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 $\{X_{i1}, \dots, X_{in_i}\}$ are i.i.d. distributed as $N(\mu_i, \Sigma)$ for $1 \leq i \leq K$. Let $\mathbf{X}_i = (X_{i1}, \dots, X_{in_i})$ for $i = 1, \dots, k$. The k samples are independent. μ_i, $i = 1 \dots, k$ 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 $\mathbf{Z} = (X_1, \dots, X_k)$.

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

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

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

[☆]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.

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

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

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

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 $J = \operatorname{diag}(n_1^{-1/2}\mathbf{1}_{n_1},\ldots,n_k^{-1/2}\mathbf{1}_{n_k})$. 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

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

It follows that

$$T(Z) = \lambda_{\max}\left(C^T \left(J^T (Z^T Z)^{-1} J\right)^{-1} C\right) \qquad (8)$$

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*

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

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

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

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 $\Sigma = U \Lambda U^T$ be the eigenvalue decomposition of Σ, where $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_p)$. Let $U = (U_1, U_2)$ where U_1 is $p \times r$ and U_2 is $p \times (p-r)$. Let $\Lambda_1 = \operatorname{diag}(\lambda_1, \ldots, \lambda_r)$ and $\Lambda_2 = \operatorname{diag}(\lambda_{r+1}, \ldots, \lambda_p)$. Then $\Sigma = U_1 \Lambda_1 U_1^T + U_2 \Lambda_2 U_2^T$.

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 $H_{Z\tilde{J}} = U_{Z\tilde{J}} U_{Z\tilde{J}}^T$. Then $T(Z) = \lambda_{\max}(C^T J^T Z^T (I_p - H_{Z\tilde{J}}) Z J C)$. Note that

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

Assumption 1. *Assume $C \geq \lambda_{r+1} \geq \ldots \geq \lambda_p \geq c$, where c and C are absolute constant.*

Theorem 1. *Suppose Assumption (1) holds. Suppose*

$$p/n \to \infty, \quad \textit{and} \quad \frac{\lambda_1^2 p}{\lambda_r^2 n^2} \to 0. \tag{11}$$

Suppose

$$\frac{\lambda_r n}{p} \to \infty. \tag{12}$$

Suppose

$$\frac{1}{\sqrt{p}} \|\mu_f\|_F^2 = O(1). \tag{13}$$

Then

$$(\operatorname{tr} \Lambda_2^2)^{-1/2} \big(C^T J^T Z^T (I_p - H_{Z\tilde{J}}) Z J C - (\operatorname{tr} \Lambda_2) I_{k-1} - \mu_f^T (I_p - H_{Z\tilde{Z}}) \mu_f \big) \xrightarrow{\mathcal{L}} W_{k-1}, \tag{14}$$

where W_{k-1} is a $(k-1) \times (k-1)$ symmetric random matrix whose entries above the main diagonal are i.i.d. $N(0,1)$ and the entries on the diagonal are i.i.d. $N(0,2)$.

4. Simulation Results

$$SNR = \frac{\|\mu_f\|_F^2}{\sqrt{\mathrm{tr}(\Sigma^2)}}$$

Table 1: $n_1 = n_2 = n_3 = 10$, non-sparse, $\Sigma = \mathrm{diag}()$

SNR	$p = 50$			$p = 75$			$p = 100$		
	SC	CX	NEW	SC	CX	NEW	SC	CX	NEW
0	2.000	3.000	4.000	5.000	6.000	7.000	8.000	9.000	10.000
0.1									
0.2									

Table 2: $n_1 = n_2 = n_3 = 25$, non-sparse

SNR	$p = 100$			$p = 150$			$p = 200$		
	SC	CX	NEW	SC	CX	NEW	SC	CX	NEW
0	2.000	3.000	4.000	5.000	6.000	7.000	8.000	9.000	10.000
0.1									
0.2									

Table 3: $n_1 = n_2 = n_3 = 10$, sparse

SNR	$p = 50$			$p = 75$			$p = 100$		
	SC	CX	NEW	SC	CX	NEW	SC	CX	NEW
0	2.000	3.000	4.000	5.000	6.000	7.000	8.000	9.000	10.000
0.1									
0.2									

5. Appendix

Proof of Theorem 1. It can be seen that ZJC is independent of $Z\tilde{J}$. Since $\mathrm{E}(Z\tilde{J}) = O_{p\times(n-k)}$, we can write $Z\tilde{J} = U\Lambda^{1/2}G_1$, where G_1 is a $p \times (n-k)$ matrix with i.i.d. $N(0,1)$ entries. We write $ZJC = \mu_f + U\Lambda^{1/2}G_2$, where G_2 is a $p \times (k-1)$ matrix with i.i.d. $N(0,1)$ entries.

Then

$$\begin{aligned} C^TJ^TZ^T(I_p - H_{Z\tilde{J}})ZJC =& G_2^T\Lambda^{1/2}U^T(I_P - H_{Z\tilde{J}})U\Lambda_{1/2}G_2 + \mu_f^T(I_p - H_{Z\tilde{J}})\mu_f + \\ & \mu_f^T(I_p - H_{Z\tilde{J}})U\Lambda^{1/2}G_2 + G_2^T\Lambda^{1/2}U^T(I_P - H_{Z\tilde{J}})\mu_f. \end{aligned} \tag{15}$$

To deal the first term, we note that

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

where $\xi_i \overset{i.i.d.}{\sim} N(0, I_{k-1})$. The key to its asymptotic behavior is the positive eigenvalues of $\Lambda^{1/2}U^T(I_p - H_{Z\tilde{J}})U\Lambda^{1/2}$, which in turn equal to the eigenvalues of $(I_p - H_{Z\tilde{J}})U\Lambda U^T(I_p - H_{Z\tilde{J}})$. Write $(I_p - H_{Z\tilde{J}})U\Lambda U^T(I_p - H_{Z\tilde{J}})$ as the sum of two terms

$$\begin{aligned} & (I_p - H_{Z\tilde{J}})U\Lambda U^T(I_p - H_{Z\tilde{J}}) \\ =& (I_p - H_{Z\tilde{J}})U_1\Lambda_1U_1^T(I_p - H_{Z\tilde{J}}) + (I_p - H_{Z\tilde{J}})U_2\Lambda_2U_2^T(I_p - H_{Z\tilde{J}}) \overset{def}{=} R_1 + R_2. \end{aligned}$$

Table 4: $n_1 = n_2 = n_3 = 25$, sparse

SNR	$p = 100$			$p = 150$			$p = 200$		
	SC	CX	NEW	SC	CX	NEW	SC	CX	NEW
0	2.000	3.000	4.000	5.000	6.000	7.000	8.000	9.000	10.000
0.1									
0.2									

Note that

$$\lambda_{\max}\big(R_1\big) = \lambda_{\max}\big(\Lambda_1^{1/2}U_1^T(I_p - H_{Z\tilde{J}})U_1\Lambda_1^{1/2}\big) \le \lambda_{\max}\big(\Lambda_1^{1/2}U_1^T(I_p - U_{Z\tilde{J}[,1:r]}U_{Z\tilde{J}[,1:r]}^T)U_1\Lambda_1^{1/2}\big)$$
$$\le \lambda_1\lambda_{\max}\big(U_1^T(I_p - U_{Z\tilde{J}[,1:r]}U_{Z\tilde{J}[,1:r]}^T)U_1\big) = \lambda_1\lambda_{\max}\big(I_r - U_1^T U_{Z\tilde{J}[,1:r]}U_{Z\tilde{J}[,1:r]}^T U_1\big).$$

To investigate the behavior of $U_{Z\tilde{J}}$, we need to investigate the behavior of $D_{Z\tilde{J}}$ first. Note that $G_1^T\Lambda G_1 = \tilde{J}^T Z^T Z\tilde{J} = V_{Z\tilde{J}}D_{Z\tilde{J}}^2 V_{Z\tilde{J}}^T$, and $G_1^T\Lambda G_1 = G_{1[1:r,]}^T\Lambda_1 G_{1[1:r,]} + G_{1[(r+1):p,]}^T\Lambda_2 G_{1[(r+1):p,]}$. We have

$$V_{Z\tilde{J}}D_{Z\tilde{J}}^2 V_{Z\tilde{J}}^T = G_{1[1:r,]}^T\Lambda_1 G_{1[1:r,]} + G_{1[(r+1):p,]}^T\Lambda_2 G_{1[(r+1):p,]}.$$

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

$$\begin{aligned} &\lambda_i(G_{1[1:r,]}^T\Lambda_1 G_{1[1:r,]}) \ge \lambda_i(G_{1[1:r,]}^T \operatorname{diag}(\lambda_i I_i, O_{(r-i)\times(r-i)})G_{1[1:r,]}) \\ =&\lambda_i\lambda_i(G_{1[1:i,]}G_{1[1:i,]}^T) = \lambda_i n(1 + o_P(1)), \end{aligned} \tag{16}$$

where 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 $i = 1, \dots, r$,

$$\begin{aligned} &\lambda_i(G_{1[1:r,]}^T\Lambda_1 G_{1[1:r,]}) \\ =&\lambda_i\Big(G_{1[1:r,]}^T\big(\operatorname{diag}(\lambda_1, \dots, \lambda_{i-1}, O_{(r-i+1)\times(r-i+1)}) + \operatorname{diag}(O_{(i-1)\times(i-1)}, \lambda_i, \dots, \lambda_r)\big)G_{1[1:r,]}\Big) \\ \le&\lambda_1(G_{1[1:r,]}^T \operatorname{diag}(O_{(i-1)\times(i-1)}, \lambda_i, \dots, \lambda_r)G_{1[1:r,]}) \le \lambda_1(G_{1[1:r,]}^T \operatorname{diag}(O_{(i-1)\times(i-1)}, \lambda_i I_{r-i+1})G_{1[1:r,]}) \\ =&\lambda_i\lambda_1(G_{1[i:r,]}G_{1[i:r,]}^T) = \lambda_i n(1 + o_P(1)) \end{aligned} \tag{17}$$

where the first inequality holds by Weyl's inequality. It follows from (16) and (17) that $\lambda_i(G_{1[1:r,]}^T\Lambda_1 G_{1[1:r,]}) = \lambda_i n(1 + o_P(1))$ for $i = 1, \dots, r$.

Note that $\lambda_{\max}(G_{1[(r+1):p,]}^T\Lambda_2 G_{1[(r+1):p,]}) \le C\lambda_{\max}(G_{1[(r+1):p,]}^T G_{1[(r+1):p,]}) = O_P(p)$ by Bai-Yin's law. By assumption $\lambda_r n/p \to \infty$, we can deduce that $D_{Z\tilde{J}[i,i]}^2 = \lambda_i(G_1^T\Lambda G_1) = \lambda_i n(1 + o_P(1))$, $i = 1, \dots, r$.

Now we are ready to investigate the behavior of $U_{Z\tilde{J}}$. Since $U\Lambda^{1/2}G_1G_1^T\Lambda^{1/2}U^T = U_{Z\tilde{J}}D_{Z\tilde{J}}^2U_{Z\tilde{J}}^T$, we have $G_1G_1^T = \Lambda^{-1/2}U^TU_{Z\tilde{J}}D_{Z\tilde{J}}^2U_{Z\tilde{J}}^TU\Lambda^{-1/2}$, which further indicates

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

Thus,

$$\lambda_{\max}(U^T_{[,(r+1):p]}U_{Z\tilde{J}[,1:r]}U^T_{Z\tilde{J}[,1:r]}U_{[,(r+1):p]}) \le \frac{C}{D^2_{Z\tilde{J}[r,r]}}\lambda_1(G_{1[(r+1):p,]}G^T_{1[(r+1):p,]}) = O_P(\frac{p}{\lambda_r n}).$$

Note that we have the simple relationship

$$\begin{aligned}
&\lambda_{\max}(U^T_{[,(r+1):p]}U_{Z\tilde{J}[,1:r]}U^T_{Z\tilde{J}[,1:r]}U_{[,(r+1):p]}) = \lambda_{\max}(U^T_{Z\tilde{J}[,1:r]}U_{[,(r+1):p]}U^T_{[,(r+1):p]}U_{Z\tilde{J}[,1:r]})\\
=&\lambda_{\max}(U^T_{Z\tilde{J}[,1:r]}(I_p - U_1U_1^T)U_{Z\tilde{J}[,1:r]}) = \lambda_{\max}(I_r - U^T_{Z\tilde{J}[,1:r]}U_1U_1^TU_{Z\tilde{J}[,1:r]})\\
=&1 - \lambda_{\min}(U^T_{Z\tilde{J}[,1:r]}U_1U_1^TU_{Z\tilde{J}[,1:r]}) = 1 - \lambda_{\min}(U_1^TU_{Z\tilde{J}[,1:r]}U^T_{Z\tilde{J}[,1:r]}U_1)\\
=&\lambda_{\max}(I_r - U_1^TU_{Z\tilde{J}[,1:r]}U^T_{Z\tilde{J}[,1:r]}U_1).
\end{aligned}$$

Therefore $\lambda_{\max}(I_r - U_1^TU_{Z\tilde{J}[,1:r]}U^T_{Z\tilde{J}[,1:r]}U_1) = O_P(\frac{p}{\lambda_r n})$, and we can conclude $\lambda_{\max}(R_1) = O_P(\frac{\lambda_1 p}{\lambda_r n})$.

We now deal with $R_1 + R_2$. For $i = 1, \dots, r$,

$$\lambda_i(R_1+R_2) \le \lambda_1(R_1+R_2) \le \lambda_1(R_1) + \lambda_1(R_2) \le O_P(\frac{\lambda_1 p}{\lambda_r n}) + C.$$

For $i = r+1, \dots, p-r$,

$$\lambda_i(R_1+R_2) \le \lambda_{i-r}(R_2) = \lambda_{i-r}\big(\Lambda_2^{1/2}U_2^T(I_p - H_{Z\tilde{J}})U_2\Lambda_2^{1/2}\big) \le \lambda_{i-r}(\Lambda_2) = \lambda_i.$$

On the other hand, for $i = 1, \dots, p-r-n+k$,

$$\begin{aligned}
&\lambda_i(R_1+R_2) \ge \lambda_i(R_2) = \lambda_i\big(\Lambda_2^{1/2}U_2^T(I_p - H_{Z\tilde{J}})U_2\Lambda_2^{1/2}\big)\\
=&\lambda_i\big(\Lambda_2 - \Lambda_2^{1/2}U_2^TH_{Z\tilde{J}}U_2\Lambda_2^{1/2}\big) \ge \lambda_{i+n-k}.
\end{aligned}$$

The last equality holds since $U_2^TH_{Z\tilde{J}}U_2$ is at most of rank $n-k$.

As a consequence of these bounds, we have

$$\sum_{i=1}^{p-r-n+k}\lambda^2_{i+n-k} \le \mathrm{tr}[(R_1+R_2)^2] \le r(O_P(\frac{\lambda_1 p}{\lambda_r n}) + C)^2 + \sum_{i=r+1}^{p-r}\lambda_i^2,$$

or

$$|\,\mathrm{tr}[(R_1+R_2)^2] - \sum_{i=r+1}^{p}\lambda_i^2| \le \sum_{i=r+1}^{n-k}\lambda_i^2 + \sum_{i=p-r+1}^{p}\lambda_i^2 + r(O_P(\frac{\lambda_1 p}{\lambda_r n}) + C)^2.$$

Similarly,

$$|\,\mathrm{tr}[(R_1+R_2)] - \sum_{i=r+1}^{p}\lambda_i| \le \sum_{i=r+1}^{n-k}\lambda_i + \sum_{i=p-r+1}^{p}\lambda_i + r(O_P(\frac{\lambda_1 p}{\lambda_r n}) + C).$$

These, conbined with the assumptions, yield

$$\operatorname{tr}[(R_1+R_2)^2] = (1+o_P(1))\sum_{i=r+1}^{p}\lambda_i^2,$$

and

$$\operatorname{tr}[(R_1+R_2)] = \sum_{i=r+1}^{p}\lambda_i + O(n) + O_P(\frac{\lambda_1 p}{\lambda_r n}).$$

Now we have the Lyapunov condition

$$\frac{\lambda_1[(R_1+R_2)^2]}{\operatorname{tr}[(R_1+R_2)^2]} = \frac{\big(O_P(\frac{\lambda_1 p}{\lambda_r n})+C\big)^2}{(1+o_P(1))\sum_{i=r+1}^{p}\lambda_i^2} \xrightarrow{P} 0.$$

Apply Lyapunov central limit theorem conditioning on $H_{Z\tilde{J}}$, we have

$$\big(\operatorname{tr}[(R_1+R_2)^2]\big)^{-1/2}\big(G_2^T\Lambda^{1/2}U^T(I_p-H_{Z\tilde{J}})U\Lambda_{1/2}G_2 - \operatorname{tr}(R_1+R_2)I_{k-1}\big) \xrightarrow{\mathcal{L}} W_{k-1}$$

where W_{k-1} is a $(k-1)\times(k-1)$ symmetric random matrix whose entries above the main diagonal are i.i.d. $N(0,1)$ and the entries on the diagonal are i.i.d. $N(0,2)$. By Slutsky's theorem, we have

$$\big(\sum_{i=r+1}^{p}\lambda_i^2\big)^{-1/2}\big(G_2^T\Lambda^{1/2}U^T(I_p-H_{Z\tilde{J}})U\Lambda_{1/2}G_2 - (\sum_{i=r+1}^{p}\lambda_i)I_{k-1}\big) \xrightarrow{\mathcal{L}} W_{k-1}$$

As for the cross term of (15), we have

$$\begin{aligned}
&\mathrm{E}[\|\mu_f^T(I_p-H_{Z\tilde{J}})U\Lambda^{1/2}G_2\|_F^2|Z\tilde{J}]\\
=&(k-1)\operatorname{tr}(\mu_f^T(I_p-H_{Z\tilde{J}})U\Lambda U^T(I_p-H_{Z\tilde{J}})\mu_f)\\
\leq&(k-1)\lambda_1\big((I_p-H_{Z\tilde{J}})U\Lambda U^T(I_p-H_{Z\tilde{J}})\big)\|\mu_f\|_F^2\\
=&(k-1)O_P(\frac{\lambda_1 p}{\lambda_r n})\|\mu_f\|_F^2\\
=&(k-1)O_P(\frac{\lambda_1\sqrt{p}}{\lambda_r n})\sqrt{p}\|\mu_f\|_F^2 = o_P(p)
\end{aligned}$$

The last equality holds when we assume $\frac{1}{\sqrt{p}}\|\mu_f\|_F^2 = O(1)$. Hence $\|\mu_f^T(I_p-H_{Z\tilde{J}})U\Lambda^{1/2}G_2\|_F^2 = o_P(p)$. This completes the proof of the theorem.

□

References