Multivariate Statistical Analysis

Lecture 13

Fudan University

luoluo@fudan.edu.cn

Outline

Outline

The Inverted Wishart Distribution

If $\mathbf{A} \sim \mathcal{W}_p(\mathbf{\Sigma}, m)$, then $\mathbf{B} = \mathbf{A}^{-1}$ has the inverted Wishart distribution with m degrees of freedom and scale parameter $\mathbf{\Psi} = \mathbf{\Sigma}^{-1}$, written as

$$\mathbf{B} \sim \mathcal{W}_p^{-1}(\mathbf{\Psi}, m).$$

The density function of $\mathbf{B}$ is

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

where

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

Quiz

Define $\bar{\mathbb{S}}^p \to \mathbb{R}^{p\times p}$ as

$$\mathbf{F}(\mathbf{X}) = \mathbf{X}^{-1},$$

where $\bar{\mathbb{S}}^p = \{\mathbf{X} \in \mathbb{R}^{p\times p} : \mathbf{X} = \mathbf{X}^\top \text{ and } \mathbf{X} \text{ is non-singular}\}$.

What is the determinant of Jacobian of $\mathbf{F}(\mathbf{X})$?

Quiz

Let $\mathbf{x} \sim \mathcal{N}_p(\boldsymbol{\mu}, \mathbf{\Sigma})$. We define $\mathbf{\Psi} = \mathbf{\Sigma}^{-1}$ as the precision matrix.

1. It is well-known that

$$\sigma_{ij} = 0 \text{ if and only if } x_i \text{ and } x_j \text{ are independent.}$$

2. What is the meaning of $\psi_{ij} = 0$?

Outline

The Conjugate Prior for the Covariance

Theorem

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

$$\mathcal{W}^{-1}(\mathbf{A} + \mathbf{\Psi}, n + m).$$

Let each of $\mathbf{x}_1, \ldots, \mathbf{x}_N$ has distribution $\mathcal{N}_p(\mathbf{0}, \mathbf{\Sigma})$ independently and $n = N - 1$, then the sample covariance

$$\mathbf{S} = \frac{1}{n} \sum_{i=1}^{n} (\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})^\top \sim \mathcal{W}_p(\mathbf{\Sigma}, n).$$

If $\mathbf{\Sigma} \sim \mathcal{W}_p^{-1}(\mathbf{\Psi}, m)$, then we have

$$\mathbf{\Sigma} \,|\, \mathbf{S} \sim \mathcal{W}^{-1}(n\mathbf{S} + \mathbf{\Psi}, n + m).$$

Theorem

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

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

respectively, 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

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

Outline

The Characteristic Function of Wishart Distribution

Theorem

Let $\mathbf{A} \sim \mathcal{W}_p(\mathbf{\Sigma}, n)$, *then the characteristic function of*

$$a_{11}, a_{22}, \ldots, a_{pp}, 2a_{12}, \ldots, 2a_{p-1,p},$$

is is given by

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

where $\mathbf{\Theta} \in \mathbb{R}^{p\times p}$ *is symmetric.*

Outline

Matrix F-Distribution

The density of F-distribution with m and n degrees of freedom in univariate case is

$$\frac{1}{B\left(\frac{m}{2},\frac{n}{2}\right)}\left(\frac{m}{n}\right)^{\frac{n}{2}} u^{\frac{n}{2}-1}\left(1+\frac{m}{n}\cdot u\right)^{-\frac{m+n}{2}},$$

where

$$B\left(\frac{m}{2},\frac{n}{2}\right)=\frac{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})}{\Gamma\left(\frac{m+n}{2}\right)}.$$

How to generalized it to multivariate case?

Matrix F-Distribution

Let $\mathbf{A} \sim \mathcal{W}_p(\mathbf{I}, n)$ and $\mathbf{B} \sim \mathcal{W}_p(\mathbf{\Sigma}^{-1}, m)$ be independent, then

$$\mathbf{U} = \mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2},$$

has matrix F-distribution with n and m degrees of freedom.

Its density function is

$$f(\mathbf{U}) = \frac{\Gamma_p\left(\frac{m+n}{2}\right)(\det(\mathbf{\Sigma}))^{-\frac{n}{2}}}{\Gamma_p(\frac{m}{2})\Gamma_p(\frac{n}{2})} \cdot (\det(\mathbf{U}))^{\frac{n-p-1}{2}} \left(\det(\mathbf{I} + \mathbf{U}\mathbf{\Sigma}^{-1})\right)^{-\frac{m+n}{2}}.$$

It is natural to define the multivariate Beta function as

$$B_p(a, b) = \frac{\Gamma_p(a)\Gamma_p(b)}{\Gamma_p(a+b)}.$$

Matrix Beta Distribution

The density of Beta distribution with parameters $m/2$ and $n/2$ in univariate case is

$$f(w) = \frac{1}{B(\frac{m}{2}, \frac{n}{2})} \cdot w^{\frac{n}{2}-1}(1-w)^{\frac{m}{2}-1},$$

where

$$B\left(\frac{m}{2}, \frac{n}{2}\right) = \frac{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})}{\Gamma\left(\frac{m+n}{2}\right)}.$$

How to generalized it to multivariate case?

Matrix Beta Distribution

Let $\mathbf{A} \sim \mathcal{W}_p(\mathbf{\Sigma}, n)$ and $\mathbf{B} \sim \mathcal{W}_p(\mathbf{\Sigma}, m)$ be independent, then

$$\mathbf{W} = (\mathbf{A} + \mathbf{B})^{-1/2}\mathbf{A}(\mathbf{A} + \mathbf{B})^{-1/2}$$

has matrix Beta distribution with parameters $n/2$ and $m/2$ if $\mathbf{0} \prec \mathbf{W} \prec \mathbf{I}$ and 0 elsewhere.

Its density function is

$$f(\mathbf{W}) = \frac{1}{B_p(\frac{n}{2}, \frac{m}{2})} \cdot (\det(\mathbf{W}))^{\frac{n-p-1}{2}} (\det(\mathbf{I} - \mathbf{W}))^{\frac{m-p-1}{2}},$$

which does not depend on $\mathbf{\Sigma}$.

Outline

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$ constitute a sample from $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $N > p$.

We shall derive T^2-Statistic

$$T^2 = N(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)^\top \mathbf{S}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)$$

from likelihood ratio criterion

$$\lambda = \frac{\max_{\boldsymbol{\Sigma} \in \mathbb{S}_p^{++}} L(\boldsymbol{\mu}_0, \boldsymbol{\Sigma})}{\max_{\boldsymbol{\mu} \in \mathbb{R}^p, \boldsymbol{\Sigma} \in \mathbb{S}_p^{++}} L(\boldsymbol{\mu}, \boldsymbol{\Sigma})}.$$

We have

$$\lambda^{\frac{2}{N}} = \frac{1}{1 + T^2/(N-1)},$$

where

$$T^2 = N(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)^\top \mathbf{S}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0), \qquad \bar{\mathbf{x}} = \frac{1}{N}\sum_{\alpha=1}^{N} \mathbf{x}_\alpha$$

and

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

The condition $\lambda^{2/N} > c$ for some $c \in (0, 1)$ is equivalent to

$$T^2 < \frac{(N-1)(1-c)}{c}.$$