Exercise 1. Let $X_1, \cdots, X_n$ be a sequence of i.i.d. random variables with common distribution $\mathcal{N}_k(\mu, \Sigma)$.

1. Write $\bar{\mathrm{X}}_\mathrm{n} := \frac{1}{n}\sum_{i=1}^n X_i$ and $A := \sum_{k=1}^n (X_i - \bar{\mathrm{X}}_\mathrm{n})(X_i - \bar{\mathrm{X}}_\mathrm{n})^\intercal$. Show that

$$\sum_{i=1}^n (X_i - \mu)(X_i - \mu)^\intercal = A + n(\bar{\mathrm{X}}_\mathrm{n} - \mu)(\bar{\mathrm{X}}_\mathrm{n} - \mu)^\intercal \tag{1}$$

2. Define $Y_i := \sum_{l=1}^n C_{il} X_l$, with $C \in \mathbb{R}^{n\times n}$ being an orthogonal matrix. Show that

$$\sum_{i=1}^n X_i {X_i}^\intercal = \sum_{i=1}^n Y_i {Y_i}^\intercal \tag{2}$$

Solution. 1.1
We rewrite the LHS of equation (1) and we will reach the RHS at the end of the solution, so we will prove that LHS = RHS.

$$\begin{aligned}
LHS &= \sum_{i=1}^n (X_i - \mu)(X_i - \mu)^\intercal \\
&= \sum_{i=1}^n (X_i {X_i}^\intercal - X_i\mu^\intercal - \mu {X_i}^\intercal + \mu\mu^\intercal) \\
&= \sum_{i=1}^n X_i {X_i}^\intercal - \sum_{i=1}^n X_i\mu^\intercal - \sum_{i=1}^n \mu {X_i}^\intercal + \sum_{i=1}^n \mu\mu^\intercal \\
&= \sum_{i=1}^n X_i {X_i}^\intercal + n(-\frac{1}{n}\sum_{i=1}^n X_i\mu^\intercal - \frac{1}{n}\sum_{i=1}^n \mu {X_i}^\intercal) + \underbrace{\sum_{i=1}^n \mu\mu^\intercal}_{n\cdot\mu\mu^\intercal} \\
&= \sum_{i=1}^n X_i {X_i}^\intercal + n\big(-\bar{\mathrm{X}}_\mathrm{n}\,\mu^\intercal - \mu\,\bar{\mathrm{X}}_\mathrm{n} + \mu\mu^\intercal\big)
\end{aligned}$$

Adding the term $\sum_{i=1}^n (2\bar{\mathrm{X}}_\mathrm{n}{\bar{\mathrm{X}}_\mathrm{n}}^\intercal - X_i{\bar{\mathrm{X}}_\mathrm{n}}^\intercal - \bar{\mathrm{X}}_\mathrm{n}X_i^\intercal)$ to the last equation keeps this equation intact, since this term equates zero as shown in the following:

$$\begin{aligned}
\sum_{i=1}^n (2\bar{\mathrm{X}}_\mathrm{n}{\bar{\mathrm{X}}_\mathrm{n}}^\intercal - X_i{\bar{\mathrm{X}}_\mathrm{n}}^\intercal - \bar{\mathrm{X}}_\mathrm{n}X_i^\intercal) &= \sum_{i=1}^n (2\bar{\mathrm{X}}_\mathrm{n}{\bar{\mathrm{X}}_\mathrm{n}}^\intercal) - (\sum_{i=1}^n X_i){\bar{\mathrm{X}}_\mathrm{n}}^\intercal - \bar{\mathrm{X}}_\mathrm{n}(\sum_{i=1}^n X_i^\intercal) \\
&= 2n\bar{\mathrm{X}}_\mathrm{n}{\bar{\mathrm{X}}_\mathrm{n}}^\intercal - n\,\bar{\mathrm{X}}_\mathrm{n}\,{\bar{\mathrm{X}}_\mathrm{n}}^\intercal - \bar{\mathrm{X}}_\mathrm{n}\cdot n{\bar{\mathrm{X}}_\mathrm{n}}^\intercal \\
&= 2n\bar{\mathrm{X}}_\mathrm{n}{\bar{\mathrm{X}}_\mathrm{n}}^\intercal - n\,\bar{\mathrm{X}}_\mathrm{n}\,{\bar{\mathrm{X}}_\mathrm{n}}^\intercal - n\,\bar{\mathrm{X}}_\mathrm{n}\,{\bar{\mathrm{X}}_\mathrm{n}}^\intercal \\
&= 0
\end{aligned}$$

Adding this term to the last equation of LHS yields:

$$
\begin{aligned}
LHS &= \sum_{i=1}^{n}(X_i X_i^{\intercal}) + n\big(-\bar{\mathrm{X}}_{\mathrm{n}}\,\mu^{\intercal} - \mu\,\bar{\mathrm{X}}_{\mathrm{n}} + \mu\mu^{\intercal}\big) + \sum_{i=1}^{n}(2\bar{\mathrm{X}}_{\mathrm{n}}\bar{\mathrm{X}}_{\mathrm{n}}^{\intercal} - X_i\bar{\mathrm{X}}_{\mathrm{n}}^{\intercal} - \bar{\mathrm{X}}_{\mathrm{n}}X_i^{\intercal}) \\
&= \sum_{i=1}^{n}(X_i X_i^{\intercal}) + \sum_{i=1}^{n}(-X_i\bar{\mathrm{X}}_{\mathrm{n}}^{\intercal} - \bar{\mathrm{X}}_{\mathrm{n}}X_i^{\intercal} + 2\bar{\mathrm{X}}_{\mathrm{n}}\bar{\mathrm{X}}_{\mathrm{n}}^{\intercal}) + n\big(-\bar{\mathrm{X}}_{\mathrm{n}}\,\mu^{\intercal} - \mu\,\bar{\mathrm{X}}_{\mathrm{n}} + \mu\mu^{\intercal}\big) \\
&= \sum_{i=1}^{n}(X_i X_i^{\intercal} - X_i\bar{\mathrm{X}}_{\mathrm{n}}^{\intercal} - \bar{\mathrm{X}}_{\mathrm{n}}X_i^{\intercal}) + \sum_{i=1}^{n} 2\bar{\mathrm{X}}_{\mathrm{n}}\bar{\mathrm{X}}_{\mathrm{n}}^{\intercal} + n\big(-\bar{\mathrm{X}}_{\mathrm{n}}\,\mu^{\intercal} - \mu\,\bar{\mathrm{X}}_{\mathrm{n}} + \mu\mu^{\intercal}\big) \\
&= \sum_{i=1}^{n}(X_i X_i^{\intercal} - X_i\bar{\mathrm{X}}_{\mathrm{n}}^{\intercal} - \bar{\mathrm{X}}_{\mathrm{n}}X_i^{\intercal}) + \sum_{i=1}^{n} \bar{\mathrm{X}}_{\mathrm{n}}\bar{\mathrm{X}}_{\mathrm{n}}^{\intercal} + n\bar{\mathrm{X}}_{\mathrm{n}}\bar{\mathrm{X}}_{\mathrm{n}}^{\intercal} + n\big(-\bar{\mathrm{X}}_{\mathrm{n}}\,\mu^{\intercal} - \mu\,\bar{\mathrm{X}}_{\mathrm{n}} + \mu\mu^{\intercal}\big) \\
&= \sum_{i=1}^{n}(X_i X_i^{\intercal} - X_i\bar{\mathrm{X}}_{\mathrm{n}}^{\intercal} - \bar{\mathrm{X}}_{\mathrm{n}}X_i^{\intercal} + \bar{\mathrm{X}}_{\mathrm{n}}\bar{\mathrm{X}}_{\mathrm{n}}^{\intercal}) + n\bar{\mathrm{X}}_{\mathrm{n}}\bar{\mathrm{X}}_{\mathrm{n}}^{\intercal} + n\big(-\bar{\mathrm{X}}_{\mathrm{n}}\,\mu^{\intercal} - \mu\,\bar{\mathrm{X}}_{\mathrm{n}} + \mu\mu^{\intercal}\big) \\
&= \sum_{i=1}^{n}\underbrace{(X_i X_i^{\intercal} - X_i\bar{\mathrm{X}}_{\mathrm{n}}^{\intercal} - \bar{\mathrm{X}}_{\mathrm{n}}X_i^{\intercal} + \bar{\mathrm{X}}_{\mathrm{n}}\bar{\mathrm{X}}_{\mathrm{n}}^{\intercal})}_{(X_i - \bar{\mathrm{X}}_{\mathrm{n}})(X_i - \bar{\mathrm{X}}_{\mathrm{n}})^{\intercal}} + n\big(\underbrace{\bar{\mathrm{X}}_{\mathrm{n}}\bar{\mathrm{X}}_{\mathrm{n}}^{\intercal} - \bar{\mathrm{X}}_{\mathrm{n}}\,\mu^{\intercal} - \mu\,\bar{\mathrm{X}}_{\mathrm{n}} + \mu\mu^{\intercal}}_{(\bar{\mathrm{X}}_{\mathrm{n}} - \mu)(\bar{\mathrm{X}}_{\mathrm{n}} - \mu)^{\intercal}}\big) \\
&= \sum_{k=1}^{n}(X_i - \bar{\mathrm{X}}_{\mathrm{n}})(X_i - \bar{\mathrm{X}}_{\mathrm{n}})^{\intercal} + n(\bar{\mathrm{X}}_{\mathrm{n}} - \mu)(\bar{\mathrm{X}}_{\mathrm{n}} - \mu)^{\intercal} \\
&= RHS
\end{aligned}
$$

Solution. 1.2

We rewrite the RHS of equation (2) and we will reach the LHS at the end of the solution, so we will prove that LHS = RHS.

$$
\begin{aligned}
RHS &= \sum_{i=1}^{n} Y_i Y_i^{\intercal} \\
&= \sum_{i=1}^{n}\big(\sum_{l=1}^{n} C_{il}X_l\big)\big(\sum_{l=1}^{n} C_{il}X_l\big)^{\intercal} && (Y_i = \sum_{l=1}^{n} C_{il}X_l) \\
&= \sum_{i=1}^{n}\sum_{l=1}^{n}\sum_{k=1}^{n}(C_{il}X_l)(C_{ik}X_k)^{\intercal} \\
&= \sum_{i=1}^{n}\sum_{l=1}^{n}\sum_{k=1}^{n} C_{il}X_l X_k^{\intercal} C_{ik}^{\intercal} \\
&= \sum_{i=1}^{n}\sum_{l=1}^{n}\sum_{k=1}^{n} C_{il}X_l X_k^{\intercal} C_{ik}^{\intercal} \\
&= \sum_{i=1}^{n}\sum_{l=1}^{n}\sum_{k=1}^{n} C_{il}C_{ik}^{\intercal} X_l X_k^{\intercal} && (C_{ij} \text{ and } C_{ik}^{\intercal} \text{ are scalars}) \\
&= \sum_{i=1}^{n}\sum_{l=1}^{n}\sum_{k=1}^{n} C_{il}C_{ik} X_l X_k^{\intercal} && (C_{ik}^{\intercal} = C_{ik})
\end{aligned}
$$

Now we note that since we aim at reaching $CC^\intercal$ to simplify the terms we have, we denote entries of C by C_{ij} and entries of $C^\intercal$ by $\bar{C}_{ij}$. Therefore, $C_{ij} = \bar{C}_{ji}$. Using this notation, we substitute C_{ik} with $\bar{C}_{ki}$. Consequently,

$$\sum_{i=1}^{n}\sum_{l=1}^{n}\sum_{k=1}^{n} C_{il}C_{ik} = \sum_{i=1}^{n}\sum_{l=1}^{n}\sum_{k=1}^{n} C_{il}\bar{C}_{ki}$$

When $l = k$, the term $\sum_{l=1}^{n} C_{il}\bar{C}_{ki}$ indicates the ith row and ith column of the matrix $CC^\intercal = I_n$ (as C is an orthogonal matrix). if $l \neq k$, the sum would be zero. We conclude that

$$\sum_{l=1}^{n}\sum_{k=1}^{n} C_{il}\bar{C}_{ki} = \sum_{l=1}^{n} C_{il}\bar{C}_{li} = 1 \tag{3}$$

Following a similar reasoning, we also have

$$\sum_{l=1}^{n}\sum_{i=1}^{n} \bar{C}_{li}C_{il} = \sum_{l=1}^{n} 1 \tag{4}$$

Back to the RHS, we have the following

$$\begin{aligned}
RHS &= \sum_{i=1}^{n} Y_i Y_i^\intercal \\
&= \sum_{i=1}^{n}\sum_{l=1}^{n}\sum_{k=1}^{n} C_{il}C_{ik}X_l X_k^\intercal \\
&= \sum_{i=1}^{n}\sum_{l=1}^{n} C_{il}\bar{C}_{li}X_l X_l^\intercal && \text{(logic for (3))} \\
&= \sum_{i=1}^{n}\sum_{l=1}^{n} \bar{C}_{li}C_{il}X_l X_l^\intercal \\
&= \sum_{l=1}^{n}\sum_{i=1}^{n} \bar{C}_{li}C_{il}X_l X_l^\intercal \\
&= \sum_{l=1}^{n}\big(\sum_{i=1}^{n} \bar{C}_{li}C_{il}\big)X_l X_l^\intercal \\
&= \sum_{l=1}^{n} 1 X_l X_l^\intercal && \text{(4)} \\
&= \sum_{l=1}^{n} X_l X_l^\intercal \\
&= LHS
\end{aligned}$$