Elsevier LaTeX template[☆]

Elsevier[1]

Radarweg 29, Amsterdam

Elsevier Inc[a,b], Global Customer Service[b,]*

[a]1600 John F Kennedy Boulevard, Philadelphia
[b]360 Park Avenue South, New York

Abstract

This template helps you to create a properly formatted LaTeX manuscript.

Keywords: `elsarticle.cls`, LaTeX, Elsevier, template

2010 MSC: 00-01, 99-00

1. GLRT

Suppose $\{X_{i1}, \ldots, X_{in_i}\}$ are i.i.d. distributed as $N(\mu_i, \Sigma)$ for $1 \leq i \leq K$. Let $\mathbf{X}_i = (X_{i1}, \ldots, X_{in_i})$ for $i = 1, \ldots, k$. The k samples are independent. μ_i, $i = 1 \ldots, k$ and $\Sigma > 0$ are unknown. An interesting problem in multivariate analysis is to test the hypotheses

$$H : \mu_1 = \mu_2 = \cdots = \mu_k \quad v.s. \quad K : \mu_i \neq \mu_j \text{ for some } i \neq j. \tag{1}$$

Let $\mathbf{Z} = (X_1, \ldots, X_k)$.

$$f(Z; \mu_1, \ldots, \mu_k, \Sigma) = \prod_{i=1}^{k} \Big[(2\pi)^{-n_i p/2} |\Sigma|^{-n_i/2} \exp(-\frac{1}{2} \mathrm{tr}\Sigma^{-1} \sum_{j=1}^{n_i} (x_{ij} - \mu_i)(x_{ij} - \mu_i)^T) \Big].$$

Assume $n = \sum_{i=1}^{p} n_i < p$. Let $a \in \mathbb{R}^p$ be a vector satisfying $a^T a = 1$. Then

$$f_a(a^T Z; \mu_1, \ldots, \mu_k, \Sigma) = (2\pi)^{-n/2} |a^T \Sigma a|^{-n/2} \exp\Big(-\frac{1}{2a^T \Sigma a} \sum_{i=1}^{k} \sum_{j=1}^{n_i} (a^T x_{ij} - a^T \mu_i)^2 \Big)$$

[☆]Fully documented templates are available in the elsarticle package on CTAN.

[*]Corresponding author
Email address: `support@elsevier.com` (Global Customer Service)
URL: `www.elsevier.com` (Elsevier Inc)

[1]Since 1880.

$$\max_{\mu_1,\dots,\mu_k,\Sigma} f_a(a^T Z, \mu_1, \dots, \mu_k, \Sigma) = (2\pi)^{-n/2} \big(\sum_{i=1}^{k} \sum_{j=1}^{n_i} (a^T x_{ij} - a^T \bar{\mathbf{X}}_i)^2 \big)^{-n/2} e^{-n/2} \tag{2}$$

Let $S_i = \sum_{j=1}^{n_i} (x_{ij} - \bar{\mathbf{X}}_i)(x_{ij} - \bar{\mathbf{X}}_i)^T$ and $S = \sum_{i=1}^{k} S_i$.

Under H, we have

$$\max_{\mu,\Sigma} f_a(a^T Z, \mu, \dots, \mu, \Sigma) = (2\pi)^{-n/2} \big(\sum_{i=1}^{k} \sum_{j=1}^{n_i} (a^T x_{ij} - a^T \bar{\mathbf{X}})^2 \big)^{-n/2} e^{-n/2} \tag{3}$$

The generalized likelihood ratio test statistic is defined as

$$T(Z) = \max_{a^T a = 1, a^T S a = 0} a^T \sum_{i=1}^{k} n_i (\bar{\mathbf{X}}_i - \bar{\mathbf{X}})(\bar{\mathbf{X}}_i - \bar{\mathbf{X}})^T a \tag{4}$$

Let $J = \operatorname{diag}(n_1^{-1/2} \mathbf{1}_{n_1}, \dots, n_k^{-1/2} \mathbf{1}_{n_k})$. Then $S = Z(I_n - JJ^T)Z^T$ and

$$\sum_{i=1}^{k} n_i (\bar{\mathbf{X}}_i - \bar{\mathbf{X}})(\bar{\mathbf{X}}_i - \bar{\mathbf{X}})^T = Z(JJ^T - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T) Z^T. \tag{5}$$

The matrix $I_n - JJ^T$, $JJ^T - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$ and $\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$ are all projection matrix and pairwise orthogonal with rank $n-k$, $k-1$ and 1.

Let $\tilde{J}$ be a $n \times (n-k)$ matrix satisfied $\tilde{J}\tilde{J}^T = I - JJ^T$. Then $S = Z\tilde{J}\tilde{J}^T Z^T$ and Note that

$$Z(JJ^T - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T)Z^T = ZJ(I_k - \frac{1}{n} J^T \mathbf{1}_n \mathbf{1}_n^T J) J^T Z^T.$$

Note that $I_k - \frac{1}{n} J^T \mathbf{1}_n \mathbf{1}_n^T J$ is a projection matrix with rank $k-1$. Let C be a $k \times (k-1)$ matrix satisfied $CC^T = I_k - \frac{1}{n} J^T \mathbf{1}_n \mathbf{1}_n^T J$.

In Proposition 1, letting $A = Z\tilde{J}$ and $B = ZJCC^T J^T Z^T$ yields

$$\begin{aligned} T(Z) &= \lambda_{max}\big((I_p - Z\tilde{J}(\tilde{J}^T Z^T Z \tilde{J})^{-1} \tilde{J}^T Z^T) ZJCC^T J^T Z^T (I_p - Z\tilde{J}(\tilde{J}^T Z^T Z\tilde{J})^{-1} \tilde{J}^T Z^T) \big) \\ &= \lambda_{max}\big(C^T J^T Z^T (I_p - Z\tilde{J}(\tilde{J}^T Z^T Z\tilde{J})^{-1} \tilde{J}^T Z^T) ZJC \big). \end{aligned}$$

Note that

$$\begin{aligned} &\Big(\begin{pmatrix} J^T \\ \tilde{J}^T \end{pmatrix} Z^T Z \begin{pmatrix} J & \tilde{J} \end{pmatrix} \Big)^{-1} \\ &= \begin{pmatrix} J^T Z^T Z J & J^T Z^T Z \tilde{J} \\ \tilde{J}^T Z^T Z J & \tilde{J}^T Z^T Z \tilde{J} \end{pmatrix}^{-1} = \begin{pmatrix} J^T (Z^T Z)^{-1} J & J^T (Z^T Z)^{-1} \tilde{J} \\ \tilde{J}^T (Z^T Z)^{-1} J & \tilde{J}^T (Z^T Z)^{-1} \tilde{J} \end{pmatrix}. \end{aligned} \tag{6}$$

It follows that

$$
\begin{aligned}
&\left(J^T (Z^T Z)^{-1} J\right)^{-1} \\
=&J^T Z^T Z J - J^T Z^T Z \tilde{J} (\tilde{J}^T Z^T Z \tilde{J})^{-1} \tilde{J}^T Z^T Z J \\
=&J^T Z^T (I_p - Z \tilde{J} (\tilde{J}^T Z^T Z \tilde{J})^{-1} \tilde{J}^T Z^T) Z J
\end{aligned}
\tag{7}
$$

It follows that

$$
T(Z) = \lambda_{\max}\Big(C^T \big(J^T (Z^T Z)^{-1} J\big)^{-1} C\Big) \tag{8}
$$

Proposition 1. *Suppose A is a $p \times r$ matrix with rank r and B is a $p \times p$ non-zero semi-definite matrix. Let $H_A = A(A^T A)^{-1} A^T$. Then*

$$
\max_{a^T a = 1, a^T A A^T a = 0} a^T B a = \lambda_{\max}\big((I_p - H_A) B (I_p - H_A)\big). \tag{9}
$$

Proof. Note that $a^T A A^T a = 0$ is equivalent to $A^T a = 0$ and is in turn equivalent to $H_A a = 0$. In this circumstance, $a = (I_p - H_A) a$. Then

$$
\begin{aligned}
\max_{a^T a = 1, a^T A A^T a = 0} a^T B a &= \max_{a^T a = 1, H_A a = 0} a^T B a \\
&= \max_{a^T a = 1, H_A a = 0} a^T (I_p - H_A) B (I_p - H_A) a.
\end{aligned}
\tag{10}
$$

It's obvious that (10) $\leq \lambda_{\max}\big((I - H_A) B (I - H_A)\big)$. On the other hand, let α_1 be one eigenvector corresponding to the largest eigenvalue of $(I - H_A) B (I - H_A)$. Note that the row of H_A are all eigenvetors of $(I - H_A) B (I - H_A)$ corresponding to eigenvalue 0. It follows that $H_A \alpha_1 = 0$. Now that α_1 satisfies the constraint of (10), (10) is maximized when $a = \alpha_1$.

□

2. Schott's method

$$
E = Z Z^T - \sum_{i=1}^{k} n_i \bar{X}_i \bar{X}_i^T.
$$

$$
H = \sum_{i=1}^{k} n_i \bar{X}_i \bar{X}_i^T - n \bar{X} \bar{X}^T.
$$

$$\operatorname{tr} E = \operatorname{tr} Z^T Z - \operatorname{tr} J^T Z^T Z J.$$

$$\operatorname{tr} H = \operatorname{tr} J^T Z^T Z J - \frac{1}{n} 1_n^T Z^T Z 1_n$$

$$T_{SC} = \frac{1}{\sqrt{n-1}} \left(\frac{1}{k-1} \operatorname{tr} H - \frac{1}{n-k} \operatorname{tr} E \right)$$

3. Theory

Let $Z\tilde{J} = U_{Z\tilde{J}} D_{Z\tilde{J}} V_{Z\tilde{J}}^T$ be the singular value decomposition of $Z\tilde{J}$. Let $H_{Z\tilde{J}} = U_{Z\tilde{J}} U_{Z\tilde{J}}^T$. $T(Z) = \lambda_{\max}(C^T J^T Z^T (I_p - H_{Z\tilde{J}}) Z J C)$.

ZJC is independent of $H_{Z\tilde{J}}$

$$\mathrm{E}(Z\tilde{J}) = O_{p\times(n-k)}$$

Let $\Sigma = U\Lambda U^T$. Then

$$Z\tilde{J} = U\Lambda^{1/2} G_1,$$

where G_1 is a $p \times (n-k)$ matrix with i.i.d. $N(0,1)$ entries.

$$\mathrm{E}(ZJ) = (\sqrt{n_1}\mu_1, \ldots, \sqrt{n_k}\mu_k) \stackrel{def}{=} \mu_f$$

And

$$ZJC = \mu_f C + U\Lambda_{1/2} G_2$$

where G_2 is a $p \times (k-1)$ matrix with i.i.d. $N(0,1)$ entries.

Then

$$\begin{aligned} T(Z) \sim \lambda_{\max}(& G_2^T \Lambda^{1/2} U^T (I_P - H_{Z\tilde{J}}) U \Lambda_{1/2} G_2 + \mu_f^T (I_p - H_{Z\tilde{J}}) \mu_f + \\ & \mu_f^T (I_p - H_{Z\tilde{J}}) U \Lambda_{1/2} G_2 + G_2^T \Lambda^{1/2} U^T (I_P - H_{Z\tilde{J}}) \mu_f) \end{aligned}$$

We have

$$G_2^T \Lambda^{1/2} U^T (I_P - H_{Z\tilde{J}}) U \Lambda_{1/2} G_2 \sim \sum_{i=1}^{p} \lambda_i (\Lambda^{1/2} U^T (I_p - H_{Z\tilde{J}}) U \Lambda^{1/2}) \xi_i \xi_i^T$$

where $\xi_i \overset{i.i.d.}{\sim} N(0, I_{k-1})$.

The eigenvalues of $\Lambda^{1/2}U^T(I_p - H_{Z\tilde{J}})U\Lambda^{1/2}$ equal to those of $(I_p - H_{Z\tilde{J}})U\Lambda U^T(I_p - H_{Z\tilde{J}})$.

Let $U = (U_1, U_2)$ and $\Sigma = \mathrm{diag}(\Lambda_1, \Lambda_2)$, where U_1 is $p \times r$ and Λ_1 is $r \times r$. Assume $cI_{p-r} \leq \Lambda_2 \leq CI_{p-r}$.

$$\begin{aligned}
&(I_p - H_{Z\tilde{J}})U\Lambda U^T(I_p - H_{Z\tilde{J}}) \\
=&(I_p - H_{Z\tilde{J}})U_1\Lambda_1 U_1^T(I_p - H_{Z\tilde{J}}) + (I_p - H_{Z\tilde{J}})U_2\Lambda_2 U_2^T(I_p - H_{Z\tilde{J}}) = R_1 + R_2
\end{aligned}$$

$$\begin{aligned}
&\lambda_{\max}\big((I_p - H_{Z\tilde{J}})U_1\Lambda_1 U_1^T(I_p - H_{Z\tilde{J}})\big) = \lambda_{\max}\big(\Lambda_1^{1/2}U_1^T(I_p - H_{Z\tilde{J}})U_1\Lambda_1^{1/2}\big) \\
&\leq \lambda_{\max}\big(\Lambda_1^{1/2}U_1^T(I_p - U_{Z\tilde{J}[,1:r]}U_{Z\tilde{J}[,1:r]}^T)U_1\Lambda_1^{1/2}\big) \leq \lambda_1\lambda_{\max}\big(U_1^T(I_p - U_{Z\tilde{J}[,1:r]}U_{Z\tilde{J}[,1:r]}^T)U_1\big) \\
&= \lambda_1\lambda_{\max}\big(I_r - U_1^T U_{Z\tilde{J}[,1:r]}U_{Z\tilde{J}[,1:r]}^T U_1\big)
\end{aligned}$$

We need to investigate the behavior of $U_{Z\tilde{J}}$.

$$G_1^T\Lambda G_1 = \tilde{J}^T Z^T Z\tilde{J} = V_{Z\tilde{J}}D_{Z\tilde{J}}^2 V_{Z\tilde{J}}^T$$

Note that

$$G_1^T\Lambda G_1 = G_{1[1:r,]}^T\Lambda_1 G_{1[1:r,]} + G_{1[(r+1):p,]}^T\Lambda_2 G_{1[(r+1):p,]}$$

For $i = 1, \ldots, r$,

$$\begin{aligned}
&\lambda_i(G_{1[1:r,]}^T\Lambda_1 G_{1[1:r,]}) \geq \lambda_i(G_{1[1:r,]}^T\,\mathrm{diag}(\lambda_i I_i, O_{(r-i)\times(r-i)})G_{1[1:r,]}) \\
=&\lambda_i\lambda_i(G_{1[1:i,]}G_{1[1:i,]}^T) = \lambda_i n(1 + o_P(1))
\end{aligned}$$

The last equality holds since $n^{-1}G_{1[1:i,]}G_{1[1:i,]}^T \xrightarrow{P} I_i$ by law of large numbers. On the other hand, for $i = 1, \ldots, r$,

$$\begin{aligned}
&\lambda_i(G_{1[1:r,]}^T\Lambda_1 G_{1[1:r,]}) \\
=&\lambda_i\Big(G_{1[1:r,]}^T\big(\mathrm{diag}(\lambda_1, \ldots, \lambda_{i-1}, O_{(r-i+1)\times(r-i+1)}) + \mathrm{diag}(O_{(i-1)\times(i-1)}, \lambda_i, \ldots, \lambda_r)\big)G_{1[1:r,]}\Big) \\
\leq&\lambda_1(G_{1[1:r,]}^T\,\mathrm{diag}(O_{(i-1)\times(i-1)}, \lambda_i, \ldots, \lambda_r)G_{1[1:r,]}) \leq \lambda_1(G_{1[1:r,]}^T\,\mathrm{diag}(O_{(i-1)\times(i-1)}, \lambda_i I_{r-i+1})G_{1[1:r,]}) \\
=&\lambda_i\lambda_1(G_{1[i:r,]}G_{1[i:r,]}^T) = \lambda_i n(1 + o_P(1))
\end{aligned}$$

where the first inequality holds by Weyl's inequality. It follows that $\lambda_i(G_{1[1:r,]}^T \Lambda_1 G_{1[1:r,]}) = \lambda_i n(1 + o_P(1))$ for $i = 1, \ldots, r$.

Note that $\lambda_{\max}(G_{1[(r+1):p,]}^T \Lambda_2 G_{1[(r+1):p,]}) \leq C\lambda_{\max}(G_{1[(r+1):p,]}^T G_{1[(r+1):p,]}) = O_P(p)$ by Bai-Yin's law, here we have assumed $n/p = O(1)$. Assume $p/(\lambda_r n) \to 0$, we can deduce that $D_{Z\tilde{J}[i,i]}^2 = \lambda_i(G_1^T \Lambda G_1) = \lambda_i n(1 + o_P(1))$, $i = 1, \ldots, r$.

Note that

$$U\Lambda^{1/2} G_1 G_1^T \Lambda^{1/2} U^T = U_{Z\tilde{J}} D_{Z\tilde{J}}^2 U_{Z\tilde{J}}^T$$

then

$$G_1 G_1^T = \Lambda^{-1/2} U^T U_{Z\tilde{J}} D_{Z\tilde{J}}^2 U_{Z\tilde{J}}^T U \Lambda^{-1/2}$$

and

$$\begin{aligned} G_{1[(r+1):p,]} G_{1[(r+1):p,]}^T &= \Lambda_2^{-1/2} U_{[,(r+1):p]}^T U_{Z\tilde{J}} D_{Z\tilde{J}}^2 U_{Z\tilde{J}}^T U_{[,(r+1):p]} \Lambda_2^{-1/2} \\ &\geq \Lambda_2^{-1/2} U_{[,(r+1):p]}^T U_{Z\tilde{J}[,1:r]} D_{Z\tilde{J}[1:r,1:r]}^2 U_{Z\tilde{J}[,1:r]}^T U_{[,(r+1):p]} \Lambda_2^{-1/2} \\ &\geq D_{Z\tilde{J}[r,r]}^2 \Lambda_2^{-1/2} U_{[,(r+1):p]}^T U_{Z\tilde{J}[,1:r]} U_{Z\tilde{J}[,1:r]}^T U_{[,(r+1):p]} \Lambda_2^{-1/2} \end{aligned}$$

Thus,

$$\lambda_{\max}(U_{[,(r+1):p]}^T U_{Z\tilde{J}[,1:r]} U_{Z\tilde{J}[,1:r]}^T U_{[,(r+1):p]}) \leq \frac{C}{D_{Z\tilde{J}[r,r]}^2} \lambda_1(G_{1[(r+1):p,]} G_{1[(r+1):p,]}^T) = O_P(\frac{p}{\lambda_r n})$$

Note that

$$\begin{aligned} &\lambda_{\max}(U_{[,(r+1):p]}^T U_{Z\tilde{J}[,1:r]} U_{Z\tilde{J}[,1:r]}^T U_{[,(r+1):p]}) = \lambda_{\max}(U_{Z\tilde{J}[,1:r]}^T U_{[,(r+1):p]} U_{[,(r+1):p]}^T U_{Z\tilde{J}[,1:r]}) \\ =&\lambda_{\max}(U_{Z\tilde{J}[,1:r]}^T (I_p - U_1 U_1^T) U_{Z\tilde{J}[,1:r]}) = \lambda_{\max}(I_r - U_{Z\tilde{J}[,1:r]}^T U_1 U_1^T U_{Z\tilde{J}[,1:r]}) \\ =&1 - \lambda_{\min}(U_{Z\tilde{J}[,1:r]}^T U_1 U_1^T U_{Z\tilde{J}[,1:r]}) = 1 - \lambda_{\min}(U_1^T U_{Z\tilde{J}[,1:r]} U_{Z\tilde{J}[,1:r]}^T U_1) \\ =&\lambda_{\max}(I_r - U_1^T U_{Z\tilde{J}[,1:r]} U_{Z\tilde{J}[,1:r]}^T U_1) \end{aligned}$$

Therefore $\lambda_{\max}(R_1) = O_P(\frac{\lambda_1 p}{\lambda_r n})$

References