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.

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

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.

$$
\begin{aligned}
RHS &= \sum_{i=1}^{n} Y_i {Y_i}^\intercal\\
&= \sum_{i=1}^{n}\big(\sum_{l=1}^{n} c_{il}X_l\big)\big(\sum_{l=1}^{n} c_{il}X_l\big)^\intercal && (Y_i = \sum_{l=1}^{n} c_{il}X_l)\\
&= \sum_{i=1}^{n}\sum_{l=1}^{n}\sum_{k=1}^{n}(c_{il}X_l)(c_{ik}X_k)^\intercal\\
&= \sum_{i=1}^{n}\sum_{l=1}^{n}\sum_{k=1}^{n}(c_{il}X_l)(c_{ik}X_k)^\intercal\\
&= \sum_{i=1}^{n}\sum_{l=1}^{n}\sum_{k=1}^{n} c_{il}c_{ik}\langle X_l, X_k\rangle && \text{norm property: } \langle \alpha X, \beta Y\rangle = \alpha\beta\langle X, Y\rangle
\end{aligned}
$$

As C is an orthogonal matrix, $C^\intercal C = I = CC^\intercal$, we have

$$\forall i, j, k \in \{1, \cdots, \}, j \neq k, \langle c_{ij}, c_{ik}\rangle = 0$$

$$\forall i, j, k \in \{1, \cdots, \}, j = k, \langle c_{ij}, c_{ik}\rangle = 1,$$

where c_{ij} denotes the entry placed at the ith row and jth columns of C.
It follows that the quantity $\langle c_{il}X_l, c_{ik}X_k\rangle$ is zero unless $il = ik \implies l = k$. Therefore,

$$\begin{aligned}
RHS &= \sum_{i=1}^{n}\sum_{l=1}^{n}\sum_{k=l} c_{il}c_{ik} \cdot \langle X_l, X_k\rangle \\
&= \sum_{i=1}^{n}\sum_{l=1}^{n} \underbrace{c_{il}^2}_{1} \langle X_l, X_l\rangle \\
&= \sum_{l=1}^{n}\sum_{i=1}^{n} \langle X_l, X_l\rangle && \text{(change poistion of sums)} \\
&= \sum_{l=1}^{n} \langle X_l, X_l\rangle && \text{(the term under sum doesn't depend on i anymore)} \\
&= \sum_{l=1}^{n} X_l {X_l}^\intercal \\
&= \sum_{i=1}^{n} X_i {X_i}^\intercal && \text{(l is dummy variable, hence we can replace it with i)} \\
&= LHS
\end{aligned}$$