Chapter 3
Wishart Distribution

Suppose the p-component vectors $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_N$ $(N > p)$ are independent and each distributed according to $N_p(\underset{\sim}{\mu}, \Sigma)$. The maximum likelihood estimators of $\underset{\sim}{\mu}$ and Σ are

$$\hat{\underset{\sim}{\mu}} = \bar{\underset{\sim}{X}} \quad \text{and} \quad \hat{\Sigma} = \frac{1}{N}\sum_{\alpha=1}^{N}(\underset{\sim}{X}_\alpha - \bar{\underset{\sim}{X}})(\underset{\sim}{X}_\alpha - \bar{\underset{\sim}{X}})' = \frac{A}{N} = S,$$

respectively, where $A = \sum_{\alpha=1}^{N}(\underset{\sim}{X}_\alpha - \bar{\underset{\sim}{X}})(\underset{\sim}{X}_\alpha - \bar{\underset{\sim}{X}})'$ and $n = N - 1$.

The distribution of A (or S) is the **Wishart distribution**.

Note that

$$S = \begin{pmatrix} s_{11} & s_{12} & \ldots & s_{1p} \\ s_{21} & s_{22} & \ldots & s_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{p1} & s_{p2} & \ldots & s_{pp} \end{pmatrix}$$

has $1 + 2 + \ldots + p = \frac{p(p+1)}{2}$ distinct elements. The Wishart distribution is the joint distribution of the elements of S.

Suppose $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_N$ are independent p-component vectors and each distributed $N_p(\underset{\sim}{\mu}, \Sigma)$. Then the density of positive definite $A = \sum_{\alpha=1}^{N} \underset{\sim}{Z}_\alpha \underset{\sim}{Z}'_\alpha$ where $\underset{\sim}{Z}_\alpha = (\underset{\sim}{X}_\alpha - \bar{\underset{\sim}{X}})$ is

$$f(A) = \frac{1}{2^{\frac{np}{2}} \pi^{\frac{p(p-1)}{4}} |\Sigma|^{\frac{n}{2}} \prod_{i=1}^{p} \Gamma(\frac{n+1-i}{2})} |A|^{\frac{n-p-1}{2}} \exp\left(-\frac{1}{2} tr(\Sigma^{-1}A)\right)$$

and $f(A) = 0$ otherwise; where $n = N - 1$. The associated distribution is denoted as $W_p(\Sigma, n)$. The parameters of the Wishart distribution are Σ and n.

Theorem 1: If $\underset{\sim}{Y}_\alpha \sim N_p(\underset{\sim}{0}, \Sigma)$, where $\underset{\sim}{Y}_\alpha$ is a p-component vector with $n \geq p$ and $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_n$ are independently distributed, then the density of $A = \sum_{\alpha=1}^{n} \underset{\sim}{Y}_\alpha \underset{\sim}{Y}'_\alpha$ is $W_p(\Sigma, n)$ for a positive definite matrix A and 0 otherwise.

The following theorem is used to derive the characteristic function of the Wishart distribution. It ensures the existence of a reduction used there.

Theorem 2: If A and B are two real symmetric matrices such that A is positive semi definite and B is positive definite, then there exists a non-singular matrix P such that

$$\begin{aligned} P'AP &= diag(\lambda_1, \lambda_2, \ldots, \lambda_n) \\ P'BP &= I_n \end{aligned}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the characteristic roots of the equation $|A - \lambda B| = 0$ in λ.

Proof: Since B is a positive definite matrix, there exists a non-singular matrix Q such that $Q'BQ = I$.

Consider $Q'AQ$, which is real and symmetric. By the spectral decomposition theorem, there exists an orthogonal matrix H such that

$$H'(Q'AQ)H = D = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

where Rank $(Q'AQ) =$ Rank (D) and $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the characteristic roots of $|Q'AQ - \lambda I| = 0$.

Let $P = QH$, P is non-singular, so this implies $P'AP = D = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Since Q is non-singular, P is also non-singular.

Since $Q'BQ = I$ and $H'H = I$,

$$\begin{aligned} & Q'BQ = I \\ \text{or} \quad & H'QBQH = H'H \\ \text{or} \quad & P'BP = I. \end{aligned}$$

Now consider

$$\begin{aligned} & |Q'AQ - \lambda I| = 0 \\ \text{or} \quad & |Q'AQ - \lambda Q'BQ| = 0 \\ \text{or} \quad & |Q'(A - \lambda B)Q| = 0 \\ \text{or} \quad & |Q'||A - \lambda B||Q| = 0 \\ \text{or} \quad & |QQ'||A - \lambda B| = 0. \\ \text{or} \quad & |A - \lambda B| = 0. \end{aligned}$$

Characteristic Function

Suppose $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_n$ $(n \geq p)$ are independently distributed and $\underset{\sim}{Y}_\alpha \sim N(\underset{\sim}{0}, \Sigma)$, $\alpha = 1, 2, \ldots, n$.

Consider $A = \sum_{\alpha=1}^{n} \underset{\sim}{Y}_\alpha \underset{\sim}{Y}'_\alpha$.

Introduce the $p \times p$ matrix $\Theta = ((\theta_{ij}))$ with $\theta_{ij} = \theta_{ji}$. The characteristic function is

$$\begin{aligned}
\phi_A(\Theta) &= E[\exp(itr(A\Theta))] \\
&= E[\exp(itr(\sum_{\alpha=1}^{n} \underset{\sim}{Y}_\alpha \underset{\sim}{Y}'_\alpha \Theta))] \\
&= E[\exp(i\sum_{\alpha=1}^{n} \underset{\sim}{Y}'_\alpha \Theta \underset{\sim}{Y}_\alpha)] \\
&= \prod_{\alpha=1}^{n} E[\exp(i\underset{\sim}{Y}'_\alpha \Theta \underset{\sim}{Y}_\alpha)] \\
&= [E\{\exp(i\underset{\sim}{Y}'\Theta\underset{\sim}{Y})\}]^n.
\end{aligned}$$

First we find $E[\exp(i\underset{\sim}{Y}'\Theta\underset{\sim}{Y})]$. Using Theorem 1, we use the following result.
Consider Σ^{-1} and for real Θ, there exists a real non-singular matrix P such that

$$\begin{aligned}
&P'\Sigma^{-1}P = I \\
&P'\Theta P = diag(\lambda_1, \lambda_2, \ldots, \lambda_p) = D
\end{aligned}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_p$ are the characteristic roots of $|\Theta - \lambda\Sigma^{-1}| = 0$.

Consider a linear transformation $\underset{\sim}{Y} = P\underset{\sim}{Z}$ where $\underset{\sim}{Z} \sim N(\underset{\sim}{0}, I)$ and Z_j's independently distributed. Then

$$\begin{aligned}
E[\exp(i\underset{\sim}{Y}'\Theta\underset{\sim}{Y})] &= E[\exp(i\underset{\sim}{Z}'P'\Theta P\underset{\sim}{Z})] \\
&= E[\exp(i\underset{\sim}{Z}'D\underset{\sim}{Z})] \\
&= E[\exp(i\sum_{j=1}^{p} \lambda_j Z_j^2)] \\
&= \prod_{j=1}^{p} E[\exp(i\lambda_j Z_j^2)].
\end{aligned}$$

Since $Z_j \sim N(0,1)$ and Z_j's are independently distributed, so $Z_j^2 \sim \chi^2_{(1)}$. The characteristic function of a Chi-square random variable with one degree of freedom is

$$\phi_{\chi^2}(t) = (1-2it)^{-1/2}.$$

So

$$E[\exp(i\lambda_j Z_j^2)] = (1-2i\lambda_j)^{-1/2}.$$

Thus

$$\begin{aligned} E[\exp(i\underset{\sim}{Y}'\Theta\underset{\sim}{Y})] &= \prod_{j=1}^{p} E\left[\exp(i\lambda_j Z_j^2)\right] \\ &= \prod_{j=1}^{p}(1-2i\lambda_j)^{-1/2} \\ &= |I-2iD|^{-1/2}. \end{aligned}$$

The matrix $I-2iD$ is a diagonal matrix. We find

$$\begin{aligned} |I-2iD| &= |P'\Sigma^{-1}P - 2iP'\Theta P| \\ &= |P'||\Sigma^{-1}-2i\Theta||P| \\ &= |P|^2|\Sigma^{-1}-2i\Theta|. \end{aligned}$$

Since $P'\Sigma^{-1}P = I$, taking the determinant on both sides and solving it further gives

$$\begin{aligned} & |P'||\Sigma^{-1}||P| = |I| \\ \text{or} \quad & |P|^2|\Sigma^{-1}| = 1 \\ \text{or} \quad & |P|^2 = \frac{1}{|\Sigma^{-1}|}. \end{aligned}$$

Also $P'(\Sigma^{-1}-2i\Theta)P = I-2iD$. Thus

$$|I-2iD| = \frac{|\Sigma^{-1}-2i\Theta|}{|\Sigma^{-1}|}.$$

Next we find

$$E[\exp(i\underset{\sim}{Y}'\Theta\underset{\sim}{Y})] = |I-2iD|^{-1/2} = \frac{|\Sigma^{-1}|^{1/2}}{|\Sigma^{-1}-2i\Theta|^{1/2}}.$$

Finally,

$$\begin{aligned}\phi_A(\Theta) &= [E\{\exp(i\underset{\sim}{Y}'\Theta\underset{\sim}{Y})\}]^n \\ &= \frac{|\Sigma^{-1}|^{n/2}}{|\Sigma^{-1}-2i\Theta|^{n/2}} \\ &= |I-2i\Theta\Sigma|^{-n/2}\end{aligned}$$

which is a function of Σ and n only.

Additive Property of Wishart Distribution

Theorem 3: If A_j $(j=1,2,\ldots,q)$ are independently distributed and $A_j \sim W_p(\Sigma, n_j)$, then the sum of matrices

$$A=\sum_{j=1}^{q} A_j \sim W_p(\Sigma, \sum_{j=1}^{q} n_j).$$

Proof: For matrix A_j, $A_j \sim W_p(\Sigma, n_j)$, the characteristic function is

$$\phi_{A_j}(\Theta) = |I-2i\Theta\Sigma|^{-\frac{n_j}{2}}.$$

Since $A_1, A_2, \ldots, A_q$ are independently distributed, so

$$\begin{aligned}\phi_A(\Theta) &= \phi_{\sum_{j=1}^{q} A_j}(\Theta) \\ &= \phi_{A_1}(\Theta)\cdot\phi_{A_2}(\Theta)\ldots\phi_{A_q}(\Theta) \\ &= \{|I-2i\Theta\Sigma|^{-\frac{n_1}{2}}\}\cdot\{|I-2i\Theta\Sigma|^{-\frac{n_2}{2}}\}\ldots\{|I-2i\Theta\Sigma|^{-\frac{n_q}{2}}\} \\ &= |I-2i\Theta\Sigma|^{-\frac{1}{2}\sum_{j=1}^{q} n_j}.\end{aligned}$$

Thus $A \sim W_p(\Sigma, \sum_{j=1}^{q} n_j)$. This is the reproductive property of the Wishart distribution.

Distribution of Linear Transformation

Theorem 4: If $A \sim W_p(\Sigma, n)$ and C is a $p \times p$ nonsingular matrix, then for $B = C'AC$, $B \sim W_p(C'\Sigma C, n)$ where C is such that $\underset{\sim}{Z}_\alpha = C'\underset{\sim}{Y}_\alpha$ and $A=\sum_{\alpha=1}^{n} \underset{\sim}{Y}_\alpha \underset{\sim}{Y}_\alpha'$, $\underset{\sim}{Y}_\alpha \sim N_p(\underset{\sim}{0}, \Sigma)$, $\quad \alpha = 1,2,\ldots,n$ and $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_n$ are independently distributed.

Proof: Consider $A = \sum_{\alpha=1}^{n} \underset{\sim}{Y}_\alpha \underset{\sim}{Y}'_\alpha$ where $\underset{\sim}{Y}_\alpha \sim N_p(\underset{\sim}{0}, \Sigma)$, $\alpha = 1, 2, \ldots, n$ and $\underset{\sim}{Y}_\alpha$'s are independently distributed.

Let $\underset{\sim}{Z}_\alpha = C'\underset{\sim}{Y}_\alpha$. Then

$$\begin{aligned} B &= C'AC \\ &= C'(\sum_{\alpha=1}^{n} \underset{\sim}{Y}_\alpha \underset{\sim}{Y}'_\alpha)C \\ &= \sum_{\alpha=1}^{n} (C'\underset{\sim}{Y}_\alpha)(C'\underset{\sim}{Y}_\alpha)' \\ &= \sum_{\alpha=1}^{n} \underset{\sim}{Z}_\alpha \underset{\sim}{Z}'_\alpha. \end{aligned}$$

Also

$$\begin{aligned} E(\underset{\sim}{Z}_\alpha) &= \underset{\sim}{0} \\ E(\underset{\sim}{Z}_\alpha \underset{\sim}{Z}'_\alpha) &= C'E(\underset{\sim}{Y}_\alpha \underset{\sim}{Y}'_\alpha)C = C'\Sigma C. \end{aligned}$$

Since $\underset{\sim}{Y}_\alpha$'s are independently distributed, so $\underset{\sim}{Z}_\alpha$ are also independent and $\underset{\sim}{Z}_\alpha \sim N_p(\underset{\sim}{0}, C'\Sigma C)$. Thus $B = \sum_{\alpha=1}^{n} \underset{\sim}{Z}_\alpha \underset{\sim}{Z}'_\alpha \sim W_p(C'\Sigma C, n)$.

Marginal Distribution

Theorem 5: If $A = \sum_{\alpha=1}^{n} \underset{\sim}{Z}_\alpha \underset{\sim}{Z}'_\alpha$ is following $W_p(\Sigma, n)$. Suppose A and Σ are partitioned as

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

The orders of A_{11}, A_{12} and A_{22} are $(q \times q)$, $q \times (p-q)$ and $(p-q) \times (p-q)$, respectively. Similarly, the orders of Σ_{11}, Σ_{12} and Σ_{22} are $(q \times q)$, $q \times (p-q)$ and $(p-q) \times (p-q)$, respectively. Then A_{11} is distributed as $W_q(\Sigma_{11}, n)$.

Proof: Let $\underset{\sim}{Z}_\alpha$'s, $\alpha = 1, 2, \ldots, n$ are independent and each $\underset{\sim}{Z}_\alpha \sim N_p(\underset{\sim}{0}, \Sigma)$. Partition $\underset{\sim}{Z}_\alpha$ as $\underset{\sim}{Z}_\alpha = [\underset{\sim}{Z}_\alpha^{(1)}, \underset{\sim}{Z}_\alpha^{(2)}]'$ where $\underset{\sim}{Z}_\alpha^{(1)}$ is of order $q \times 1$ and $\underset{\sim}{Z}_\alpha^{(2)}$ is of order $(p-q) \times 1$. From the property of marginal distribution of $N_p(\underset{\sim}{\mu}, \Sigma)$, we know that $\underset{\sim}{Z}_1^{(1)}, \underset{\sim}{Z}_2^{(1)}, \ldots, \underset{\sim}{Z}_n^{(1)}$ are independent and each $\underset{\sim}{Z}_\alpha^{(1)} \sim N_q(\underset{\sim}{0}, \Sigma_{11}), \alpha = 1, 2, \ldots, n$.

Also A_{11} is distributed as $\sum_{\alpha=1}^{n} \underset{\sim}{Z}_{\alpha}^{(1)} \underset{\sim}{Z}_{\alpha}^{(1)'}$, so A_{11} has $W_q(\Sigma_{11}, n)$ distribution.

Similarly, A_{22} is distributed as $\sum_{\alpha=1}^{n} \underset{\sim}{Z}_{\alpha}^{(2)} \underset{\sim}{Z}_{\alpha}^{(2)'}$, so A_{22} has $W_{p-q}(\Sigma_{22}, n)$ distribution.

This result can be generalized to the case when $\underset{\sim}{Z}_{\alpha}$ is partitioned into q sub-vectors as follows.

Theorem 6: Let A and Σ are partitioned into $p_1, p_2, \ldots, p_q$ rows and columns $(p_1 + p_2 + \ldots + p_q = p)$ as follows:

$$A = \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1q} \\ A_{21} & A_{22} & \ldots & A_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ A_{q1} & A_{q2} & \ldots & A_{qq} \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \ldots & \Sigma_{1q} \\ \Sigma_{21} & \Sigma_{22} & \ldots & \Sigma_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{q1} & \Sigma_{q2} & \ldots & \Sigma_{qq} \end{pmatrix}.$$

If A is following $W_p(\Sigma, n)$ and if $\Sigma_{ij} = 0$ for all $i, j = 1, 2, \ldots, q$ $(i \neq j)$, then $A_{11}, A_{22}, \ldots, A_{qq}$ are independently distributed and $A_{jj} \sim W_{p_j}(\Sigma_{jj}, n), j = 1, 2, \ldots, q$.

Proof: The random matrix A is distributed as $A = \sum_{\alpha=1}^{n} \underset{\sim}{Z}_{\alpha} \underset{\sim}{Z}'_{\alpha}$ where $\underset{\sim}{Z}_{\alpha}$'s are independently distributed and each $\underset{\sim}{Z}_{\alpha} \sim N(\underset{\sim}{0}, \Sigma)$.

Partition $\underset{\sim}{Z}_{\alpha}$ as $\underset{\sim}{Z}_{\alpha} = (\underset{\sim}{Z}_{\alpha}^{(1)}, \underset{\sim}{Z}_{\alpha}^{(2)}, \ldots, \underset{\sim}{Z}_{\alpha}^{(q)})'$ where the orders of $\underset{\sim}{Z}_{\alpha}^{(1)}, \underset{\sim}{Z}_{\alpha}^{(2)}, \ldots, \underset{\sim}{Z}_{\alpha}^{(q)}$ are $p_1 \times 1, p_2 \times 1, \ldots, p_q \times 1$, respectively.

Since $\Sigma_{ij} = 0$, so this implies that $\underset{\sim}{Z}_{1}^{(1)}, \underset{\sim}{Z}_{2}^{(1)}, \ldots, \underset{\sim}{Z}_{n}^{(1)}, \underset{\sim}{Z}_{1}^{(1)}, \underset{\sim}{Z}_{2}^{(2)}, \ldots, \underset{\sim}{Z}_{2}^{(n)}, \ldots, \underset{\sim}{Z}_{1}^{(q)}, \underset{\sim}{Z}_{2}^{(q}, \ldots, \underset{\sim}{Z}_{n}^{(q)}$, (i.e., $\underset{\sim}{Z}_{\alpha}^{(1)}, \underset{\sim}{Z}_{\alpha}^{(2)}, \ldots, \underset{\sim}{Z}_{\alpha}^{(n)}, \alpha = 1, 2, \ldots, n$) are also independent. Then $A_{11} = \sum_{\alpha=1}^{n} \underset{\sim}{Z}_{\alpha}^{(1)} \underset{\sim}{Z}_{\alpha}^{(1)'}$, $A_{22} = \sum_{\alpha=1}^{n} \underset{\sim}{Z}_{\alpha}^{(2)} \underset{\sim}{Z}_{\alpha}^{(2)'}, \ldots, A_{qq} = \sum_{\alpha=1}^{n} \underset{\sim}{Z}_{\alpha}^{(q)} \underset{\sim}{Z}_{\alpha}^{(q)'}$ are also independent. Hence using Theorem 5 it is concluded that $A_{jj} \sim W_{p_j}(\Sigma_{jj}, n)$.

Conditional Distribution

Suppose the p-component vectors $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_n$ are independently distributed and each $\underset{\sim}{Y}_{\alpha} \sim N_p(\underset{\sim}{0}, \Sigma)$, $\alpha = 1, 2, \ldots, n$. Then $A = \sum_{\alpha=1}^{n} \underset{\sim}{Y}_{\alpha} \underset{\sim}{Y}'_{\alpha} \sim W_p(\Sigma, n)$.

Suppose $\underset{\sim}{Y}_{\alpha}$'s are partitioned as $\underset{\sim}{Y}_{\alpha} = (\underset{\sim}{Y}_{\alpha}^{(1)}, \underset{\sim}{Y}_{\alpha}^{(2)})'$ such that $\underset{\sim}{Y}_{\alpha}^{(1)}$ and $\underset{\sim}{Y}_{\alpha}^{(2)}$ are p and $(p-q)$ component vectors, respectively. The A and Σ matrices are suitably partitioned as

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

where the orders of A_{11} and Σ_{11} are $p \times p$, orders of each A_{22} and Σ_{22} are $(p-q) \times (p-q)$ and orders of each A_{12} and Σ_{12} are $p \times (p-q)$.

Theorem 7: The distribution of $A_{22.1} = A_{22} - A_{21}A_{11}^{-1}A_{12}$ is $W_{p-q}(\Sigma_{22.1}, n-p)$ where $\Sigma_{22.1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$.

Proof: The representation $A = \sum_{\alpha=1}^{n} \underset{\sim}{Y}_\alpha \underset{\sim}{Y}_\alpha'$ can be expressed as $A = Y'Y$ where Y is a $n \times p$ matrix. The rows of Y are i.i.d. and follow $N_p(\underset{\sim}{0}, \Sigma)$. Similarly, $A_{11} = \sum_{\alpha=1}^{n} \underset{\sim}{Y}_\alpha^{(1)} \underset{\sim}{Y}_\alpha^{(1)'}$, $A_{22} = \sum_{\alpha=1}^{n} \underset{\sim}{Y}_\alpha^{(2)} \underset{\sim}{Y}_\alpha^{(2)'}$ and $A_{12} = \sum_{\alpha=1}^{n} \underset{\sim}{Y}_\alpha^{(1)} \underset{\sim}{Y}_\alpha^{(2)'}$ can be expressed as $A_{11} = Y_1'Y_1$, $A_{22} = Y_2'Y_2$ and $A_{12} = Y_1'Y_2$, respectively.

Then $A_{22.1}$ can be expressed as

$$\begin{aligned} A_{22.1} &= A_{22} - A_{21}A_{11}^{-1}A_{12} \\ &= Y_2'Y_2 - Y_2'Y_1A_{11}^{-1}Y_1'Y_2 \\ &= Y_2'[I - Y_1A_{11}^{-1}Y_1']Y_2 \\ &= Y_2'PY_2 \end{aligned}$$

where $P = I - Y_1A_{11}^{-1}Y_1'$ is a symmetric and idempotent matrix with rank $(n-p)$.

Note that

$$PY_1 = Y_1 - Y_1A_{11}^{-1}Y_1'Y_1 = Y_1 - Y_1A_{11}^{-1}A_{11} = 0.$$

Let $Y_{2.1} = Y_2 - Y_1\Sigma_{11}^{-1}\Sigma_{12}$.

Consider

$$\begin{aligned} Y_{2.1}'PY_{2.1} &= (Y_2' - \Sigma_{21}\Sigma_{11}^{-1}Y_1')(I - Y_1A_{11}^{-1}Y_1')(Y_2 - Y_1\Sigma_{11}^{-1}\Sigma_{12}) \\ &= (Y_2' - Y_2'Y_1A_{11}^{-1}Y_1' - \Sigma_{21}\Sigma_{11}^{-1}Y_1' + \Sigma_{21}\Sigma_{11}^{-1}Y_1'Y_1A_{11}^{-1}Y_1')(Y_2 - Y_1\Sigma_{11}^{-1}\Sigma_{12}) \\ &= Y_2'Y_2 - Y_2'Y_1A_{11}^{-1}Y_1'Y_2 \\ &= Y_2'PY_2. \end{aligned}$$

Also, the covariance between $Y_{2.1}$ and Y_1 is

$$Cov(Y_1, Y_2 - Y_1\Sigma_{11}^{-1}\Sigma_{12}) = \Sigma_{12} - \Sigma_{11}\Sigma_{11}^{-1}\Sigma_{12} = 0.$$

So $Y_{2.1}$ and Y_1 are independent.

Thus $Y_{2.1} \sim N_{p-q}(\underset{\sim}{0}, \Sigma_{22.1})$.

Next, we use the following result:

Suppose X follows $N(0, \Sigma)$ where X is an $n \times p$ matrix and C is an $n \times n$ symmetric matrix, then $X'CX$ has a Wishart distribution if and only if C is idempotent and $X'CX \sim W_p(\Sigma, r)$ where $r = tr(C) = rank(C)$.

Using this result, we can conclude that

$$Y_{2.1}'PY_{2.1} \sim W_{p-q}(\Sigma_{22.1}, n-p)$$

for any given Y_1 as P is idempotent matrix and its rank is $(n-p)$.

This conditional distribution is free of Y_1. Therefore, it is the unconditional distribution, and moreover, $A_{22.1}$ is independent of X_1.

The Generalized Variance

The multivariate analogue of the variance σ^2 of a univariate distribution is the covariance matrix Σ. Another multivariate analogue is the scalar $|\Sigma|$, which is called the generalised variance of the multivariate distribution.

The generalized variance of the sample of vectors $\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_N$ is defined as

$$|S| = \left| \frac{1}{N-1} \sum_{\alpha=1}^{N} (\underset{\sim}{x}_\alpha - \bar{\underset{\sim}{x}})(\underset{\sim}{x}_\alpha - \bar{\underset{\sim}{x}})' \right|$$

In some sense, it is a measure of spread in a multivariate setup. Such a term arises when constructing the likelihood ratio test, and we will use it in the topic of the Hotelling T^2 distribution.

Similarly, we denote

$$|A| = \left| \sum_{\alpha=1}^{N} (\underset{\sim}{x}_\alpha - \bar{\underset{\sim}{x}})(\underset{\sim}{x}_\alpha - \bar{\underset{\sim}{x}})' \right|.$$

The probability distribution of $|A|$ is the same as the distribution of the product of χ^2 variables with $n, n-1, n-2, \ldots, n-p+1$ degrees of freedom. The exact form of the distribution of the product of χ^2 variables with $n, n-1, \ldots, n-p+1$ degrees of freedom is not known.

The following theorem describes this result.

Theorem 8: Given a random sample $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_N$ from $N_p(\underset{\sim}{\mu}, \Sigma)$, the distribution of the sample generalized variance $|S|$ is the same as the distribution of $\frac{|\Sigma|}{(N-1)^p}$ (or $\frac{|\Sigma|}{n^p}$) times the product of p independent factors where the distribution of i^{th} factor being χ^2 with $(N-i)$ degrees of freedom.

Also, teh asymptotic distribution is

$$\sqrt{n}\left(\frac{|S|}{|\Sigma|} - 1\right) \sim N(0, 2p)$$

and

$$E(|S|) = \frac{|\Sigma|}{n^p} \prod_{i=1}^{p} E(\chi^2_{N-i}) = \frac{|\Sigma|}{n^p} \prod_{i=1}^{p} (N-i).$$

Quadratic Forms in Normal Variables

Let $\underset{\sim}{X} \sim N_p(\underset{\sim}{0}, I)$, the necessary and sufficient condition that $\underset{\sim}{X}' A \underset{\sim}{X}$ to be distributed as χ^2 distribution is that A is idempotent, and the degrees of freedom of the χ^2 distribution will be the same as rank of A, i.e., trace of A matrix.

Theorem 9 (Craig's theorem): Two quadratic forms $\underset{\sim}{X}' A \underset{\sim}{X}$ and $\underset{\sim}{X}' B \underset{\sim}{X}$ are independently distributed if and only if $AB = 0$ where A and B are symmetric and idempotent matrices.

The necessity of this condition is proved by equating the joint characteristic function to the product of the two characteristic functions. The sufficiency of the condition $AB = 0$ is proved by simultaneously diagonalizing A and B with the help of an orthogonal matrix and showing that no non-zero terms appear at the same positions in each of the diagonal matrices.

Theorem 10 (James theorem): Let $Q = \underset{\sim}{X}'\underset{\sim}{X} = \sum_{i=1}^{k} \underset{\sim}{X}'A_i\underset{\sim}{X}$, then for the quadratic forms $\underset{\sim}{X}'A_i\underset{\sim}{X}$ to be independently distributed as χ^2, any one of the following three equivalent conditions is necessary and sufficient:

(I) $A_i^2 = A_i$, $i = 1, 2, \ldots, k$

(II) $A_i A_j = 0$, $i \neq j$, $j = 1, 2, \ldots, k$

(III) $\sum_{i=1}^{k} rank(A_i) = p$.

This is a generalized version of Cochran's theorem, which states only condition (III). The theorem also holds when $Q = \underset{\sim}{X}'A\underset{\sim}{X} = \sum_{i=1}^{k} \underset{\sim}{X}'A_i\underset{\sim}{X}$ where A is idempotent.

Theorem 11 (Cochran's theorem): Suppose $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_N$ are independently distributed and $\underset{\sim}{Y}_\alpha \sim N_p(\underset{\sim}{0}, \Sigma)$, $\alpha = 1, 2, \ldots, N$. Suppose C_i is a $N \times N$ symmetric matrix with rank $(C_i) = r_i$, $i = 1, 2, \ldots, m$ and $\sum_{i=1}^{m} C_i = I_N$. Suppose the matrix $C_i = ((C_{\alpha\beta}^i))$ is used in forming $Q_i = \sum_{\alpha=1}^{N}\sum_{\beta=1}^{N} C_{\alpha\beta}^i \underset{\sim}{Y}_\alpha \underset{\sim}{Y}_\beta'$ and suppose $\sum_{i=1}^{m} Q_i = \sum_{\alpha=1}^{N} \underset{\sim}{Y}_\alpha \underset{\sim}{Y}_\alpha'$ $(i = 1, 2, , \ldots, m)$. Then a necessary and sufficient condition that $Q_1, Q_2, \ldots, Q_m$ are independently distributed with $Q_i \sim W_p(\Sigma, r_i)$ is $\sum_{i=1}^{m} r_i = N$.

Inverted Wishart Distribution

If $A \sim W_p(\Sigma, n)$, then $B = A^{-1}$ has the inverted Wishart distribution and its density is

$$\frac{1}{2^{\frac{np}{2}} \pi^{\frac{p(p-1)}{4}} \prod_{i=0}^{p-1} \Gamma\left(\frac{n-i}{2}\right)} |\Sigma^{-1}|^{\frac{n}{2}} |B|^{-\frac{n+p+1}{2}} \exp\left(-\frac{1}{2} tr\Sigma^{-1}B^{-1}\right).$$

The proof follows by $f(B) = f(A) \cdot |J|$, where J is the Jacobian of the transformation.

Noncentral Wishart Distribution

Suppose the columns of a matrix X are independently distributed about a non-null mean vector given by a matrix M, then the distribution of $A = XX'$ is a noncentral Wishart distribution with $\Omega = \Sigma^{-1}M'M$ as the noncentrality matrix of the noncentral Wishart distribution.

If $\underset{\sim}{X}_\alpha \sim N_p(M, \Sigma)$ and $A = \sum_{\alpha=1}^{n} \underset{\sim}{X}_\alpha \underset{\sim}{X}'_\alpha$ then A has a noncentral Wishart distribution with PDF

$$f(A) = \frac{|A|^{\frac{n-p-1}{2}} \exp(-\frac{1}{2} tr \Sigma^{-1} A) \exp(-\frac{\Omega}{2})}{2^{\frac{np}{2}} \pi^{\frac{p(p-1)}{4}} \prod_{i=0}^{p-1} \Gamma(\frac{n-i}{2}) |\Sigma|^{\frac{n}{2}}} {}_0F_1\left(\frac{n}{2}, \frac{1}{4}\Omega\Sigma^{-1}A\right)$$

where the Hypergeometric function is defined as

$${}_pF_q(a_1, \ldots, a_p; b_1, \ldots, b_q; Z) = \sum_{k=0}^{\infty} \frac{(a_1)_k \ldots (a_p)_k}{(b_1)_k \ldots (b_q)_k} \cdot \frac{z^k}{k!}$$

and $(a)_k = a(a+1)\ldots(a+k-1)$.

It is denoted as $A \sim W_p(\Sigma, n, \Omega)$.

Some Useful Results for Wishart Distribution

Lemma 11: If $A \sim W_p(\Sigma, n)$ then $E(A) = n\Sigma$ and $E(A^{-1}) = \frac{1}{n-p-1}\Sigma^{-1}$.

Lemma 12: If $A \sim W_p(\Sigma, n)$, $\underset{\sim}{h}$ is a $p \times 1$ vector and is independently distributed of A, then $u = \frac{\underset{\sim}{h}' A \underset{\sim}{h}}{\underset{\sim}{h}' \Sigma \underset{\sim}{h}}$ has the χ^2 distribution with n degrees of freedom and is independent of $\underset{\sim}{h}$.

Lemma 13: If $A \sim W_p(\Sigma, n)$, then $\frac{\sigma^{pp}}{a^{pp}}$ has the χ^2 distribution with $(n-p+1)$ degrees of freedom where a^{pp} and σ^{pp} are the last (p^{th}) diagonal elements of A^{-1} and Σ^{-1} matrices, respectively.

Lemma 14: If $A \sim W_p(\Sigma, n)$ and $\underset{\sim}{h}$ is a $p \times 1$ vector distributed independently of A, then

$$\frac{\underset{\sim}{h}' \Sigma^{-1} \underset{\sim}{h}}{\underset{\sim}{h}' A^{-1} \underset{\sim}{h}}$$

has the χ^2 distribution with $(n-p+1)$ degrees of freedom and is independent of $\underset{\sim}{h}$.

Some Useful Results for Jacobian of Transformation

Suppose we want to find the distribution of Y given the distribution of X.

1. $Y = AXB$ where the order of Y is $p \times q$, the order of A is $p \times p$, the order of X is $p \times q$ and the order of B is $q \times q$. Then the Jacobian of transformation $|J|$, denoted as $J(Y \to X) = |\frac{\partial Y}{\partial X}| = |A|^q |B|^p$ where the order of $\frac{\partial Y}{\partial X}$ is $pq \times pq$.

2. $Y = AXA'$ where the order of X is $p \times p$ symmetric, the order of Y is $p \times p$, the order of A is $p \times p$ and the order of X is $p \times p$ then $J(Y \to X) = |A|^{p+1}$.

3. If $|X| \neq 0$, $Y = X^{-1}$, where the order of X is $p \times p$, then

   - $J(Y \to X) = |X|^{-(p+1)}$ if X is symmetric
   - $J(Y \to X) = |X|^{-2p}$ if X is non-symmetric.

4. For lower triangular matrices X, Y and A, each of order $p \times p$,

   - If $Y = AX$ then $J(Y \to X) = \prod_{i=1}^{p} a_{ii}^{i}$.
   - If $Y = XA$ then $J(Y \to X) = \prod_{i=1}^{p} a_{ii}^{p-1}$.
   - If $Y = XA' + AX'$ then $J(Y \to X) = 2^p \prod_{i=1}^{p} a_{ii}^{p-i+1}$.
   - If $Y > 0,\ Y = XX'$ then $J(Y \to X) = 2^p \prod_{i=1}^{p} a_{ii}^{p-i+1}$.

5. For upper triangular matrices X and A, if $Y = XA' + AX'$, then $J(Y \to X) = 2^p \prod_{i=1}^{p} a_{ii}^{i}$.