Multivariate Statistics

Lecture 10

Fudan University

Outline

Outline

1. The Density of the Wishart Distribution

2. Properties of the Wishart Distribution

Outline

1. The Density of the Wishart Distribution
2. Properties of the Wishart Distribution
3. The Generalized Variance

Outline

Outline

Outline

The Wishart Distribution

We shall obtain the distribution of

$$\mathbf{A} = \sum_{\alpha=1}^{N} (\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})^\top,$$

where $\mathbf{x}_1, \ldots, \mathbf{x}_N$ are independent, each with the distribution $\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $N > p$.

We have shown that $\mathbf{A}$ is distributed as $\sum_{\alpha=1}^{n} \mathbf{z}_\alpha \mathbf{z}_\alpha^\top$ where $n = N - 1$ and $\mathbf{z}_1, \ldots, \mathbf{z}_n$ are independent, each with the distribution $\mathcal{N}_p(\mathbf{0}, \boldsymbol{\Sigma})$.

We shall show that the density of $\mathbf{A}$ for $\mathbf{A}$ positive definite is

$$\frac{(\det(\mathbf{A}))^{\frac{n-p-1}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\mathbf{A}\right)\right)}{2^{\frac{np}{2}} \pi^{\frac{p(p-1)}{4}} (\det(\boldsymbol{\Sigma}))^{\frac{n}{2}} \prod_{i=1}^{p} \Gamma\left(\frac{1}{2}(n+1-i)\right)}.$$

The Wishart Distribution

We shall first consider the case of $\boldsymbol{\Sigma} = \mathbf{I}$. Let

$$\begin{bmatrix} \mathbf{z}_1 & \dots & \mathbf{z}_n \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1^\top \\ \vdots \\ \mathbf{v}_p^\top \end{bmatrix} \in \mathbb{R}^{p \times n}.$$

Then the (i, j)-th elements of $\mathbf{A}$ can be written as

$$a_{ij} = \mathbf{v}_i^\top \mathbf{v}_j$$

and vectors $\mathbf{v}_1, \dots, \mathbf{v}_p$ are independently distributed according to $\mathcal{N}_n(\mathbf{0}, \mathbf{I})$.

The Wishart Distribution

Applying Gram-Schmidt orthogonalization on $\mathbf{v}_1, \ldots, \mathbf{v}_p$.

1. Let $\mathbf{w}_1 = \mathbf{v}_1$ and $\mathbf{w}_i = \mathbf{v}_i - \sum_{j=1}^{i-1} \frac{\mathbf{w}_j^\top \mathbf{v}_i}{\|\mathbf{w}_j\|_2^2} \cdot \mathbf{w}_j$ for $i = 2, \ldots, p$.
2. We can prove by induction that $\mathbf{w}_k$ is orthogonal to $\mathbf{w}_i$ for $k < i$.
3. We can show that $\Pr\left(\|\mathbf{w}_i\|_2 = 0\right) = \Pr\left(\mathrm{rank}(\mathbf{A}) < p\right) = 0$.

Define the $p \times p$ lower triangular matrix $\mathbf{T}$ $(t_{ij} = 0$ for $i < j)$ with

$$t_{ii} = \|\mathbf{w}_i\|_2 \quad \text{for} \quad i = 1, \ldots, p;$$

$$t_{ij} = \frac{\mathbf{w}_j^\top \mathbf{v}_i}{\|\mathbf{w}_j\|_2} \quad \text{for} \quad j = 1, \ldots, i-1, \quad i = 2, \ldots, p.$$

Then we have

$$\mathbf{v}_i = \sum_{j=1}^{i} \frac{t_{ij}\mathbf{w}_j}{\|\mathbf{w}_j\|_2}, \quad \begin{bmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_p \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|_2} & \cdots & \frac{\mathbf{w}_p}{\|\mathbf{w}_p\|_2} \\ | & & | \end{bmatrix} \mathbf{T}^\top \text{ and } \mathbf{A} = \mathbf{T}\mathbf{T}^\top.$$

The formula

$$\mathbf{v}_i = \sum_{j=1}^{i} \frac{t_{ij}}{\|\mathbf{w}_j\|_2} \cdot \mathbf{w}_j$$

means t_{ij} for $j = 1, \ldots, i-1$ are the first $i-1$ coordinates of $\mathbf{v}_i$ in the coordinate system with $\mathbf{w}_1, \ldots, \mathbf{w}_{i-1}$.

The sum of the other $n - i + 1$ coordinates squared is

$$\|\mathbf{v}_i\|_2^2 - \sum_{j=1}^{i-1} t_{ij}^2 = t_{ii}^2 = \|\mathbf{w}_i\|_2^2 .$$

The Wishart Distribution

There exist $\mathbf{w}_i', \ldots, \mathbf{w}_n'$ and $t_{ii}', \ldots, t_{in}'$ such that

$$\mathbf{v}_i = \sum_{j=1}^{i-1} \frac{t_{ij}}{\|\mathbf{w}_j\|_2} \cdot \mathbf{w}_j + \sum_{j=i}^{n} \frac{t_{ij}'}{\|\mathbf{w}_j'\|} \cdot \mathbf{w}_j' = \mathbf{W}_i \mathbf{t}_i'$$

where

$$\mathbf{t}_i' = \begin{bmatrix} t_{i1} \\ \vdots \\ t_{ii-1} \\ t_{ii}' \\ \vdots \\ t_{in}' \end{bmatrix} \quad \text{and} \quad \mathbf{W}_i = \begin{bmatrix} | & & | & | & & | \\ \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} & \cdots & \frac{\mathbf{w}_{i-1}}{\|\mathbf{w}_{i-1}\|} & \frac{\mathbf{w}_i'}{\|\mathbf{w}_i'\|} & \cdots & \frac{\mathbf{w}_n'}{\|\mathbf{w}_n'\|} \\ | & & | & | & & | \end{bmatrix} \in \mathbb{R}^{n \times n}$$

is orthogonal. Then we have $\mathbf{t}_i' = \mathbf{W}_i^\top \mathbf{v}_i$.

Lemma 1

Conditional on $\mathbf{w}_1, \ldots, \mathbf{w}_{i-1}$ (or equivalently on $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$), then random variables $t_{i1}, \ldots, t_{ii-1}$ are independently distributed and t_{ij} is distributed according to $\mathcal{N}(0, 1)$ for $i > j$; and t_{ii}^2 has the χ^2-distribution with $n - i + 1$ degrees of freedom.

The sketch of the proof:

1. Conditional on $\mathbf{w}_1, \ldots, \mathbf{w}_{i-1}$, the matrix $\mathbf{W}_i$ is fixed.
2. We have $\mathbf{t}'_i = \mathbf{W}_i^\top \mathbf{v}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ since $\mathbf{v}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and $\mathbf{W}^\top \mathbf{W} = \mathbf{I}$.
3. We have $t_{ii}^2 = \|\mathbf{v}_i\|_2^2 - \sum_{j=1}^{i-1} t_{ij}^2 = \sum_{j=i}^{n} {t'_{ij}}^2$, where each t'_{ij} are independently distributed according to $\mathcal{N}(0, 1)$ for $j = i, \ldots, n$.

The Wishart Distribution

Since the conditional distribution of $t_{i1}, \ldots, t_{ii}$ does not depend on $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$, they are distributed independently of $t_{11}, t_{21}, t_{22}, \ldots, t_{i-1,i-1}$.

Corollary 1

Let $\mathbf{z}_1, \ldots, \mathbf{z}_n$ be independently distributed, each according to $\mathcal{N}_p(\mathbf{0}, \mathbf{I})$, where $n \geq p$; let

$$\mathbf{A} = \sum_{\alpha=1}^{n} \mathbf{z}_\alpha \mathbf{z}_\alpha^\top = \mathbf{T}\mathbf{T}^\top,$$

where $t_{ij} = 0$ for $i < j$, and $t_{ii} > 0$ for $i = 1, \ldots, p$. Then $t_{11}, t_{21}, \ldots, t_{pp}$ are independently distributed; t_{ij} is distributed according to $\mathcal{N}(0, 1)$ for $i > j$; and t_{ii}^2 has the χ^2-distribution with $n - i + 1$ degrees of freedom.

The Wishart Distribution

Theorem 2

Let $\mathbf{z}_1, \ldots, \mathbf{z}_n$ be independently distributed, each according to $\mathcal{N}_p(\mathbf{0}, \mathbf{\Sigma})$, where $n \geq p$; let

$$\mathbf{A} = \sum_{\alpha=1}^{n} \mathbf{z}_\alpha \mathbf{z}_\alpha^\top = \mathbf{T}^* \mathbf{T}^{*\top},$$

where $t_{ij}^* = 0$ for $i < j$, and $t_{ii}^* > 0$ for $i = 1, \ldots, p$. Then the density of $\mathbf{T}^*$ is

$$\frac{\prod_{i=1}^{p} t_{ii}^{*\,n-i} \exp\left(-\frac{1}{2}\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\mathbf{T}^*\mathbf{T}^{*\top}\right)\right)}{2^{\frac{p(n-2)}{2}} \pi^{\frac{p(p-1)}{4}} (\det(\mathbf{\Sigma}))^{\frac{n}{2}} \prod_{i=1}^{p} \Gamma\left(\frac{1}{2}(n+1-i)\right)}.$$

The Wishart Distribution

Theorem 3

Let $\mathbf{z}_1, \ldots, \mathbf{z}_n$ be independently distributed, each according to $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$, where $n \geq p$. Then the density of $\mathbf{A} = \sum_{\alpha=1}^{n} \mathbf{z}_\alpha \mathbf{z}_\alpha^\top$ is

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

for $\mathbf{A}$ positive definite, and 0 otherwise.

Corollary 2

Let $\mathbf{x}_1, \ldots, \mathbf{x}_N$ be independently distributed, each according to $\mathcal{N}_p(\boldsymbol{\mu}, \mathbf{\Sigma})$, where $N > p$; Then the density of $\mathbf{A} = \sum_{\alpha=1}^{N} (\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})^\top$ is (1), where $n = N - 1$ and $\mathbf{x} = \frac{1}{N}\sum_{\alpha=1}^{N} \mathbf{x}_\alpha$.

The Wishart Distribution

The multivariate gamma function is defined as

$$\Gamma_p(t) = \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^{p} \Gamma\Big(t - \frac{1}{2}(i-1)\Big).$$

Then the Wishart density can be written as

$$\frac{(\det(\mathbf{A}))^{\frac{n-p-1}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\mathbf{A}\right)\right)}{2^{\frac{np}{2}} (\det(\mathbf{\Sigma}))^{\frac{n}{2}} \Gamma_p\left(\frac{n}{2}\right)}.$$

The Wishart Distribution

We denote the density of the Wishart distribution as

$$w(\mathbf{A} \mid \boldsymbol{\Sigma}, n) = \frac{(\det(\mathbf{A}))^{\frac{n-p-1}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\mathbf{A}\right)\right)}{2^{\frac{np}{2}} (\det(\boldsymbol{\Sigma}))^{\frac{n}{2}} \Gamma_p\left(\frac{n}{2}\right)}$$

and the associated distribution will be termed

$$\mathbf{A} \sim \mathcal{W}(\boldsymbol{\Sigma}, n).$$

If $n < p$, then $\mathbf{A}$ does not have a density, but its distribution is nevertheless defined, and we shall refer to it as $\mathcal{W}(\boldsymbol{\Sigma}, n)$.

Corollary 3

Let $\mathbf{x}_1, \dots, \mathbf{x}_N$ be independently distributed, each according to $\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $N > p$. Then the distribution of $\mathbf{S} = \frac{1}{n}\sum_{\alpha=1}^{N}(\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})^\top$ is $\mathcal{W}\left(\frac{1}{n}\boldsymbol{\Sigma}, n\right)$.

Outline

The Characteristic Function of the Wishart Distribution

Lemma 2

Given $\mathbf{B}$ positive semidefinite and $\mathbf{A}$ positive definite, there exists a non-singular matrix $\mathbf{F}$ such that $\mathbf{F}^\top\mathbf{B}\mathbf{F} = \mathbf{D}$ and $\mathbf{F}^\top\mathbf{A}\mathbf{F} = \mathbf{I}$, where $\mathbf{D}$ is diagonal.

Lemma 3

The characteristic function of chi-square distribution with the degree of freedom n is

$$\phi(t) = (1 - 2\mathrm{i}t)^{-\frac{n}{2}}.$$

The Characteristic Function of the Wishart Distribution

Theorem 4

If $\mathbf{z}_1, \ldots, \mathbf{z}_n$ are independent, each with distribution $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$, then the characteristic function of $a_{11}, \ldots, a_{pp}, 2a_{12}, \ldots, 2a_{p-1,p}$, where a_{ij} is the (i,j)-th element of

$$\mathbf{A} = \sum_{\alpha=1}^{n} \mathbf{z}_\alpha \mathbf{z}_\alpha^\top$$

is given by

$$\mathbb{E}\left[\exp(\mathrm{i}\,\mathrm{tr}(\mathbf{A}\mathbf{\Theta}))\right] = \left(\det\left(\mathbf{I} - 2\mathrm{i}\mathbf{\Theta}\mathbf{\Sigma}\right)\right)^{-\frac{n}{2}}.$$

The Sum of Wishart Matrices

If $\mathbf{A}_1, \ldots, \mathbf{A}_q$ are independently distributed with $\mathbf{A}_i \sim \mathcal{W}(\boldsymbol{\Sigma}, n_i)$ for $i = 1, \ldots, q$, then

$$\mathbf{A} = \sum_{i=1}^{q} \mathbf{A}_i \sim \mathcal{W}\left(\boldsymbol{\Sigma}, \sum_{i=1}^{q} n_i\right).$$

If $p = 1$ and $\boldsymbol{\Sigma} = 1$, then $\mathcal{W}(\boldsymbol{\Sigma}, n)$ is a χ^2-distribution with n degrees of freedom.

Certain Linear Transformation

We shall frequently make the transformation

$$\mathbf{A} = \mathbf{C}\mathbf{B}\mathbf{C}^{-1},$$

where $\mathbf{C} \in \mathbb{R}^{p \times p}$ is non-singular.

If the random matrix $\mathbf{A}$ is distributed according to $\mathcal{W}(\mathbf{\Sigma}, n)$, then $\mathbf{B}$ is distributed according to $\mathcal{W}(\mathbf{\Phi}, n)$ where

$$\mathbf{\Phi} = \mathbf{C}^{-1}\mathbf{\Sigma}\left(\mathbf{C}^{\top}\right)^{-1}.$$

Marginal Distributions

Let $\mathbf{A}$ and $\boldsymbol{\Sigma}$ be partitioned into q and $p-q$ rows and columns,

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \qquad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

If $\mathbf{A}$ is distributed according to $\mathcal{W}(\boldsymbol{\Sigma}, n)$, then $\mathbf{A}_{11}$ is distributed according to $\mathcal{W}(\boldsymbol{\Sigma}_{11}, n)$.

Marginal Distributions

Let $\mathbf{A}$ and $\mathbf{\Sigma}$ be partitioned into $p_1, \ldots, p_q$ rows and $p_1, \ldots, p_q$ columns as

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \cdots & \mathbf{A}_{1q} \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{q1} & \cdots & \mathbf{A}_{qq} \end{bmatrix} \quad \text{and} \quad \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \cdots & \mathbf{\Sigma}_{1q} \\ \vdots & \ddots & \vdots \\ \mathbf{\Sigma}_{q1} & \cdots & \mathbf{\Sigma}_{qq}. \end{bmatrix}$$

If $\mathbf{\Sigma} = \mathbf{0}$ for $i \neq j$ and if $\mathbf{A} \sim \mathcal{W}(\mathbf{\Sigma}, n)$, then $\mathbf{A}_{11}, \ldots, \mathbf{A}_{qq}$ are independently distributed and $\mathbf{A}_{jj} \sim \mathcal{W}(\mathbf{\Sigma}_{jj}, n)$ for $j = 1, \ldots, q$.

Conditional Distributions

Let $\mathbf{A}$ and $\mathbf{\Sigma}$ be partitioned into q and $p-q$ rows and columns as

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \qquad \text{and} \qquad \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}.$$

If $\mathbf{A}$ is distributed according to $\mathcal{W}(\mathbf{\Sigma}, n)$, then the distribution of

$$\mathbf{A}_{11.2} = \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}$$

is distributed according to $\mathcal{W}(\mathbf{\Sigma}_{11.2}, n-p+q)$, where $\mathbf{\Sigma}_{11.2} = \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}$ and $n \geq p-q$.

Follow the analysis in the section of partial correlation coefficient.

Outline

The Generalized Variance

The multivariate analog of the variance of the univariate distribution:

1. Covariance matrix $\mathbf{\Sigma}$.
2. The scalar $\det(\mathbf{\Sigma})$, which is called the generalized variance.

The generalized variance of the sample of vectors $\mathbf{x}_1, \ldots, \mathbf{x}_N$ is

$$\det(\mathbf{S}) = \det\left(\frac{1}{N-1}\sum_{\alpha=1}^{N}(\mathbf{x}_\alpha - \bar{\mathbf{x}}_\alpha)(\mathbf{x}_\alpha - \bar{\mathbf{x}}_\alpha)^\top\right)$$

The Generalized Variance

Let

$$\mathbf{A} = \sum_{\alpha=1}^{N} (\mathbf{x}_\alpha - \bar{\mathbf{x}}_\alpha)(\mathbf{x}_\alpha - \bar{\mathbf{x}}_\alpha)^\top = (N-1)\mathbf{S}$$

and

$$\mathbf{X} - \bar{\mathbf{x}}\mathbf{1} = \begin{bmatrix} | & & | \\ \mathbf{x}_1 - \bar{\mathbf{x}} & \cdots & \mathbf{x}_N - \bar{\mathbf{x}} \\ | & & | \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1^\top \\ \vdots \\ \mathbf{v}_p^\top \end{bmatrix} = \mathbf{V} \in \mathbb{R}^{p \times N}.$$

The sample generalized variance comes p rows of $\mathbf{V} = \mathbf{X} - \bar{\mathbf{x}}\mathbf{1}$ as p vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$ in N-dimensional space.

We have $\det(\mathbf{S}) = \det(\mathbf{A})/(N-1)^p$.

Distribution of the Sample Generalized Variance

Consider that $\mathbf{x}_1, \ldots, \mathbf{x}_N$ are independently sampled from $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then

$$\mathbf{A} = \sum_{\alpha=1}^{n} \mathbf{z}_\alpha \mathbf{z}_\alpha^\top$$

where $\mathbf{z}_1, \ldots, \mathbf{z}_n$ are distributed independently according to $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$, and $n = N - 1$.

Let $\mathbf{z}_\alpha = \mathbf{C}\mathbf{y}_\alpha$ for $\alpha = 1, \ldots, n$, where $\mathbf{C}\mathbf{C}^\top = \boldsymbol{\Sigma}$. Then $\mathbf{y}_1, \ldots, \mathbf{y}_n$ are independently distributed, each with distribution $\mathcal{N}(\mathbf{0}, \mathbf{I})$. Let

$$\mathbf{B} = \sum_{\alpha=1}^{n} \mathbf{y}_\alpha \mathbf{y}_\alpha^\top = \sum_{\alpha=1}^{n} \mathbf{C}^{-1}\mathbf{z}_\alpha \mathbf{z}_\alpha^\top (\mathbf{C}^{-1})^\top = \mathbf{C}^{-1}\mathbf{A}(\mathbf{C}^{-1})^\top,$$

then $\det(\mathbf{A}) = \det(\mathbf{C})\det(\mathbf{B})\det(\mathbf{C}^\top) = \det(\mathbf{B})\det(\boldsymbol{\Sigma})$.

We have shown that $\det(\mathbf{B}) = \prod_{i=1}^{p} t_{ii}^2$, where $t_{11}^2, \ldots, t_{pp}^2$ are independent and t_{ii}^2 are distributed according to χ^2-distribution with $N - i$ degrees of freedom.

Distribution of the Sample Generalized Variance

$\det(\mathbf{S}) = \det(\mathbf{B})\det(\boldsymbol{\Sigma})/(N-1)^p$ equals to

$$\frac{\det(\boldsymbol{\Sigma})\prod_{i=1}^{p} t_{ii}^2}{(N-1)^p},$$

where $t_{11}^2, \ldots, t_{pp}^2$ are independent and t_{ii}^2 are distributed according to χ^2-distribution with $N-i$ degrees of freedom.

Distribution of the Sample Generalized Variance

Let $\det(\mathbf{B})/n^p = \prod_{i=1}^{p} V_i(n)$, where $V_1(n), \ldots, V_p(n)$ are independently distributed and $nV_i(n)$ is distributed according to χ^2-distribution with $n-p+i$ degrees of freedom.

Since $nV_i(n)$ is distributed as $\sum_{\alpha=1}^{n-p+i} w_\alpha^2$ where the w_α are independent, each with distribution $\mathcal{N}(0,1)$, the central limit theorem states that

$$\frac{nV_i(n) - (n-p+i)}{\sqrt{2(n-p+i)}} = \sqrt{n} \cdot \frac{V_i(n) - 1 + \frac{p-1}{n}}{\sqrt{2}\sqrt{1 - \frac{p-i}{n}}}$$

is asymptotically distributed according to $\mathcal{N}(0,1)$.

Then $\sqrt{n}(V_i(n) - 1)$ is asymptotically distributed according to $\mathcal{N}(0,2)$.

Distribution of the Sample Generalized Variance

Theorem 5 [Serfling (1980), Section 3.3]

Let $\{\mathbf{u}(n)\}$ be a sequence of m-component random vectors and $\mathbf{b}$ a fixed vector such that

$$\lim_{n\to\infty} \sqrt{n}(\mathbf{u}(n) - \mathbf{b}) \sim \mathcal{N}(\mathbf{0}, \mathbf{T}).$$

Let $\mathbf{f}(\mathbf{u})$ be a vector-valued function of $\mathbf{u}$ such that each component $f_j(\mathbf{u})$ has a nonzero differential at $\mathbf{u} = \mathbf{b}$, and let

$$\left.\frac{\partial f_j(\mathbf{u})}{\partial u_i}\right|_{\mathbf{u}=\mathbf{b}}$$

be the (i, j)-th component of $\mathbf{\Phi_b}$. Then $\sqrt{n}(\mathbf{f}(\mathbf{u}(n)) - f(\mathbf{b}))$ has the limiting distribution $\mathcal{N}(\mathbf{0}, \mathbf{\Phi_b}^\top \mathbf{T} \mathbf{\Phi_b})$.

Distribution of the Sample Generalized Variance

Let $\det(\mathbf{B})/n^p = f(\mathbf{u}) = \prod_{i=1}^{p} u_i$,

$$\mathbf{u}(n) = \begin{bmatrix} V_1(n) \\ \vdots \\ V_p(n) \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{T} = 2\mathbf{I}.$$

Then we have

$$\frac{\partial f}{\partial u_i}\bigg|_{\mathbf{u}=\mathbf{b}} = 1, \quad \phi_{\mathbf{b}} = \mathbf{1} \quad \text{and} \quad \phi_{\mathbf{b}}^\top \mathbf{T} \phi_{\mathbf{b}} = 2p,$$

which implies

$$\sqrt{n}\left(\frac{\det(\mathbf{S})}{\det(\mathbf{\Sigma})} - 1\right) = \sqrt{n}\left(\frac{\det(\mathbf{B})}{n^p} - 1\right)$$

is asymptotically distributed according to $\mathcal{N}(0, 2p)$.

Outline

Distribution of the Set of Correlation Coefficients

Recall that

$$r_{ij} = \frac{a_{ij}}{\sqrt{a_{ii}}\sqrt{a_{jj}}}.$$

When the covariance matrix is diagonal, that is

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_{11} & 0 & \cdots & 0 \\ 0 & \sigma_{22} & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{pp} \end{bmatrix} \quad \text{and} \quad \det(\mathbf{\Sigma}) = \prod_{i=1}^{p} \sigma_{ii},$$

then the density of $\{r_{ij} : i < j,\ i, j = 1, ..., p\}$ is

$$\frac{\left(\Gamma\left(\frac{n}{2}\right)\right)^p \left(\det\left([r_{ij}]_{ij}\right)\right)^{\frac{n-p-1}{2}}}{\Gamma_p\left(\frac{n}{2}\right)}.$$

Distribution of the Set of Correlation Coefficients

Sketch of the proof:

1. We consider the transformation

$$\begin{cases} a_{ij} = \sqrt{a_{ii}}\sqrt{a_{jj}}\, r_{ij} & i < j, \\ a_{ii} = a_{ii} & i = j, \end{cases}$$

   which is from $\{r_{ij} : i < j, \quad i,j = 1,\ldots,p\} \cup \{a_{ii} : i = 1,\ldots,p\}$ to $\{a_{ij} : i < j, \quad i,j = 1,\ldots,p\} \cup \{a_{ii} : i = 1,\ldots,p\}$.

2. The joint density of $\{r_{ij} : i < j, \quad i,j = 1,\ldots,p\} \cup \{a_{ii} : i = 1,\ldots,p\}$ is

$$\frac{\left(\det\left([r_{ij}]_{ij}\right)\right)^{\frac{n-p-1}{2}}}{\Gamma_p\left(\frac{n}{2}\right)} \frac{\prod_{i=1}^{p} a_{ii}^{\frac{n}{2}-1} \exp\left(-\frac{a_{ii}}{2\sigma_{ii}}\right)}{\prod_{i=1}^{p} 2^{\frac{n}{2}} \sigma_{ii}^{\frac{n}{2}}}.$$

3. Integrate out a_{ii}.

Outline

The Inverted Wishart Distribution

If $\mathbf{A}$ has the distribution $\mathcal{W}(\boldsymbol{\Sigma}, m)$, then $\mathbf{B} = \mathbf{A}^{-1}$ has the density is

$$w^{-1}(\mathbf{B} \mid \boldsymbol{\Psi}, m) = \frac{(\det(\boldsymbol{\Psi}))^{\frac{m}{2}} (\det(\mathbf{B}))^{-\frac{m+p+1}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Psi}\mathbf{B}^{-1}\right)\right)}{2^{\frac{mp}{2}}\Gamma_p\left(\frac{m}{2}\right)}.$$

for $\mathbf{B}$ positive definite and 0 elsewhere, where $\boldsymbol{\Psi} = \boldsymbol{\Sigma}^{-1}$.

1. We call $\mathbf{B}$ has the inverted Wishart distribution with m degrees of freedom and denote $\mathbf{B} \sim \mathcal{W}^{-1}(\boldsymbol{\Psi}, m)$.
2. We call $\boldsymbol{\Psi}$ the precision matrix or concentration matrix.
3. The derivation of $w^{-1}(\boldsymbol{\Psi}, m)$ are based on the determinant for Jacobian of transformation $\mathbf{A} = \mathbf{B}^{-1}$ is $(\det(\mathbf{B}))^{-(p+1)}$.

The Inverted Wishart Distribution

If the posterior distribution $p(\boldsymbol{\theta} \mid \mathbf{x})$ is in the same probability distribution family as the prior probability distribution $p(\boldsymbol{\theta})$, the prior and posterior are then called conjugate distributions, and the prior is called a conjugate prior.

Theorem 6

If $\mathbf{A}$ has the distribution $\mathcal{W}(\boldsymbol{\Sigma}, n)$ and $\boldsymbol{\Sigma}$ has the a prior distribution $\mathcal{W}^{-1}(\boldsymbol{\Psi}, m)$, then the conditional distribution of $\boldsymbol{\Sigma}$ given $\mathbf{A}$ is the inverted Wishart distribution $\mathcal{W}^{-1}(\mathbf{A} + \boldsymbol{\Psi}, n + m)$.

Corollary 4

If $n\mathbf{S}$ has the distribution $\mathcal{W}(\boldsymbol{\Sigma}, n)$ and $\boldsymbol{\Sigma}$ has the a prior distribution $\mathcal{W}^{-1}(\boldsymbol{\Psi}, m)$, then the conditional distribution of $\boldsymbol{\Sigma}$ given $\mathbf{S}$ is the inverted Wishart distribution $\mathcal{W}^{-1}(n\mathbf{S} + \boldsymbol{\Psi}, n + m)$.

The Inverted Wishart Distribution

Theorem 7

Let $\mathbf{x}_1, \ldots, \mathbf{x}_N$ be observations from $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Suppose $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ have the a prior density

$$n\left(\boldsymbol{\mu} \;\middle|\; \boldsymbol{\nu}, \frac{\boldsymbol{\Sigma}}{K}\right) \times w^{-1}(\boldsymbol{\Sigma} \mid \boldsymbol{\Psi}, m),$$

where $n = N - 1$. Then the posterior density of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ given

$$\bar{\mathbf{x}} = \frac{1}{N}\sum_{\alpha=1}^{N} \mathbf{x}_\alpha \quad \text{and} \quad \mathbf{S} = \frac{1}{N-1}\sum_{\alpha=1}^{N} (\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})^\top$$

is

$$n\left(\boldsymbol{\mu} \;\middle|\; \frac{N\bar{\mathbf{x}} + K\boldsymbol{\nu}}{N+K}, \frac{\boldsymbol{\Sigma}}{N+K}\right) \cdot w^{-1}\left(\boldsymbol{\Sigma} \mid \boldsymbol{\Psi} + n\mathbf{S} + \frac{NK(\bar{\mathbf{x}} - \boldsymbol{\nu})(\bar{\mathbf{x}} - \boldsymbol{\nu})^\top}{N+K}, N + m\right).$$