -2    0
## [2,]   -2    5    0
## [3,]    0    0    2
```

Calculate the matrix square root:

```r
Sigma_sqrt <- expm::sqrtm(Sigma)
print(Sigma_sqrt)
```

```
##            [,1]       [,2]     [,3]
## [1,]  0.7071068 -0.7071068 0.000000
## [2,] -0.7071068  2.1213203 0.000000
## [3,]  0.0000000  0.0000000 1.414214
```

Check if the matrix square root times the matrix square root equals A:

```r
print(Sigma_sqrt %*% Sigma_sqrt)
```

```
##      [,1] [,2] [,3]
## [1,]    1   -2    0
## [2,]   -2    5    0
## [3,]    0    0    2
```

So it follows:

$$Z = \begin{pmatrix} 0.7071068 \cdot x_1 + (-0.7071068) \cdot x_2 \\ (-0.7071068) \cdot x_1 + 2.1213203 \cdot x_2 \\ 1.414214 \cdot x_3 \end{pmatrix}$$

Problem 3

$X = (X_1, X_2)$ with joint pdf:

$$f/(x_1, x_2) = \begin{cases} 2\varphi(X_1)\varphi(X_2), & X_1 \cdot X_2 > 0, \\ 0, & \text{otherwise} \end{cases}$$

where

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} exp(\frac{-x}{2})$$

$$\implies f(x_1, x_2) = 2 \cdot \frac{1}{\sqrt{2\pi}} exp(\frac{-x_1}{2}) \cdot \frac{1}{\sqrt{2\pi}} exp(\frac{-x_2}{2}) \cdot I_{\{x_1 \cdot x_2 > 0\}}(x_1, x_2)$$

Case 1 (x_1, x_2 > 0):

$$
\begin{aligned}
f_{X_1}(x_1) &= \int_0^\infty 2 \cdot \frac{1}{\sqrt{2\pi}} exp(\frac{-x_1}{2}) \cdot \frac{1}{\sqrt{2\pi}} exp(\frac{-x_2}{2}) dx_2 \\
&= 2 \cdot \frac{1}{\sqrt{2\pi}} exp(\frac{-x_1}{2}) \int_0^\infty \frac{1}{\sqrt{2\pi}} exp(\frac{-x_2}{2}) dx_2 \\
&= 2 \cdot \frac{1}{\sqrt{2\pi}} exp(\frac{-x_1}{2}) \cdot \frac{1}{2} \cdot \underbrace{\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} exp(\frac{-x_2}{2}) dx_2}_{\text{=1 (pdf of } \mathcal{N}(0,1) \text{ over it's whole support)}} \\
&= \frac{1}{\sqrt{2\pi}} exp(\frac{-x_1}{2}) \cdot I_{\{0,\infty\}}(x_1). \\
f_{X_2}(x_2) &= \int_0^\infty 2 \cdot \frac{1}{\sqrt{2\pi}} exp(\frac{-x_2}{2}) \cdot \frac{1}{\sqrt{2\pi}} exp(\frac{-x_1}{2}) dx_1 \\
&= 2 \cdot \frac{1}{\sqrt{2\pi}} exp(\frac{-x_2}{2}) \int_0^\infty \frac{1}{\sqrt{2\pi}} exp(\frac{-x_1}{2}) dx_1 \\
&= 2 \cdot \frac{1}{\sqrt{2\pi}} exp(\frac{-x_2}{2}) \cdot \frac{1}{2} \cdot \underbrace{\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} exp(\frac{-x_1}{2}) dx_1}_{\text{=1(pdf of } \mathcal{N}(0,1) \text{ over it's whole support)}} \\
&= \frac{1}{\sqrt{2\pi}} exp(\frac{-x_2}{2}) \cdot I_{\{0,\infty\}}(x_2).
\end{aligned}
$$

Case 2 (x_1, x_2 < 0):

$$
\begin{aligned}
f_{X_1}(x_1) &= \int_0^\infty 2 \cdot \frac{1}{\sqrt{2\pi}} exp(\frac{-x_1}{2}) \cdot \frac{1}{\sqrt{2\pi}} exp(\frac{-x_2}{2}) dx_2 \\
&= 2 \cdot \frac{1}{\sqrt{2\pi}} exp(\frac{-x_1}{2}) \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}} exp(\frac{-x_2}{2}) dx_2 \\
&= 2 \cdot \frac{1}{\sqrt{2\pi}} exp(\frac{-x_1}{2}) \cdot \frac{1}{2} \cdot \underbrace{\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} exp(\frac{-x_2}{2}) dx_2}_{\text{=1 (pdf of } \mathcal{N}(0,1) \text{ over it's whole support)}} \\
&= \frac{1}{\sqrt{2\pi}} exp(\frac{-x_1}{2}) \cdot I_{\{-\infty,0\}}(x_1). \\
f_{X_2}(x_2) &= \int_0^\infty 2 \cdot \frac{1}{\sqrt{2\pi}} exp(\frac{-x_2}{2}) \cdot \frac{1}{\sqrt{2\pi}} exp(\frac{-x_1}{2}) dx_1 \\
&= 2 \cdot \frac{1}{\sqrt{2\pi}} exp(\frac{-x_2}{2}) \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}} exp(\frac{-x_1}{2}) dx_1 \\
&= 2 \cdot \frac{1}{\sqrt{2\pi}} exp(\frac{-x_2}{2}) \cdot \frac{1}{2} \cdot \underbrace{\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} exp(\frac{-x_1}{2}) dx_1}_{\text{=1(pdf of } \mathcal{N}(0,1) \text{ over it's whole support)}} \\
&= \frac{1}{\sqrt{2\pi}} exp(\frac{-x_2}{2}) \cdot I_{\{-\infty,0\}}(x_2).
\end{aligned}
$$

Then

$$
\begin{aligned}
f_{X_1}(x_1) &= \frac{1}{\sqrt{2\pi}} exp(\frac{-x_1}{2}) \cdot I_{\{-\infty,\infty\}}(x_1), \\
f_{X_2}(x_2) &= \frac{1}{\sqrt{2\pi}} exp(\frac{-x_2}{2}) \cdot I_{\{-\infty,\infty\}}(x_2).
\end{aligned}
$$

Thus, $f_{X_1}(x_1) \sim \mathcal{N}(0,1)$ and $f_{X_2}(x_2) \sim \mathcal{N}(0,1)$, but X is not multivariate Normal since the support of $f(X_1, X_2)$ is not $\mathbb{R}^2$.