STAT 3690 Lecture Note

Part III: Multivariate normal distribution

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca, zhiyanggeezhou.github.io)

2023/Mar/11 22:07:26

Multivariate normal (MVN) distribution (J&W Sec 4.2)

Definition

- Standard MVN
  - $\boldsymbol{Z} = [Z_1, \ldots, Z_p]^\top \sim \mathrm{MVN}_p(\mathbf{0}, \mathbf{I}) \Leftrightarrow Z_1, \ldots, Z_p \overset{\text{iid}}{\sim} \mathcal{N}(0, 1)$
  - pdf
$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = (2\pi)^{-p/2} \exp(-\boldsymbol{z}^\top \boldsymbol{z}/2) \cdot \mathbf{1}_{\mathbb{R}^p}(\boldsymbol{z})$$
- General MVN
  - $\boldsymbol{X} = [X_1, \ldots, X_p]^\top \sim \mathrm{MVN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Leftrightarrow$ there exists $\boldsymbol{\mu} \in \mathbb{R}^p$, $\mathbf{A} \in \mathbb{R}^{p \times p}$ and $\boldsymbol{Z} \sim \mathrm{MVN}_p(\mathbf{0}, \mathbf{I})$ such that $\boldsymbol{X} = \mathbf{A}\boldsymbol{Z} + \boldsymbol{\mu}$ and $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^\top$
    - Limited to non-degenerate cases, i.e., invertible $\mathbf{A}$ ($\Leftrightarrow \boldsymbol{\Sigma} > 0$)
  - pdf
$$f_{\boldsymbol{X}}(\boldsymbol{x}) = (2\pi)^{-p/2} (\det \boldsymbol{\Sigma})^{-1/2} \exp\{-(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu})/2\} \cdot \mathbf{1}_{\mathbb{R}^p}(\boldsymbol{x})$$

- Exercise 3.1: Density of $\mathrm{MVN}_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ evaluated at $(4, 7)$, where
$$\boldsymbol{\mu} = [3, 6]^\top, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 10 & 2 \\ 2 & 5 \end{bmatrix}.$$

```r
options(digits = 4)
(Mu = matrix(c(3, 6), ncol = 1, nrow = 2))
(Sigma = matrix(c(10, 2 ,2, 5), ncol = 2, nrow = 2))
(x = c(4,7))
# Method 1: following the pdf
(2*pi)^{-length(Mu)/2}*det(Sigma)^{-.5}*exp(-drop(t(x-Mu)%*%solve(Sigma)%*%(x-Mu))/2)
# Method 2: via mvtnorm::dmvnorm()
mvtnorm::dmvnorm(x, mean = Mu, sigma = Sigma)
```

Properties of MVN

- $\boldsymbol{X}$ is of MVN $\Leftrightarrow a^\top \boldsymbol{X}$ is normally distributed for ALL non-zero $a \in \mathbb{R}^p$.
  - Warning: the marginal normality do not imply the joint normality.
- If $\boldsymbol{X} \sim \mathrm{MVN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then $\mathbf{A}\boldsymbol{X} + \boldsymbol{b} \sim \mathrm{MVN}_q(\mathbf{A}\boldsymbol{\mu} + \boldsymbol{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top)$ for $\mathbf{A} \in \mathbb{R}^{q \times p}$ of full-row-rank. Specifically, if $\boldsymbol{X} \sim \mathrm{MVN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then
  - $\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{X} - \boldsymbol{\mu}) \sim \mathrm{MVN}_p(\mathbf{0}, \mathbf{I})$ AND

  - (Stochastic representation of MVN) there is $\boldsymbol{Z} \sim \text{MVN}_p(\mathbf{0}, \mathbf{I})$ such that $\boldsymbol{X} = \boldsymbol{\Sigma}^{1/2}\boldsymbol{Z} + \boldsymbol{\mu}$.
- $(\boldsymbol{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{X} - \boldsymbol{\mu}) \sim \chi^2(p)$ if $\boldsymbol{X} \sim \text{MVN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

- Exercise 3.2: Generate six iid samples following bivariate normal $\text{MVN}_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with

$$\boldsymbol{\mu} = [3, 6]^\top, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 10 & 2 \\ 2 & 5 \end{bmatrix}.$$

```r
options(digits = 4)
set.seed(1)
(Mu = matrix(c(3, 6), ncol = 1, nrow = 2))
(Sigma = matrix(c(10, 2 ,2, 5), ncol = 2, nrow = 2))
n = 10
# Method 1: following the stochastic representation
sample1 = matrix(0, nrow = n, ncol = length(Mu))
for (i in 1:n) {
  sample1[i, ] = t(
    expm::sqrtm(Sigma) %*%
    matrix(rnorm(length(Mu)), nrow = length(Mu), ncol = 1) +
    Mu
 )
}
sample1
# Method 2: via MASS::mvrnorm()
(sample2 = MASS::mvrnorm(n, Mu, Sigma))
```

- Exercise 3.3: Suppose $X_1 \sim \mathcal{N}(0, 1)$. In the following two cases, verify that $X_2 \sim \mathcal{N}(0, 1)$ as well. Does $\boldsymbol{X} = [X_1, X_2]^\top$ follow an MVN in both cases?
  a. $X_2 = -X_1$;
  b. $X_2 = (2Y - 1)X_1$, where $Y \sim \text{Ber}(p)$ and $Y \perp\!\!\!\perp X_1$.

```r
options(digits = 4)
set.seed(1)
xsize = 1e4L
x1 = rnorm(xsize)
# case a
x2 = -x1
plot3D::hist3D(z=table(cut(x1, 100), cut(x2, 100)), border = "black") # 3d histogram of (x1, x2)
plot3D::image2D(z=table(cut(x1, 100), cut(x2, 100)), border = "black") # plot the support of joint pdf
# case b
Y = rbinom(n = xsize, 1, .3)
x2 = (2 * Y - 1) * x1
plot3D::hist3D(z=table(cut(x1, 100), cut(x2, 100)), border = "black") # 3d histogram of (x1, x2)
plot3D::image2D(z=table(cut(x1, 100), cut(x2, 100)), border = "black") # plot the support of joint pdf
```

Marginal and conditional MVN

- If $\boldsymbol{X} \sim \mathrm{MVN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

  with
  - random p_i-vector $\boldsymbol{X}_i$, $i = 1, 2$,
  - p_i-vector $\boldsymbol{\mu}_i$, $i = 1, 2$,
  - $p_i \times p_i$ matrix $\boldsymbol{\Sigma}_{ii} > 0$, $i = 1, 2$,
- then
  - (Marginals of MVN are still MVN) $\boldsymbol{X}_i \sim \mathrm{MVN}_{p_i}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_{ii})$
  - $\boldsymbol{X}_i \mid \boldsymbol{X}_j = \boldsymbol{x}_j \sim \mathrm{MVN}_{p_i}(\boldsymbol{\mu}_{i|j}, \boldsymbol{\Sigma}_{i|j})$
    - $\boldsymbol{\mu}_{i|j} = \boldsymbol{\mu}_i + \boldsymbol{\Sigma}_{ij}\boldsymbol{\Sigma}_{jj}^{-1}(\boldsymbol{x}_j - \boldsymbol{\mu}_j)$
    - $\boldsymbol{\Sigma}_{i|j} = \boldsymbol{\Sigma}_{ii} - \boldsymbol{\Sigma}_{ij}\boldsymbol{\Sigma}_{jj}^{-1}\boldsymbol{\Sigma}_{ji}$
  - $\boldsymbol{X}_1 \perp\!\!\!\perp \boldsymbol{X}_2 \Leftrightarrow \boldsymbol{\Sigma}_{12} = \boldsymbol{0}$
    - Warning: the prerequisite for this equivalence is the joint normal of $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$.

- Exercise 3.4: The argument $\boldsymbol{X}_1 \perp\!\!\!\perp \boldsymbol{X}_2 \Leftrightarrow \boldsymbol{\Sigma}_{12} = 0$ is based on $[\boldsymbol{X}_1^\top, \boldsymbol{X}_2^\top]^\top \sim \mathrm{MVN}$. That is, if $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are both MVN BUT they are not jointly normal, the zero $\boldsymbol{\Sigma}_{12}$ doesn't suffice for the independence between $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$. Recall the case b. in Exercise 3.3: $X_1 \sim \mathcal{N}(0, 1)$ and $X_2 = (2Y - 1)X_1$, where $Y \sim \mathrm{Ber}(p)$ and $Y \perp\!\!\!\perp X_1$. Verify that X_1 and X_2 are not independent of each other. (Hint: assume the independence and then check the support of $[X_1, X_2]^\top$.)

Hypothesis testing

- Is it a squirrel?

Figure 1: Squirrel (Photograph by the Lacoste Garden Centre)

Figure 2: Flying Squirrel (Photograph by Joel Sartore)

Figure 3: Flying Squirrel (Photograph by Alex Badyaev)

- Null and alternative hypotheses, say H_0 and H_1, resp.
- Name of testing method
- Test statistic (varying with the testing method) and corresponding level α rejection region R_α
  - Pr(test statistic $\in R_\alpha \mid H_0) \leq \alpha$
  - Reject H_0 if the value of test statistic $\in R_\alpha$
    - Type I error: H_0 is incorrectly rejected; i.e., H_0 is correct but rejected

        * Type II error: H_0 is incorrectly accepted i.e., H_0 is wrong but NOT rejected

- p-value: a special test statistic with a default level α rejection region $[0, \alpha]$
- Necessary components in reporting a testing result
    1. Hypotheses
    2. Name of approach
    3. Level α
    4. (Value of test statistic AND rejection region) OR p-value
    5. Conclusion: e.g., at the α level, we reject/do not reject H_0, i.e., we believe that…

Checking/testing the normality (J&W Sec 4.6)

- Checkcing the univariate normality
    - Normal Q-Q plot
        * qqnorm(); car::qqPlot()
    - Univariate normality test
        * shapiro.test(); nortest::ad.test(); MVN::mvn()
- Checkcing the multivariate normality
    - χ^2 Q-Q plot
        * $D_i^2 = (\boldsymbol{X}_i - \bar{\boldsymbol{X}})^\top \mathbf{S}^{-1}(\boldsymbol{X}_i - \bar{\boldsymbol{X}}) \approx \chi^2(p)$ if $\boldsymbol{X}_i \overset{\text{iid}}{\sim} \text{MVN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$
        * qqplot(); car::qqPlot()
    - Multivariate normality test
        * MVN::mvn()

```
options(digits = 4)
library(datasets)
data(iris)
head(iris)
(iris_setosa = iris[iris$Species=='setosa', 1:3])
p = ncol(iris_setosa)
n = nrow(iris_setosa)

# Marginal normal Q-Q plot
car::qqPlot(rnorm(n), id = F)
car::qqPlot(iris_setosa[,1], id = F)
car::qqPlot(iris_setosa[,2], id = F)
car::qqPlot(iris_setosa[,3], id = F)

# Univariate normality test
## Shapiro-Wilk Normality Test
shapiro.test(rnorm(n))
shapiro.test(iris_setosa[,1])
shapiro.test(iris_setosa[,2])
shapiro.test(iris_setosa[,3])
## Anderson-Darling test for normality
nortest::ad.test(iris_setosa[,1])
nortest::ad.test(iris_setosa[,2])
nortest::ad.test(iris_setosa[,3])
## via MVN::mvn()
MVN::mvn(
  iris_setosa,
  univariateTest = "AD" # "SW"/"CVM"/"Lillie"/"SF"/"AD"
```

```
)$univariateNormality

# chi^2 Q-Q plot
d_square = diag(
  as.matrix(sweep(iris_setosa, 2, colMeans(iris_setosa))) %*%
    solve(var(iris_setosa)) %*%
    t(as.matrix(sweep(iris_setosa, 2, colMeans(iris_setosa))))
)
car::qqPlot(d_square, dist="chisq", df = p, id = F)
MVN::mvn(
  iris_setosa,
  multivariatePlot = "qq"
)

# Multivariate normality test
MVN::mvn(
  iris_setosa,
  mvnTest = "dh" # "mardia"/"hz"/"royston"/"dh"/"energy"
)$multivariateNormality
```

Detecting outliers (J&W Sec 4.7)

- Scatter plot of standardized values
- Checking the points farthest from the origin in χ^2 Q-Q plot

Improving normality (J&W Sec 4.8)

- (Original) Box-Cox (power) transformation: transform positive x into

$$X^* = \begin{cases} (X^\lambda - 1)/\lambda & \lambda \neq 0 \\ \ln(X) & \lambda = 0 \end{cases}$$

  with λ selected with certain criterion
  - If $X \leq 0$, change it to be positive first.
  - See J. Tukey (1977). *Exploratory Data Analysis*. Boston: Addison-Wesley.

```
library(datasets)
data(iris)
head(iris)
iris_setosa = iris[iris$Species=='setosa', 1:3]

iris_setosa = iris_setosa - min(iris_setosa) + 1 # make sure all the entries are positive

(lambda = EnvStats::boxcox(iris_setosa[,2], optimize=T)$lambda)
if (lambda != 0){
  df_new = (iris_setosa[,2]^lambda-1)/lambda
}else df_new = log(iris_setosa[,2])

car::qqPlot(df_new, id = F)
shapiro.test(df_new)
nortest::ad.test(df_new)
```

- Multivariate Box-Cox transformation

```r
(lambdas = MVN::mvn(
  iris_setosa,
  bc = T,
  bcType = 'optimal'
)$BoxCoxPowerTransformation)
iris_setosa_new = iris_setosa
for (i in 1:length(lambdas)){
  if (lambdas[i] != 0){
    iris_setosa_new[,i] = (iris_setosa[,i]^lambdas[i]-1)/lambdas[i]
  }else iris_setosa_new[,i] = log(iris_setosa[,i])
}
MVN::mvn(
  iris_setosa_new,
  mvnTest = "energy" # "mardia"/"hz"/"royston"/"dh"/"energy"
)$multivariateNormality
```

Maximum likelihood (ML) estimation of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ (J&W Sec 4.3)

- Sample: $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n \overset{\text{iid}}{\sim} \text{MVN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $n > p$
- Likelihood function

$$\begin{aligned} L(\boldsymbol{\mu}, \boldsymbol{\Sigma}) =& \prod_{i=1}^{n} \left[\frac{1}{\sqrt{(2\pi)^p \det(\boldsymbol{\Sigma})}} \exp\Big\{ -\frac{1}{2} (\boldsymbol{X}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{X}_i - \boldsymbol{\mu}) \Big\} \right] \\ =& \frac{1}{\sqrt{(2\pi)^{np} \{\det(\boldsymbol{\Sigma})\}^n}} \exp\Big\{ -\frac{1}{2} \sum_{i=1}^{n} (\boldsymbol{X}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{X}_i - \boldsymbol{\mu}) \Big\} \end{aligned}$$

- Log likelihood

$$\ell(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \ln L(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = -\frac{np}{2} \ln(2\pi) - \frac{n}{2} \ln\{\det(\boldsymbol{\Sigma})\} - \frac{1}{2} \sum_{i=1}^{n} (\boldsymbol{X}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{X}_i - \boldsymbol{\mu})$$

- ML estimator

$$(\hat{\boldsymbol{\mu}}_{\text{ML}}, \widehat{\boldsymbol{\Sigma}}_{\text{ML}}) = \arg \max_{\boldsymbol{\mu} \in \mathbb{R}^p, \boldsymbol{\Sigma} \in \mathbb{R}^{p \times p}, \boldsymbol{\Sigma} > 0} \ell(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = (\bar{\boldsymbol{X}}, \frac{n-1}{n} \mathbf{S})$$

  - Consistency: $(\hat{\boldsymbol{\mu}}_{\text{ML}}, \widehat{\boldsymbol{\Sigma}}_{\text{ML}})$ approaches $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ (in certain sense) as $n \to \infty$
  - Efficiency: the covariance matrix of $(\hat{\boldsymbol{\mu}}_{\text{ML}}, \widehat{\boldsymbol{\Sigma}}_{\text{ML}})$ is approximately optimal (in certain sense) as $n \to \infty$
  - Invariance: for any function g, the ML estimator of $g(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is $g(\hat{\boldsymbol{\mu}}_{\text{ML}}, \widehat{\boldsymbol{\Sigma}}_{\text{ML}})$.

Sampling distributions of $\bar{\boldsymbol{X}}$ and S (J&W Sec 4.4)

- Recall the univariate case: if $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$, then
  - $s^2 \perp\!\!\!\perp \bar{X}$
    - Sample variance $s^2 = (n-1)^{-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$
  - $\sqrt{n}(\bar{X} - \mu)/\sigma \sim \mathcal{N}(0, 1)$
  - $(n-1)s^2/\sigma^2 \sim \chi^2(n-1)$
  - $\sqrt{n}(\bar{X} - \mu)/s \sim t(n-1)$
- The multivariate case: if $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n \overset{\text{iid}}{\sim} \text{MVN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $n > p$, then
  - $\mathbf{S} \perp\!\!\!\perp \bar{\boldsymbol{X}}$, i.e., $\widehat{\boldsymbol{\Sigma}}_{\text{ML}} \perp\!\!\!\perp \hat{\boldsymbol{\mu}}_{\text{ML}}$
  - $\sqrt{n} \boldsymbol{\Sigma}^{-1/2} (\bar{\boldsymbol{X}} - \boldsymbol{\mu}) \sim \text{MVN}_p(\mathbf{0}, \mathbf{I})$

- $(n-1)\mathbf{S} = n\widehat{\mathbf{\Sigma}}_{\mathrm{ML}} \sim W_p(\mathbf{\Sigma}, n-1)$
- $n(\bar{\boldsymbol{X}} - \boldsymbol{\mu})^\top \mathbf{S}^{-1}(\bar{\boldsymbol{X}} - \boldsymbol{\mu}) \sim$ Hotelling's $T^2(p, n-1)$

- Wishart distribution
  - $W_p(\mathbf{\Sigma}, n)$ is the distribution of $\sum_{i=1}^n \boldsymbol{Y}_i \boldsymbol{Y}_i^\top$ with $\boldsymbol{Y}_1, \dots, \boldsymbol{Y}_n \overset{\text{iid}}{\sim} \mathrm{MVN}_p(\mathbf{0}, \mathbf{\Sigma})$
    - A generalization of χ^2-distribution: $W_p(\mathbf{\Sigma}, n) = \chi^2(n)$ if $p = \mathbf{\Sigma} = 1$
  - Propoties
    - $\mathbf{A}\mathbf{A}^\top > 0$ and $\mathbf{W} \sim W_p(\mathbf{\Sigma}, n) \Rightarrow \mathbf{A}\mathbf{W}\mathbf{A}^\top \sim W_p(\mathbf{A}\mathbf{\Sigma}\mathbf{A}^\top, n)$
    - $\mathbf{W}_i \overset{\text{iid}}{\sim} W_p(\mathbf{\Sigma}, n_i) \Rightarrow \mathbf{W}_1 + \mathbf{W}_2 \sim W_p(\mathbf{\Sigma}, n_1 + n_2)$
    - $\mathbf{W}_1 \perp\!\!\!\perp \mathbf{W}_2$, $\mathbf{W}_1 + \mathbf{W}_2 \sim W_p(\mathbf{\Sigma}, n)$ and $\mathbf{W}_1 \sim W_p(\mathbf{\Sigma}, n_1) \Rightarrow \mathbf{W}_2 \sim W_p(\mathbf{\Sigma}, n - n_1)$
    - $\mathbf{W} \sim W_p(\mathbf{\Sigma}, n)$ and $\boldsymbol{a} \in \mathbb{R}^p \Rightarrow$
    $$\frac{\boldsymbol{a}^\top \mathbf{W} \boldsymbol{a}}{\boldsymbol{a}^\top \mathbf{\Sigma} \boldsymbol{a}} \sim \chi^2(n)$$
    - $\mathbf{W} \sim W_p(\mathbf{\Sigma}, n)$, $\boldsymbol{a} \in \mathbb{R}^p$ and $n \geq p \Rightarrow$
    $$\frac{\boldsymbol{a}^\top \mathbf{\Sigma}^{-1} \boldsymbol{a}}{\boldsymbol{a}^\top \mathbf{W}^{-1} \boldsymbol{a}} \sim \chi^2(n - p + 1)$$
    - $\mathbf{W} \sim W_p(\mathbf{\Sigma}, n) \Rightarrow$
    $$\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{W}) \sim \chi^2(np)$$

- Hotelling's T^2 distribution
  - A generalization of (Student's) t-distribution
  - If $\boldsymbol{X} \sim \mathrm{MVN}_p(\mathbf{0}, \mathbf{I})$ and $\mathbf{W} \sim W_p(\mathbf{I}, n)$, then
  $$\boldsymbol{X}^\top \mathbf{W}^{-1} \boldsymbol{X} \sim T^2(p, n)$$
  - $Y \sim T^2(p, n) \Leftrightarrow \frac{n-p+1}{np} Y \sim F(p, n-p+1)$

- Wilk's lambda distribution
  - Wilks's lambda is to Hotelling's T^2 as F distribution is to Student's t in univariate statistics.
  - Given independent $\mathbf{W}_1 \sim W_p(\mathbf{\Sigma}, n_1)$ and $\mathbf{W}_2 \sim W_p(\mathbf{\Sigma}, n_2)$ with $n_1 \geq p$,
  $$\Lambda = \frac{\det(\mathbf{W}_1)}{\det(\mathbf{W}_1 + \mathbf{W}_2)} = \frac{1}{\det(\mathbf{I} + \mathbf{W}_1^{-1}\mathbf{W}_2)} \sim \Lambda(p, n_1, n_2)$$
    - Resort to an approximation in computation: $\{(p - n_2 + 1)/2 - n_1\} \ln \Lambda(p, n_1, n_2) \approx \chi^2(n_2 p)$