

Let  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$  be a random sample of size  $n$  from a  $p$ -variate normal distribution with mean vector  $\boldsymbol{\mu}$  and covariance matrix  $\boldsymbol{\Sigma}$ . What's the distribution of  $(n-1)\mathbf{S}$ ?

$$\begin{aligned}
(n-1)\mathbf{S} &= \frac{n-1}{n-1} \sum_{j=1}^n (\mathbf{X}_j - \bar{\mathbf{X}}) (\mathbf{X}_j - \bar{\mathbf{X}})' = \\
&= \sum_{j=1}^n (\mathbf{X}_j + \boldsymbol{\mu} - \boldsymbol{\mu} - \bar{\mathbf{X}}) (\mathbf{X}_j + \boldsymbol{\mu} - \boldsymbol{\mu} - \bar{\mathbf{X}})' = \\
&= \sum_{j=1}^n ((\mathbf{X}_j - \boldsymbol{\mu}) - (\bar{\mathbf{X}} - \boldsymbol{\mu})) ((\mathbf{X}_j - \boldsymbol{\mu}) - (\bar{\mathbf{X}} - \boldsymbol{\mu}))' = \\
&= \sum_{j=1}^n (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' - (\bar{\mathbf{X}} - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' - (\mathbf{X}_j - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' + (\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' = \\
&= \sum_{j=1}^n (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' - 2 \sum_{j=1}^n (\mathbf{X}_j - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' + \sum_{j=1}^n (\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' = \\
&= \sum_{j=1}^n (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' - 2(\bar{\mathbf{X}} - \boldsymbol{\mu}) \sum_{j=1}^n (\mathbf{X}_j - \boldsymbol{\mu})' + n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' = \\
&= \sum_{j=1}^n (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' - 2(\bar{\mathbf{X}} - \boldsymbol{\mu})((\sum_{j=1}^n \mathbf{X}_j) - n\boldsymbol{\mu})' + n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' = \\
&= \sum_{j=1}^n (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' - 2(\bar{\mathbf{X}} - \boldsymbol{\mu})(n\bar{\mathbf{X}} - n\boldsymbol{\mu})' + n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' = \\
&= \sum_{j=1}^n (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' - 2n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' + n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' = \\
&= \sum_{j=1}^n (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' - n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' = \\
&= \sum_{j=1}^n (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' - [\sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu})] [\sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu})]' =
\end{aligned}$$

so looking at the first and second part of this equation

$$\begin{aligned}
\mathbf{X}_j &\sim N\left(\begin{smallmatrix} \boldsymbol{\mu} \\ (p \times 1) \end{smallmatrix}, \begin{smallmatrix} \boldsymbol{\Sigma} \\ (p \times p) \end{smallmatrix}\right) \Rightarrow (\mathbf{X}_j - \boldsymbol{\mu}) \sim N(\mathbf{0}, \boldsymbol{\Sigma}) \Rightarrow \\
&\Rightarrow (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' \sim W_p(1, \boldsymbol{\Sigma}) \Rightarrow \\
&\Rightarrow \sum_{j=1}^n (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' \sim W_p(n, \boldsymbol{\Sigma})
\end{aligned}$$

That is, the first term in the sum is distributed as a Wishart distribution with  $n$  degrees of freedom. Reminder that a  $p$ -dimensional Wishart distribution with  $m$  degrees of freedom,  $W_p(m, \boldsymbol{\Sigma})$ , is equal to the distribution of  $\sum_{j=1}^m \mathbf{Z}_j \mathbf{Z}_j'$ , where the  $\mathbf{Z}_j$  are independently distributed as  $N_p(\mathbf{0}, \boldsymbol{\Sigma})$ . For the second term,

$$\begin{aligned}
\bar{\mathbf{X}} &\sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}/n) \Rightarrow (\bar{\mathbf{X}} - \boldsymbol{\mu}) \sim N(\mathbf{0}, \boldsymbol{\Sigma}/n) \Rightarrow \\
&\Rightarrow \sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \sim N(\mathbf{0}, \boldsymbol{\Sigma}) \Rightarrow
\end{aligned}$$

$$\left[\sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu})\right] \left[\sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu})'\right] \sim W_p(1, \boldsymbol{\Sigma})$$

That is, the second term in the sum is distributed as a Wishart distribution with 1 degree of freedom.

Combining the two, we now have a Wishart distribution with  $n$  degrees of freedom minus another Wishart distribution with 1 degree of freedom to create a Wishart distribution with  $n - 1$  distribution, so we finally have

$$(n - 1)\mathbf{S} \sim W_p\left(n - 1, \boldsymbol{\Sigma}_{(p \times p)}\right)$$