Likelihood Ratio Tests in Multivariate Variance Components Models

Satoshi Kuriki

Faculty of Engineering, University of Tokyo

October 1992

Acknowledgements

In preparing this thesis I am indebted to many people. First of all, I wish to express my sincere gratitude to Professor Chihiro Hirotsu, my thesis supervisor, for his guidance and encouragement.

I would like to express my thanks to Professor Kei Takeuchi, Professor Shun-ichi Amari, Professor Yasuo Ohashi, Professor Akimichi Takemura and Dr. Tatsuya Kubokawa for serving as members of the judging committee.

Thanks are also due to Professor Kazuo Murota for his generous advise on English writing, and to Professor Yasunori Fujikoshi, Professor Masaaki Sugihara, Dr. Hiroe Tsubaki, Dr. Michiko Tsubaki, Dr. Toshio Honda and Dr. Fumiyasu Komaki for their valuable comments and suggestions.

Finally I am very grateful to my parents for their continual encouragement, and to my wife, Hisae, who has shared both my joy and frustration in the course of my research.

Contents

Chapter I. Introduction 1

1. Multivariate variance components models 1
2. Asymptotic theory of the likelihood ratio test under restricted alternatives 2
3. Scope of the thesis 6

Appendix A. A proof of Lemma 2.1 7

Chapter II. One-sided test for the equality of two covariance matrices 9

1. Introduction 9
2. LRT statistics and their least favorable distributions 10
3. Limiting null distributions of LRT 13
    3.1. Derivation of the limiting null distribution 13
    3.2. Calculation of U_k's by recurrence formula 17
    3.3. An expression in terms of Pfaffian 19
4. Asymptotic expansion of the null distribution of LRT 24
5. Significance points and biases 26
6. Power comparisons 29

Appendix A. Examples of the limiting distributions 33
Appendix B. Distributions of the maximum and minimum roots 34

Chapter III. One-sided test for the equality of two covariance matrices concerning the complex multivariate normal population 36

1. Introduction 36
2. LRT statistics and their properties 37
3. Limiting null distributions of LRT 38

Appendix A. Examples of the limiting distributions 43

Chapter IV. Tests for covariance structure in random coefficient regression model 44

1. Introduction 44
2. Properties of power functions of LRT 45
    2.1. LRT statistics 45
    2.2. Monotonicity of power function of the LRT T_{12} 46
    2.3. Unbiasedness of the LRT T_{01} 49
3. Limiting null distribution of the test statistics of LRT 50

4. Local unbiasedness of a general class of tests for H_0 against $H_1 - H_0$ 55
5. Applications to testing problems in random coefficient regression model ... 58
6. Power comparisons ... 60
Appendix A. Examples of the limiting distributions ... 63
Appendix B. A sufficient condition for the FKG condition ... 63

References ... 67

I. Introduction.

1. Multivariate variance components model.

When effects of factors are random, the analysis of variance model is called random effects model or variance components model. There is long history of studying the statistical inference based on the variance components model, see Rao and Kleffe (1988), Searle, et al. (1992) and their references therein. Among them, however, neither the maximum likelihood nor the likelihood ratio test in the the multivariate variance components models is discussed well. Exact derivation of the maximum likelihood estimators and the likelihood ratio criteria are current topics, and discussions of their statistical properties such as distribution or optimality can be hardly found. The aim of the thesis is to derive the limiting null distributions of the likelihood ratio test statistics to determine the significance points, and to prove some optimalities of the likelihood ratio tests in multivariate variance components models.

One of the most typical multivariate variance components model is the following one-way classification model with random effects discussed by Anderson, et al. (1986):

$$\boldsymbol{X}_{ij} = \boldsymbol{\mu} + \boldsymbol{V}_i + \boldsymbol{U}_{ij}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, k, \tag{1.1}$$

where $\boldsymbol{X}_{ij}$ is a $p \times 1$ observed vector, $\boldsymbol{\mu}$ an unknown mean vector, $\boldsymbol{V}_i$ an unobserved random effect vector of group i, and $\boldsymbol{U}_{ij}$ an unobserved measurement error. $\boldsymbol{V}_i$ and $\boldsymbol{U}_{ij}$ are assumed to be independently distributed according to the normal distributions $N_p(\mathbf{0}, \boldsymbol{\Theta})$ and $N_p(\mathbf{0}, \boldsymbol{\Psi})$, respectively. The covariance matrix $\boldsymbol{\Theta}$ of the effect vector $\boldsymbol{V}_i$ is said to be effect matrix or multivariate components of variance. Here we discuss the two cases: the case where $\boldsymbol{\Psi}$ is fully unknow positive definite matrix, and the case where $\boldsymbol{\Psi} = \sigma^2 \boldsymbol{I}_p$ with $\sigma^2 > 0$ unknown. $\boldsymbol{\mu}$ and $\boldsymbol{\Theta}$ are assumed to be fully unknown. Note that the model (1.1) with $\boldsymbol{\Psi} = \sigma^2 \boldsymbol{I}_p$ is Scheffé's mixed model with replications (Scheffé, 1959).

Concerning the model (1.1) we shall be interested in estimating $\boldsymbol{\Theta}$. The uniformly minimum variance unbiased estimator $\hat{\boldsymbol{\Theta}}_{\mathrm{UMVU}}$ is

$$\begin{aligned}
\hat{\boldsymbol{\Theta}}_{\mathrm{UMVU}} &= \frac{1}{k}\left\{\frac{1}{n-1}\boldsymbol{H} - \frac{1}{n(k-1)}\boldsymbol{G}\right\} && \text{if } \boldsymbol{\Psi} \text{ is fully unknown,} \\
&= \frac{1}{k}\left\{\frac{1}{n-1}\boldsymbol{H} - \frac{1}{n(k-1)}\frac{\mathrm{tr}\boldsymbol{G}}{p}\boldsymbol{I}_p\right\} && \text{if } \boldsymbol{\Psi} = \sigma^2\boldsymbol{I}_p,
\end{aligned}$$

where $\boldsymbol{H}$ and $\boldsymbol{G}$ are between and within sum of squares matrices, distributed independently according to the Wishart distributions $W_p(M, \boldsymbol{\Phi})$ and $W_p(N, \boldsymbol{\Psi})$ with

$\boldsymbol{\Phi} = \boldsymbol{\Psi} + k\boldsymbol{\Theta}$, $M = n-1$ and $N = n(k-1)$, respectively. Since $\boldsymbol{H}$ and $\boldsymbol{G}$ are independent, $\hat{\boldsymbol{\Theta}}_{\mathrm{UMVU}}$ has negative latent roots with positive probability. This seems to mean that $\hat{\boldsymbol{\Theta}}_{\mathrm{UMVU}}$ is not suitable for the estimator of the covariance matrix $\boldsymbol{\Theta}$. The maximum likelihood estimation is one method to avoid this difficulty. The maximum likelihood estimator (MLE) of $\boldsymbol{\Theta}$ was given by Klotz and Putter (1969) if $\boldsymbol{\Psi}$ is fully unknown; Anderson, et al. (1986) if $\boldsymbol{\Psi} = \sigma^2 \boldsymbol{I}_p$ with σ^2 unknown. Calvin and Dykstra (1991) obtained an algorithm to get the MLE concerning the two-factor random effects model. The MLE of the effect matrix is the point maximizing the likelihood function under Löwner order restriction (i.e. $\boldsymbol{\Theta} \geq \boldsymbol{O}$), and as a result both the algorithm to maximize the likelihood function and the distribution of the obtained estimator are complex generally. ($\boldsymbol{A} \geq \boldsymbol{B}$ denotes Löwner order meaning that $\boldsymbol{A} - \boldsymbol{B}$ is nonnegative definite.)

A similar difficulty appears in the testing problem concerning $\boldsymbol{\Theta}$. Consider testing the homogeneity hypothesis $\boldsymbol{\Theta} = \boldsymbol{O}$. The hypothesis $\boldsymbol{\Theta} = \boldsymbol{O}$ is equivalent to the equality of two covariance matrices $\boldsymbol{\Phi} = \boldsymbol{\Psi}$, however, the usual methods of testing the equality of two covariance matrices based on the sample covariance matrices $\hat{\boldsymbol{\Phi}} = (1/M)\boldsymbol{H}$ and $\hat{\boldsymbol{\Psi}} = (1/N)\boldsymbol{G}$ (e.g. Anderson, 1984b, Section 10.6; Nagao, 1973) are ineffective for our purpose, because the alternative $\boldsymbol{\Theta} \geq \boldsymbol{O}$ ($\boldsymbol{\Phi} \geq \boldsymbol{\Psi}$) is not taken into account. The likelihood ratio test (LRT) is one useful method of testing when the alternative is restricted. Anderson, et al. (1986) obtained the likelihood ratio criterion for testing the hypothesis $\operatorname{rank} \boldsymbol{\Theta} \leq r$, which reduces to a one-sided test for the equality of two covariance matrices when $r = 0$. As the case of the maximum likelihood estimation, both the algorithm to obtain the likelihood ratio criterion and its (non)null distribution are complex. In particular, as Anderson, et al. (1986) mentioned, the null distributions of (-2) times the logarithm of the likelihood ratio criteria are not chi-squared distribution even though asymptotically. They are shown to be mixtures of chi-squared distributions by means of the general theory of Chernoff (1954) described in the following section.

2. Asymptotic theory of the likelihood ratio test under restricted alternatives.

Chernoff (1954) discussed the limiting null distribution of the likelihood ratio criterion when the null hypothesis lies on the boundary of the alternative hypothesis in the parameter space. Following theorem is based on Chernoff (1954, Theorem 1), and its rearrangement by Self and Liang (1987, Theorem 3).

Definition 2.1 The set $\omega \subset \mathcal{R}^p$ is said to be approximated at $\theta_0 \in \mathcal{R}^p$ by a cone $\mathcal{C} \subset \mathcal{R}^p$ if

$$\inf_{x \in \mathcal{C}} \|x - y\| = \mathrm{o}(\|y\|) \quad \text{for } y \in \omega - \theta_0,$$

and

$$\inf_{y\in\omega-\theta_0} \|x-y\| = \mathrm{o}(\|x\|) \quad \text{for } x \in \mathcal{C}$$

holds. Here $\mathrm{o}(\,\cdot\,)$ means $\mathrm{o}(t)/t \to 0$ as $t \to 0$, and $y \in \omega - \theta_0$ means $y + \theta_0 \in \omega$.

Theorem 2.1 Let $X_i \in \mathcal{R}^k$, $i = 1, \ldots, n$, be n independent observation with common density $f(x, \theta)$, $\theta \in \Omega \subset \mathcal{R}^p$. Assume that the true value θ_0 of the parameter θ is an interior point of Ω. Moreover, we assume the regularity conditions of Lehmann (1983, (A0)-(A2) of Section 6.2, and (A)-(D) of Section 6.4) for the density function f. The Fisher information matrix is denoted by $I(\theta)$. Consider the problem of testing the null hypothesis $H_0 : \theta \in \omega_0 \subset \Omega$ against the alternative $H_1 : \theta \in \omega_1 \subset \Omega$, and suppose that the sets ω_0 and ω_1 are approximated at θ_0 by the closed cones $\mathcal{C}_0$ and $\mathcal{C}_1$, respectively. Then, as $n \to \infty$, $-2 \log \Lambda_n$ with

$$\Lambda_n = \frac{\sup_{\theta\in\omega_1} \prod_{i=1}^n f(X_i, \theta)}{\sup_{\theta\in\omega_0} \prod_{i=1}^n f(X_i, \theta)}$$

converges to

$$\overline{\chi}^2 = \min_{\theta\in\mathcal{C}_0} (Z-\theta)' I(\theta_0)(Z-\theta) - \min_{\theta\in\mathcal{C}_1} (Z-\theta)' I(\theta_0)(Z-\theta) \tag{2.1}$$

in distribution, where Z is a $p \times 1$ random vector distributed as $N_p(0, I(\theta_0)^{-1})$.

We summarize what is known about the distribution of (2.1) as a following theorem. (e.g. Shapiro, 1988; Shapiro, 1985, Theorem 3.1.)

Theorem 2.2 Suppose that $\mathcal{C}_0$ and $\mathcal{C}_1$ are closed convex cone such that $\mathcal{C}_0 \subset \mathcal{C}_1$, and that either of them is a linear subspace. Then the distribution of $\overline{\chi}^2$ in (2.1) is a mixture of chi-squared distributions

$$\Pr(\overline{\chi}^2 \le c) = \sum_{i=0}^{\dim\mathcal{C}_1 - \dim\mathcal{C}_0} w_i \Pr(\chi_i^2 \le c) \quad \text{with} \;\; w_i \ge 0, \;\; \sum_i w_i = 1,$$

where χ_i^2 is a chi-squared random variable with i degrees of freedom, $\chi_0^2 = 0$, and $\dim \mathcal{C}$ denotes the dimension of the smallest linear subspace containing the cone $\mathcal{C}$.

Remark 2.1 For the mixing probability $\{w_i\}$ the relation

$$\sum_i (-1)^i w_i = 0 \tag{2.2}$$

holds. One proof of (2.2) is found in McMullen (1975, Theorems 2, 3) in terms of the internal and external angles. For the geometrical interpretation of $\{w_i\}$, see also Wynn (1975) and Shapiro (1987).

Now we apply the general theory of the likelihood ratio test to the problem of testing

$$H_0: \quad \operatorname{rank}\boldsymbol{\Theta} \le r \tag{2.3}$$

for a specified r $(0 \le r < p)$ under the model (1.1) with $\boldsymbol{\Psi}$ fully unknown. We regard

$$\boldsymbol{Y}_i = (\boldsymbol{X}_{i1}{}', \ldots, \boldsymbol{X}_{ik}{}')'{}_{kp\times 1}, \quad i = 1, \ldots, n,$$

as the i.i.d. n samples from the normal population $N_{kp}(\boldsymbol{\xi}, \boldsymbol{\Sigma})$ with a mean vector $\boldsymbol{\xi} = (\boldsymbol{\mu}', \ldots, \boldsymbol{\mu}')'{}_{kp\times 1}$ and a covariance matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Psi}+\boldsymbol{\Theta} & \boldsymbol{\Theta} & \cdots & \boldsymbol{\Theta} \\ \boldsymbol{\Theta} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{\Theta} \\ \boldsymbol{\Theta} & \cdots & \boldsymbol{\Theta} & \boldsymbol{\Psi}+\boldsymbol{\Theta} \end{pmatrix}_{kp\times kp}.$$

Noting that $\boldsymbol{\Sigma}$ is positive definite (p.d.) if and only if both $\boldsymbol{\Phi} = \boldsymbol{\Psi} + k\boldsymbol{\Theta}$ and $\boldsymbol{\Psi}$ are positive definite, we put

$$\Omega = \{(\boldsymbol{\mu}, \boldsymbol{\Phi}, \boldsymbol{\Psi}) \in \mathcal{R}^p \times \mathcal{R}^{p(p+1)/2} \times \mathcal{R}^{p(p+1)/2} \mid \boldsymbol{\Phi}, \boldsymbol{\Psi} : \text{p.d.}\}.$$

Here $\mathcal{R}^{p(p+1)/2}$ denotes the set of $p \times p$ symmetric matrices. Testing H_0 in (2.3) under the model (1.1) reduces to testing the null hypothesis $H_0 : (\boldsymbol{\mu}, \boldsymbol{\Phi}, \boldsymbol{\Psi}) \in \omega_0 \subset \Omega$ with

$$\omega_0 = \{(\boldsymbol{\mu}, \boldsymbol{\Phi}, \boldsymbol{\Psi}) \in \mathcal{R}^p \times \mathcal{R}^{p(p+1)/2} \times \mathcal{R}^{p(p+1)/2} \mid \boldsymbol{\Phi} \ge \boldsymbol{\Psi}, \operatorname{rank}(\boldsymbol{\Phi} - \boldsymbol{\Psi}) \le r, \boldsymbol{\Psi} : \text{p.d.}\}$$

against the alternative hypothesis $H_1 : (\boldsymbol{\mu}, \boldsymbol{\Phi}, \boldsymbol{\Psi}) \in \omega_1 \subset \Omega$ with

$$\omega_1 = \{(\boldsymbol{\mu}, \boldsymbol{\Phi}, \boldsymbol{\Psi}) \in \mathcal{R}^p \times \mathcal{R}^{p(p+1)/2} \times \mathcal{R}^{p(p+1)/2} \mid \boldsymbol{\Phi} \ge \boldsymbol{\Psi}, \boldsymbol{\Psi} : \text{p.d.}\}.$$

Suppose that the true value of the parameters $(\boldsymbol{\mu}_0, \boldsymbol{\Phi}_0, \boldsymbol{\Psi}_0)$ are in ω_0 (i.e. H_0 holds) and satisfy

$$\operatorname{rank}\boldsymbol{\Theta}_0 = r \quad \text{with} \quad \boldsymbol{\Theta}_0 = (1/k)(\boldsymbol{\Phi}_0 - \boldsymbol{\Psi}_0).$$

Since H_0 and H_1 remain invariant under a transformation $\boldsymbol{X}_{ij} \mapsto \boldsymbol{G}\boldsymbol{X}_{ij}$ with $\boldsymbol{G}$ $p \times p$ nonsingular matrix, by virtue of the invariance of likelihood ratio (Lehmann, 1986, p.341, Problem 17), we can put $\boldsymbol{\Psi}_0 = \boldsymbol{I}_p$ and

$$\boldsymbol{\Theta}_0 = \begin{pmatrix} \boldsymbol{\Delta} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}_{p\times p} \quad \text{with} \ \ \boldsymbol{\Delta} = \operatorname{diag}(\delta_i)_{r\times r}, \quad \delta_i > 0,$$

without loss of generality. The cones approximating ω_0 and ω_1 are given as follows.

Lemma 2.1 The set ω_0 is approximated at $(\boldsymbol{\mu}_0, \boldsymbol{\Phi}_0, \boldsymbol{\Psi}_0)$ by the linear subspace

$$\mathcal{C}_0 = \{(\boldsymbol{\mu}, \boldsymbol{\Phi}, \boldsymbol{\Psi}) \in \mathcal{R}^p \times \mathcal{R}^{p(p+1)/2} \times \mathcal{R}^{p(p+1)/2} \mid \boldsymbol{\Phi}_{22} = \boldsymbol{\Psi}_{22}\},$$

where $\boldsymbol{\Phi}_{22}$ and $\boldsymbol{\Psi}_{22}$ denote the $(p-r)\times(p-r)$ lower right matrices of $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$, respectively. The set ω_1 is approximated at $(\boldsymbol{\mu}_0, \boldsymbol{\Phi}_0, \boldsymbol{\Psi}_0)$ by the closed convex cone

$$\mathcal{C}_1 = \{(\boldsymbol{\mu}, \boldsymbol{\Phi}, \boldsymbol{\Psi}) \in \mathcal{R}^p \times \mathcal{R}^{p(p+1)/2} \times \mathcal{R}^{p(p+1)/2} \mid \boldsymbol{\Phi}_{22} \geq \boldsymbol{\Psi}_{22}\}.$$

Here we define that $\|\boldsymbol{A}\| = \sqrt{\mathrm{tr}\boldsymbol{A}\boldsymbol{A}'} = \sqrt{\sum_{ij} {a_{ij}}^2}$ for a matrix $\boldsymbol{A} = (a_{ij})$.

(Proof) Proof is given in Appendix A. $\square$

Let Λ_n be the likelihood ratio criterion for testing H_0 against H_1. By virtue of Theorem 2.1 and Lemma 2.1, evaluating the Fisher information matrix at $(\boldsymbol{\mu}_0, \boldsymbol{\Phi}_0, \boldsymbol{\Psi}_0)$, we immediately obtain the limiting distribution of $-2\log\Lambda_n$ as

$$\begin{aligned}\overline{\chi}^2 = &\min_{\Phi_{22}=\Psi_{22}} \{(\boldsymbol{Z}-\boldsymbol{\Phi}_{22})(\boldsymbol{Z}-\boldsymbol{\Phi}_{22})' + (k-1)(\boldsymbol{W}-\boldsymbol{\Psi}_{22})(\boldsymbol{W}-\boldsymbol{\Psi}_{22})'\} \\ &- \min_{\Phi_{22}\geq\Psi_{22}} \{(\boldsymbol{Z}-\boldsymbol{\Phi}_{22})(\boldsymbol{Z}-\boldsymbol{\Phi}_{22})' + (k-1)(\boldsymbol{W}-\boldsymbol{\Psi}_{22})(\boldsymbol{W}-\boldsymbol{\Psi}_{22})'\}, \quad (2.4)\end{aligned}$$

where $\boldsymbol{Z} = (z_{ij})$ is a $(p-r)\times(p-r)$ symmetric random matrix whose diagonal element z_{ii} and off-diagonal element z_{ij} $(i<j)$ are independently distributed as $N(0,1)$ and $N(0,1/2)$, respectively; $\boldsymbol{W}$ is a $(p-r)\times(p-r)$ symmetric random matrix distributed according to the same distribution as $(1/\sqrt{k-1})\boldsymbol{Z}$, independently of $\boldsymbol{Z}$. Putting

$$\boldsymbol{A} = \sqrt{\frac{k-1}{k}}(\boldsymbol{Z}-\boldsymbol{W}) \quad \text{and} \quad \boldsymbol{M} = \sqrt{\frac{k-1}{k}}(\boldsymbol{\Phi}_{22}-\boldsymbol{\Psi}_{22}),$$

we have

$$\begin{aligned}\text{R.H.S. of (2.4)} &= \mathrm{tr}\boldsymbol{A}\boldsymbol{A}' - \min_{M\geq O} \mathrm{tr}(\boldsymbol{A}-\boldsymbol{M})(\boldsymbol{A}-\boldsymbol{M})' \\ &= \sum_{b_i>0} {b_i}^2, \qquad (2.5)\end{aligned}$$

where $b_1 > \cdots > b_{p-r}$ are the latent roots of $\boldsymbol{A}$. Note that $\boldsymbol{A}$ has the same distribution as $\boldsymbol{Z}$. Anderson (1989), Anderson and Amemiya (1991) showed that $-2\log\Lambda_n$ converges to R.H.S. of (2.5) in distribution by more straightforward calculations. Theorem 2.2 states that the distribution of $\overline{\chi}^2$ in (2.4) is a mixture of chi-squared distributions with ν $(0 \leq \nu \leq (p-r)(p-r+1)/2)$ degrees of freedom.

3. Scope of the thesis.

As shown in the previous section, the limiting null distribution of (-2) times the logarithm of the likelihood ratio criterion for testing the rank of the effect matrix is a mixture of chi-squared distributions. But its mixing probability can not be derived by Chernoff's general theory. In this thesis, in three typical multivariate variance components models, the mixing probabilities of the limiting null distributions of the likelihood ratio test statistics for effect matrices are derived, and some properties of their power functions are proved.

In Chapter II, concerning the model (1.1) with $\boldsymbol{\Psi}$ fully unknown, we discuss the one-sided likelihood ratio test based on the two sample covariances $\hat{\boldsymbol{\Phi}} = (1/M)\boldsymbol{H}$ and $\hat{\boldsymbol{\Psi}} = (1/N)\boldsymbol{G}$ for testing

(i) the hypothesis $\boldsymbol{\Theta} = \boldsymbol{O}$, i.e. the hypothesis $\boldsymbol{\Phi} = \boldsymbol{\Psi}$,

(ii) the hypothesis $\operatorname{rank} \boldsymbol{\Theta} \leq r$, i.e. the hypothesis that $\boldsymbol{\Phi} \geq \boldsymbol{\Psi}$ and $\operatorname{rank}(\boldsymbol{\Phi} - \boldsymbol{\Psi}) \leq r$, for a specified r $(0 < r < p)$ and

(iii) the goodness of fit of the covariance structure that $\boldsymbol{\Phi} = \boldsymbol{\Psi} + k\boldsymbol{\Theta}$, i.e. the hypothesis $\boldsymbol{\Phi} \geq \boldsymbol{\Psi}$.

The LRT's for testing (i) and (ii) are one-sided tests for two covariance matrices $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ because the alternative is $\boldsymbol{\Phi} \geq \boldsymbol{\Psi}$. The hypothesis that $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ are unrestricted is settled as the alternative of the LRT for testing (iii). Some properties of the power functions such as unbiasedness, monotonicity and consistency of the LRT's are proved. We derive the limiting null distributions as mixtures of chi-squared distributions, and give the table of quantiles based on the limiting distribution. In addition, the asymptotic expansions of the null distributions are given. The amounts of the bias of the LRT's for testing (ii) and (iii), which are shown to be biased, are evaluated. A Monte Carlo study to compare the power of several tests for (i) including the LRT is given.

In Chapter III we discuss the one-sided likelihood ratio tests for testing (i)-(iii) of Chapter II in the model (1.1) with the unobserved random variables $\boldsymbol{U}_i$ and $\boldsymbol{V}_{ij}$ distributed according to the complex normal distributions. This multivariate variance components model concerning the complex normal population is shown to appear in the frequency representation of the stationary Gaussian multiple time series model with replications.

In Chapter IV we discuss two likelihood ratio tests for covariance structure in the random coefficient model introduced by Rao (1965). This model includes the model (1.1) with $\boldsymbol{\Psi} = \sigma^2 \boldsymbol{I}_p$ as a special case. In terms of the model (1.1) with $\boldsymbol{\Psi} = \sigma^2 \boldsymbol{I}_p$, two LRT's treated in this chapter are based on the statistics $\hat{\boldsymbol{\Phi}} = (1/M)\boldsymbol{H}$ and $\hat{\sigma}^2 = (1/Np)\operatorname{tr} \boldsymbol{G}$ for testing

(i) the hypothesis $\boldsymbol{\Theta} = \boldsymbol{O}$, i.e. the hypothesis $\boldsymbol{\Phi} = \sigma^2 \boldsymbol{I}_p$, and

(ii) the goodness of fit of the covariance structure that $\boldsymbol{\Phi} = \sigma^2 \boldsymbol{I}_p + k\boldsymbol{\Theta}$, i.e. the hypothesis $\boldsymbol{\Phi} \geq \sigma^2 \boldsymbol{I}_p$.

As in Chapter II, the LRT for testing (i) is one-sided test because the alternative is $\boldsymbol{\Phi} \geq \sigma^2 \boldsymbol{I}_p$, and the the hypothesis that $\boldsymbol{\Phi}$ and σ^2 are unrestricted is settled as the alternative of the LRT for testing (ii). The unbiasedness of the LRT for (i) and the monotonicity of power function of the LRT for (ii) are proved. The limiting null distribution of these LRT statistics are also obtained. For a general class of tests for (i) including the LRT, the local unbiasedness is proved using FKG inequality. Here a new sufficient condition for the FKG condition is posed. A Monte Carlo study to compare the power of several tests for (i) including the LRT is also given.

Appendix A. A proof of Lemma 2.1.

Noting that $\boldsymbol{\Psi}$ is positive definite when $\boldsymbol{\Psi}$ is in a neighborhood in $\mathcal{R}^{p(p+1)/2}$ with center at the true value $\boldsymbol{\Psi}_0 = \boldsymbol{I}_p$, we see it sufficient to show the following (a) and (b).

(a) The set

$$\Omega_0 = \{\boldsymbol{\Theta} \in \mathcal{R}^{p(p+1)/2} \mid \boldsymbol{\Theta} \geq \boldsymbol{O}, \operatorname{rank} \boldsymbol{\Theta} \leq r\}$$

is approximated at

$$\boldsymbol{\Theta}_0 = \begin{pmatrix} \boldsymbol{\Delta} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}_{p\times p} \quad \text{with} \quad \boldsymbol{\Delta} = \operatorname{diag}(\delta_i)_{r\times r}, \quad \delta_i > 0,$$

by the linear subspace

$$\mathcal{C}_{\Omega_0} = \{\boldsymbol{\Theta} \in \mathcal{R}^{p(p+1)/2} \mid \boldsymbol{\Theta}_{22} = \boldsymbol{O}\},$$

where $\boldsymbol{\Theta}_{22}$ denotes the $(p-r)\times(p-r)$ lower right matrix of $\boldsymbol{\Theta}$.

(b) The set

$$\Omega_1 = \{\boldsymbol{\Theta} \in \mathcal{R}^{p(p+1)/2} \mid \boldsymbol{\Theta} \geq \boldsymbol{O}\}$$

is approximated at $\boldsymbol{\Theta}_0$ by the closed convex cone

$$\mathcal{C}_{\Omega_1} = \{\boldsymbol{\Theta} \in \mathcal{R}^{p(p+1)/2} \mid \boldsymbol{\Theta}_{22} \geq \boldsymbol{O}\}.$$

Proof of (a): Fix

$$\boldsymbol{Y} = \begin{pmatrix} \boldsymbol{Y}_{11} & \boldsymbol{Y}_{12} \\ \boldsymbol{Y}_{12}' & \boldsymbol{Y}_{22} \end{pmatrix} \in \Omega_0 - \boldsymbol{\Theta}_0$$

such that $\|\boldsymbol{Y}\| \leq \sqrt{r}\delta/2$ with $\delta = \min \delta_i > 0$. Noting that $\boldsymbol{Y}_{11} + \boldsymbol{\Delta} \geq (\delta/2)\boldsymbol{I}_r$ and hence $\boldsymbol{Y}_{22} = \boldsymbol{Y}_{12}'(\boldsymbol{Y}_{11} + \boldsymbol{\Delta})^{-1}\boldsymbol{Y}_{12}$, we have

$$\inf_{X\in\mathcal{C}_{\Omega_0}} \|\boldsymbol{X} - \boldsymbol{Y}\| = \|\boldsymbol{Y}_{22}\| \leq \frac{2}{\delta}\, \|\boldsymbol{Y}_{12}\|^2 = \mathrm{o}(\|\boldsymbol{Y}\|).$$

Next, fix

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_{11} & \boldsymbol{X}_{12} \\ \boldsymbol{X}_{12}' & \boldsymbol{O} \end{pmatrix} \in \mathcal{C}_{\Omega_0}$$

such that $\|\boldsymbol{X}\| \le \sqrt{r}\delta/2$. Noting that $\boldsymbol{X}_{11} + \boldsymbol{\Delta} \ge (\delta/2)\boldsymbol{I}_r$ and

$$\begin{pmatrix} \boldsymbol{X}_{11} & \boldsymbol{X}_{12} \\ \boldsymbol{X}_{12}' & \boldsymbol{X}_{12}'(\boldsymbol{X}_{11} + \boldsymbol{\Delta})^{-1}\boldsymbol{X}_{12} \end{pmatrix} \in \Omega_0 - \boldsymbol{\Theta}_0,$$

we have

$$\inf_{Y \in \Omega_0 - \Theta} \|\boldsymbol{X} - \boldsymbol{Y}\| \le \|\boldsymbol{X}_{12}'(\boldsymbol{X}_{11} + \boldsymbol{\Delta})^{-1}\boldsymbol{X}_{12}\| \le \frac{2}{\delta}\|\boldsymbol{X}_{12}\|^2 = \mathrm{o}(\|\boldsymbol{X}\|).$$

□

Proof of (b): We see first that

$$\inf_{X \in \mathcal{C}_{\Omega_1}} \|\boldsymbol{X} - \boldsymbol{Y}\| = 0 \quad \text{for any } \boldsymbol{Y} \in \Omega_1 - \boldsymbol{\Theta}_0,$$

because $\Omega_1 - \boldsymbol{\Theta}_0 \subset \mathcal{C}_{\Omega_1}$.

Next, fix

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_{11} & \boldsymbol{X}_{12} \\ \boldsymbol{X}_{12}' & \boldsymbol{X}_{22} \end{pmatrix} \in \mathcal{C}_{\Omega_1}$$

such that $\|\boldsymbol{X}\| \le \sqrt{r}\delta/2$. Noting that $\boldsymbol{X}_{11} + \boldsymbol{\Delta} \ge (\delta/2)\boldsymbol{I}_r$, $\boldsymbol{X}_{22} \ge \boldsymbol{O}$ and

$$\begin{pmatrix} \boldsymbol{X}_{11} & \boldsymbol{X}_{12} \\ \boldsymbol{X}_{12}' & \boldsymbol{X}_{22} + \boldsymbol{X}_{12}'(\boldsymbol{X}_{11} + \boldsymbol{\Delta})^{-1}\boldsymbol{X}_{12} \end{pmatrix} \in \Omega_1 - \boldsymbol{\Theta}_0,$$

we have

$$\inf_{Y \in \Omega_1 - \Theta} \|\boldsymbol{X} - \boldsymbol{Y}\| \le \|\boldsymbol{X}_{12}'(\boldsymbol{X}_{11} + \boldsymbol{\Delta})^{-1}\boldsymbol{X}_{12}\| \le \frac{2}{\delta}\|\boldsymbol{X}_{12}\|^2 = \mathrm{o}(\|\boldsymbol{X}\|).$$

□

II. One-sided test for the equality of two covariance matrices.

1. Introduction.

Let $\boldsymbol{H}$ and $\boldsymbol{G}$ be $p \times p$ random matrices which are independently distributed according to the Wishart distributions $W_p(M, \boldsymbol{\Phi})$ and $W_p(N, \boldsymbol{\Psi})$, respectively, where $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ are assumed to be positive definite and $M \geq p$, $N \geq p$. Consider the hierarchical hypotheses $H_0 \subset H_0^{(r)} \subset H_1 \subset H_2$ with

$$\begin{aligned} &H_0 : \boldsymbol{\Phi} = \boldsymbol{\Psi}, && H_0^{(r)} : \boldsymbol{\Phi} \geq \boldsymbol{\Psi},\ \text{rank}\,(\boldsymbol{\Phi} - \boldsymbol{\Psi}) \leq r \quad (0 < r < p), \\ &H_1 : \boldsymbol{\Phi} \geq \boldsymbol{\Psi}, && H_2 : \boldsymbol{\Phi}, \boldsymbol{\Psi} \ \text{ are unrestricted.} \end{aligned}$$

Here $\boldsymbol{A} \geq \boldsymbol{B}$ denotes Löwner order meaning that $\boldsymbol{A} - \boldsymbol{B}$ is nonnegative definite. In this chapter the likelihood ratio tests (LRT's) for the following hypotheses:

(i) T_{01} : LRT for testing H_0 against $H_1 - H_0$,
(ii) $T_{01}^{(r)}$: LRT for testing $H_0^{(r)}$ against $H_1 - H_0^{(r)}$, and
(iii) T_{12} : LRT for testing H_1 against $H_2 - H_1$

are discussed. The main purpose is to derive the limiting null distributions of test statistics of these LRT's.

These testing problems appear in the multivariate variance components model:

$$\boldsymbol{X}_{ij} = \boldsymbol{\mu} + \boldsymbol{V}_i + \boldsymbol{U}_{ij}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, k, \tag{1.1}$$

where $\boldsymbol{X}_{ij}$ is a $p \times 1$ observed vector, $\boldsymbol{\mu}$ an unknown mean vector, $\boldsymbol{V}_i$ an unobserved random effect vector of group i, and $\boldsymbol{U}_{ij}$ an unobserved measurement error. $\boldsymbol{V}_i$ and $\boldsymbol{U}_{ij}$ are assumed to be independently distributed according to the normal distributions $N_p(\boldsymbol{0}, \boldsymbol{\Theta})$ and $N_p(\boldsymbol{0}, \boldsymbol{\Psi})$, respectively. The complete sufficient statistics of the model (1.1) are $\overline{\overline{\boldsymbol{X}}} = \sum_{i=1}^n \overline{\boldsymbol{X}}_i / n$ with $\overline{\boldsymbol{X}}_i = \sum_{j=1}^k \boldsymbol{X}_{ij}/k$,

$$\boldsymbol{H} = k \sum_{i=1}^{n} (\overline{\boldsymbol{X}}_i - \overline{\overline{\boldsymbol{X}}})(\overline{\boldsymbol{X}}_i - \overline{\overline{\boldsymbol{X}}})', \tag{1.2}$$

and

$$\boldsymbol{G} = \sum_{i=1}^{n} \sum_{j=1}^{k} (\boldsymbol{X}_{ij} - \overline{\boldsymbol{X}}_i)(\boldsymbol{X}_{ij} - \overline{\boldsymbol{X}}_i)'. \tag{1.3}$$

$\boldsymbol{H}$ and $\boldsymbol{G}$ are distributed according to $W_p(M, \boldsymbol{\Phi})$ and $W_p(N, \boldsymbol{\Psi})$, respectively, where $\boldsymbol{\Phi} = \boldsymbol{\Psi} + k\boldsymbol{\Theta}$, $M = n - 1$ and $N = n(k-1)$. Then testing the null hypothesis of no effect: $\boldsymbol{\Theta} = \boldsymbol{O}$ reduces to testing H_0 based on $\boldsymbol{H}$ (1.2) and $\boldsymbol{G}$ (1.3) against the one-sided alternative hypothesis $H_1 - H_0$. We are also interested in testing the null hypothesis that the effect vectors are linear combinations of r or less factors, which reduces to $H_0^{(r)}$, and the alternative should be $H_1 - H_0^{(r)}$. Testing the goodness of fit of the model (1.1) gives another type of restricted inference, which amounts to testing H_1 against the alternative $H_2 - H_1$.

Anderson (1984b, Section 10.6.2) and Anderson, et al. (1986) discussed the LRT's T_{01} and $T_{01}^{(r)}$, and pointed out that the null distributions of (-2) times the logarithm of the likelihood ratio criteria are not chi-squared distribution even though asymptotically. Anderson (1984a) discussed the same tests in the context of structural relationship models. We shall derive the limiting null distributions of the test statistics of T_{01}, $T_{01}^{(r)}$ and T_{12} which was obtained by Sakata (1987) and Anderson (1989) only for $p = 2$ $(p - r = 2)$; and give the table of quantiles of the distributions, which is the exact version of the table by Amemiya, et al. (1990) who estimated the quantiles by Monte Carlo simulations.

Outline of this chapter is as follows. In Section 2, the likelihood ratio test statistics and some of their properties are listed. In Section 3, we derive the limiting null distributions of the likelihood ratio test statistics by two different methods. Moreover the asymptotic expansions of the null distributions of the likelihood ratio test statistics are derived in Section 4. In Section 5, limiting significance points and biases of the tests are tabulated. Appendix A illustrates several examples of the asymptotic null distributions. Appendix B gives the formulae of the distribution function of the maximum latent root of a random matrix that appears as the limit of a Wishart matrix.

2. LRT statistics and their least favorable distributions.

We give the test statistics and their least favorable distributions to calculate the significance points here. Let Λ_{01}, $\Lambda_{01}^{(r)}$ and Λ_{12} be the likelihood ratio criteria for T_{01}, $T_{01}^{(r)}$ and T_{12}, respectively. Anderson (1984a, 1984b) and Anderson, et al. (1986) showed that

$$\begin{aligned}
\Lambda_{01}^{(r)} &= \prod_{i=r+1}^{R} \{\frac{{l_i}^{\rho}}{\rho l_i + 1 - \rho}\}^{(M+N)/2} && \text{if } R \geq r+1, \\
&= 1 && \text{if } R \leq r, \\
\Lambda_{01} &= \Lambda_{01}^{(0)}
\end{aligned}$$

with $\rho = M/(M+N)$, $l_1 > \cdots > l_p$ the latent roots of $(N/M)\boldsymbol{H}\boldsymbol{G}^{-1}$, and R the number of $l_i > 1$. Obviously the likelihood ratio criterion for testing H_0 against

$H_2 - H_0$ is

$$\Lambda_{02} = \prod_{i=1}^{p} \{ \frac{l_i{}^{\rho}}{\rho l_i + 1 - \rho} \}^{(M+N)/2},$$

and hence we have

$$\begin{aligned} \Lambda_{12} = \frac{\Lambda_{02}}{\Lambda_{01}} &= \prod_{i=R+1}^{p} \{ \frac{l_i{}^{\rho}}{\rho l_i + 1 - \rho} \}^{(M+N)/2} \qquad && \text{if } R \leq p-1, \\ &= 1 && \text{if } R = p. \end{aligned}$$

Since these statistics are functions of l_i's, their distributions depend only on the latent roots $\delta_1 \geq \cdots \geq \delta_p$ of $\boldsymbol{\Phi\Psi}^{-1}$. The hypothesis H_0 reduces to the simple hypothesis $\delta_1 = \cdots = \delta_p = 1$. The hypotheses $H_0^{(r)}$ and H_1 reduce to the composite hypotheses $\delta_1 \geq \cdots \geq \delta_r \geq \delta_{r+1} = \cdots = \delta_p = 1$ and $\delta_1 \geq \cdots \geq \delta_p \geq 1$, respectively. To calculate the significance points for $T_{01}^{(r)}$ and T_{12}, we need the least favorable distributions, which are obtained by the straightforward application of Anderson and Das Gupta (1964).

Define an increasing function $f_{MN}^{+}(\,\cdot\,)$ and a decreasing function $f_{MN}^{-}(\,\cdot\,)$ on $(0, \infty)$ by

$$\begin{aligned} f_{MN}^{+}(x) &= (M+N)\{\log(\rho x + 1 - \rho) - \rho \log x\} I(x > 1), \\ f_{MN}^{-}(x) &= (M+N)\{\log(\rho x + 1 - \rho) - \rho \log x\} I(x < 1), \end{aligned}$$

where $I(\,\cdot\,)$ denotes the indicator function. Then we have

$$-2 \log \Lambda_{01} = \sum_{i=1}^{p} f_{MN}^{+}(l_i), \qquad -2 \log \Lambda_{01}^{(r)} = \sum_{i=r+1}^{p} f_{MN}^{+}(l_i),$$

$$\text{and} \qquad -2 \log \Lambda_{12} = \sum_{i=1}^{p} f_{MN}^{-}(l_i).$$

Then by Theorem 2 of Anderson and Das Gupta (1964) we have that the power function of T_{01} and $T_{01}^{(r)}$, namely $\beta_{01}(\boldsymbol{\delta}) = \mathrm{P}_{\delta}(-2 \log \Lambda_{01} > c)$ and $\beta_{01}^{(r)}(\boldsymbol{\delta}) = \mathrm{P}_{\delta}(-2 \log \Lambda_{01}^{(r)} > c)$ with $c > 0$, are monotonically increasing in each component of $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_p)$. And the power function of the test T_{12}, namely
$\beta_{12}(\boldsymbol{\delta}) = \mathrm{P}_{\delta}(-2 \log \Lambda_{12} > c)$ with $c > 0$, is monotonically decreasing in each of the component of $\boldsymbol{\delta}$. Since $\beta_{01}(\boldsymbol{\delta})$ is increasing, T_{01} is unbiased. Because of the increasing property, $\sup_{H_0^{(r)}} \beta_{01}^{(r)}(\boldsymbol{\delta})$ is attained when $\delta_1, \ldots, \delta_r \uparrow +\infty$ and $\delta_{r+1} = \cdots = \delta_p = 1$; $\inf_{H_1 - H_0^{(r)}} \beta_{01}^{(r)}(\boldsymbol{\delta})$ is attained when $\delta_1, \ldots, \delta_{r+1} \downarrow 1$ and $\delta_{r+2} = \cdots = \delta_p = 1$. Because of the decreasing property, $\sup_{H_1} \beta_{12}(\boldsymbol{\delta})$ is attained at $\delta_1 = \cdots = \delta_p = 1$; $\inf_{H_2 - H_1} \beta_{12}(\boldsymbol{\delta})$ is attained when $\delta_1, \ldots, \delta_{p-1} \uparrow +\infty$

and $\delta_p \uparrow 1$. Theorem 2.1 below determines the distribution when some δ_i's go to infinity.

Theorem 2.1 (Schott and Saw, 1984) Let $d_1 > \cdots > d_p$ be the latent roots of $\boldsymbol{W}_1\boldsymbol{W}_2{}^{-1}$ where $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ are independently distributed according to $W_p(\nu_1, \boldsymbol{\Sigma}_1)$ and $W_p(\nu_2, \boldsymbol{\Sigma}_2)$, respectively, with the latent roots of $\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2{}^{-1}$ being $\delta_1 \geq \cdots \geq \delta_r \geq \delta_{r+1} = \cdots = \delta_p = 1$. Let $d_1^* > \cdots > d_{p-r}^*$ be the latent roots of $\boldsymbol{W}_1^*\boldsymbol{W}_2^{*-1}$ where $\boldsymbol{W}_1^*$ and $\boldsymbol{W}_2^*$ are independently distributed according to $W_{p-r}(\nu_1 - r, \boldsymbol{I}_{p-r})$ and $W_{p-r}(\nu_2, \boldsymbol{I}_{p-r})$ respectively. Then $(d_{r+1}, \ldots, d_p)$ converges to $(d_1^*, \ldots, d_{p-r}^*)$ in distribution as $\delta_1, \ldots, \delta_r \uparrow +\infty$.

We can summarize the properties of the likelihood ratio tests as follows.

Theorem 2.2
(i) The power function of T_{01} is monotonically increasing in the components of $\boldsymbol{\delta}$. T_{01} is unbiased.
(ii) The power function of $T_{01}^{(r)}$ $(0 < r < p)$ is monotonically increasing in the components of $\boldsymbol{\delta}$. The α-significance point $c_{01}^{(r)}(\alpha; p, M, N)$ is determined by

$$\sup_{H_0^{(r)}} \beta_{01}^{(r)}(\boldsymbol{\delta}) = \mathrm{P}(\sum_{i=1}^{p-r} f_{MN}^{+}(l_i^*) > c_{01}^{(r)}(\alpha; p, M, N)) = \alpha,$$

where $l_1^* > \cdots > l_{p-r}^*$ are the latent roots of $(N/M)\boldsymbol{W}_1^*\boldsymbol{W}_2^{*-1}$ with $\boldsymbol{W}_1^*$ and $\boldsymbol{W}_2^*$ being independently distributed according to $W_{p-r}(M - r, \boldsymbol{I}_{p-r})$ and $W_{p-r}(N, \boldsymbol{I}_{p-r})$, respectively. $T_{01}^{(r)}$ is biased since

$$\inf_{H_1 - H_0^{(r)}} \beta_{01}^{(r)}(\boldsymbol{\delta}) = \mathrm{P}_{H_0}(-2\log\Lambda_{01}^{(r)} > c_{01}^{(r)}(\alpha; p, M, N)) < \alpha.$$

(iii) The power function of T_{12} is monotonically decreasing in the components of $\boldsymbol{\delta}$. The α-significance point $c_{12}(\alpha; p, M, N)$ is determined by

$$\sup_{H_1} \beta_{12}(\boldsymbol{\delta}) = \mathrm{P}_{H_0}(-2\log\Lambda_{12} > c_{12}(\alpha; p, M, N)) = \alpha.$$

T_{12} is biased since

$$\inf_{H_2 - H_1} \beta_{12}(\boldsymbol{\delta}) = \mathrm{P}(f_{MN}^{-}(l^{**}) > c_{12}(\alpha; p, M, N)) < \alpha,$$

where l^{**} is a random variable such that $\{M/(M - p + 1)\}l^{**}$ is distributed as the F-distribution F_N^{M-p+1}.

The following statements on the consistency can be established easily.

Theorem 2.3 T_{01}, $T_{01}^{(r)}$ and T_{12} are consistent in the following sense.

(i) $$\lim_{\delta_1 \to +\infty} \beta_{01}(\boldsymbol{\delta}) = 1 \quad \text{uniformly in} \quad \delta_2, \ldots, \delta_p.$$

(ii) $$\lim_{\delta_{r+1} \to +\infty} \beta_{01}^{(r)}(\boldsymbol{\delta}) = 1 \quad \text{uniformly in} \quad \delta_{r+2}, \ldots, \delta_p.$$

(iii) $$\lim_{\delta_p \to +0} \beta_{12}(\boldsymbol{\delta}) = 1 \quad \text{uniformly in} \quad \delta_1, \ldots, \delta_{p-1}.$$

3. Limiting null distribution of LRT.

3.1. Derivation of the limiting null distribution.

Theorem 2.2 (i), (iii) shows that to get the significance points of T_{01} and T_{12}, we only need the distributions of Λ_{01} and Λ_{12} under $H_0 : \delta_1 = \cdots = \delta_p = 1$. We shall derive the limiting joint distribution of Λ_{01} and Λ_{12} under H_0.

We give a lemma which is used in proving Theorems 3.1 and 3.2.

Lemma 3.1 Let $\{f_n(\,\cdot\,)\}$ be a sequence of densities on $\mathbf{R}^p$ with $\lim_{n\to\infty} f_n(\boldsymbol{x}) = f(\boldsymbol{x})$ for each $\boldsymbol{x} \in \mathbf{R}^p$. And let $\{T_n^j(\,\cdot\,),\ 1 \le j \le k\}$ be a sequence of measurable functions on $\mathbf{R}^p$ with $\lim_{n\to\infty} T_n^j(\boldsymbol{x}) = T^j(\boldsymbol{x})$ for each $\boldsymbol{x} \in \mathbf{R}^p$. If $f(\,\cdot\,)$ is a density, then

$$\lim_{n\to\infty} \int \exp\{i \sum_{j=1}^{k} t_j T_n^j(\boldsymbol{x})\} f_n(\boldsymbol{x}) d\boldsymbol{x} = \int \exp\{i \sum_{j=1}^{k} t_j T^j(\boldsymbol{x})\} f(\boldsymbol{x}) d\boldsymbol{x}.$$

(Proof) Note that

$$\begin{aligned} &\left| \int \exp\{i \sum_{j=1}^{k} t_j T_n^j(\boldsymbol{x})\} f_n(\boldsymbol{x}) d\boldsymbol{x} - \int \exp\{i \sum_{j=1}^{k} t_j T^j(\boldsymbol{x})\} f(\boldsymbol{x}) d\boldsymbol{x} \right| \\ &\le \int |f_n(\boldsymbol{x}) - f(\boldsymbol{x})| d\boldsymbol{x} + \int h_n(\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x} \end{aligned}$$

with

$$h_n(\boldsymbol{x}) = \left| \exp\{i \sum_{j=1}^{k} t_j T_n^j(\boldsymbol{x})\} - \exp\{i \sum_{j=1}^{k} t_j T^j(\boldsymbol{x})\} \right| \quad (\le 2).$$

As $n \to \infty$, the first term converges to zero by Scheffé's theorem. The second term converges to zero by the bounded convergence theorem. □

Let $\phi(s,t)$ denote the joint characteristic function of $-2\log\Lambda_{01}$ and $-2\log\Lambda_{12}$ under H_0, i.e. $\phi(s,t) = \mathrm{E}_{H_0} {\Lambda_{01}}^{-2is} {\Lambda_{12}}^{-2it}$. The joint density function of $l_1 > \cdots > l_p > 0$ is given by

$$c(p; M, N) \prod_{i=1}^{p} {l_i}^{(M-p-1)/2} (\frac{M}{M+N} l_i + \frac{N}{M+N})^{-(M+N)/2} \prod_{i<j} (l_i - l_j), \quad (3.1)$$

where

$$c(p;M,N) = \frac{\pi^{p^2/2}\Gamma_p(\frac{M+N}{2})}{\Gamma_p(\frac{M}{2})\Gamma_p(\frac{N}{2})\Gamma_p(\frac{p}{2})} \cdot \frac{M^{pM/2}N^{pN/2}}{(M+N)^{p(M+N)/2}} \tag{3.2}$$

with $\Gamma_p(a) = \pi^{p(p-1)/4}\prod_{i=1}^{p}\Gamma(a-\frac{1}{2}(i-1))$. Define

$$b_i = \sqrt{\frac{MN}{2(M+N)}}(l_i - 1), \quad 1 \le i \le p. \tag{3.3}$$

By letting $M, N \to \infty$ with $\rho = M/(M+N) \to \rho_0$ $(0 \le \rho_0 \le 1)$ and b_i fixed (i.e. $l_i \to 1$), the limit of the joint density of $\boldsymbol{b} = (b_1, \dots, b_p)'$, $b_1 > \cdots > b_p$, is

$$\varphi_0(\boldsymbol{b}) = d(p)\exp\{-\frac{1}{2}\sum_{i=1}^{p} {b_i}^2\}\prod_{i<j}(b_i - b_j) \tag{3.4}$$

with

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

$\varphi_0(\boldsymbol{b})$ in (3.4) turns out to be the density of the latent roots of a $p \times p$ symmetric random matrix $\boldsymbol{A}$ with normal density

$$\frac{1}{2^{p/2}\pi^{p(p+1)/4}}\exp\{-\frac{1}{2}\mathrm{tr}\boldsymbol{A}^2\}. \tag{3.5}$$

(Anderson, 1984b, Theorem 13.3.5.)

On the other hand it is easy to show that

$$\lim_{M,N\to\infty} -2\log\Lambda_{01} = \lim_{M,N\to\infty}\sum_{i=1}^{p} f_{MN}^{+}(l_i) = \sum_{i=1}^{p}(b_i \vee 0)^2,$$

$$\lim_{M,N\to\infty} -2\log\Lambda_{12} = \lim_{M,N\to\infty}\sum_{i=1}^{p} f_{MN}^{-}(l_i) = \sum_{i=1}^{p}(b_i \wedge 0)^2,$$

where $x \vee y$ and $x \wedge y$ are the maximum and the minimum of x and y, respectively. Therefore by Lemma 3.1 we get the characteristic function of the limiting null distribution as

$$\begin{aligned}
\phi_0(s,t) &= \lim_{M,N\to\infty}\phi(s,t) \\
&= \int_{b_1>\cdots>b_p}\exp\{is\sum_{i=1}^{p}(b_i \vee 0)^2 + it\sum_{i=1}^{p}(b_i \wedge 0)^2\}\varphi_0(\boldsymbol{b})d\boldsymbol{b} \\
&= \sum_{r=0}^{p}\int_{B_r\times\overline{B}_{p-r}}\exp\{is\sum_{i=1}^{r}{b_i}^2 + it\sum_{i=r+1}^{p}{b_i}^2\}\varphi_0(\boldsymbol{b})d\boldsymbol{b}
\end{aligned} \tag{3.6}$$

with $d\boldsymbol{b} = \prod_{i=1}^{p} db_i$,

$$B_r = \{(b_1, \ldots, b_r) \mid b_1 > \cdots > b_r > 0\},$$

$$\overline{B}_{p-r} = \{(b_{r+1}, \ldots, b_p) \mid 0 > b_{r+1} > \cdots > b_p\}.$$

By the Laplace expansion of the linkage factor $\prod_{i<j}(b_i - b_j)$, which is the Vandermonde determinant $\det({b_i}^{p-j})_{1 \le i,j \le p}$, (3.6) is

$$d(p) \sum_{r=0}^{p} \sum_{\lambda} (-1)^{\sum_{i=1}^{r}(i+\lambda_i)} \int_{B_r} \exp\{-\frac{1}{2\theta^2} \sum_{i=1}^{r} {b_i}^2\} \det({b_i}^{p-\lambda_j})_{1 \le i,j \le r} db_1 \ldots db_r$$
$$\times \int_{\overline{B}_{p-r}} \exp\{-\frac{1}{2\varphi^2} \sum_{i=1}^{p-r} {b_{i+r}}^2\} \det({b_{i+r}}^{p-\overline{\lambda}_j})_{1 \le i,j \le p-r} db_{r+1} \ldots db_p \qquad (3.7)$$

where $\theta = (1-2is)^{-\frac{1}{2}}$ and $\varphi = (1-2it)^{-\frac{1}{2}}$, $\sum_{\lambda}$ is summation over all combinations of $\lambda_1 < \cdots < \lambda_r$, $\overline{\lambda}_1 < \cdots < \overline{\lambda}_{p-r}$ such that

$$\{\lambda_1, \ldots, \lambda_r, \overline{\lambda}_1, \ldots, \overline{\lambda}_{p-r}\} = \{1, \ldots, p\}.$$

Putting

$$U_k(q_1, \ldots, q_k) = \int_{B_k} \exp\{-\frac{1}{2} \sum_{i=1}^{k} {b_i}^2\} \det({b_i}^{q_j})_{1 \le i,j \le k} db_1 \ldots db_k \qquad (3.8)$$

$(k \ge 1)$ and $U_0 = 1$, we have that (3.7) is

$$d(p) \sum_{r=0}^{p} \sum_{\lambda} U_r(p-\lambda_1, \ldots, p-\lambda_r) U_{p-r}(p-\overline{\lambda}_1, \ldots, p-\overline{\lambda}_{p-r}) \theta^{Q} \varphi^{\overline{Q}} \qquad (3.9)$$

with $Q = \sum_{i=1}^{r}(p-\lambda_i)+r$ and $\overline{Q} = p(p+1)/2 - Q$. (3.9) is a characteristic function of a mixture of the bivariate chi-squared distributions with Q and $\overline{Q}$ $(0 \le Q, \overline{Q} \le p(p+1)/2)$ degrees of freedom. By inverting the limiting characteristic function (3.9), we get the theorem:

Theorem 3.1 As $M, N \to \infty$ with $M/(M+N) \to \rho_0$ $(0 \le \rho_0 \le 1)$, the limiting joint distribution function of $-2\log\Lambda_{01}$ and $-2\log\Lambda_{12}$ under H_0 is given by

$$\lim_{M,N\to\infty} \mathrm{P}_{H_0}(-2\log\Lambda_{01} \le y, -2\log\Lambda_{12} \le z)$$
$$= d(p) \sum_{r=0}^{p} \sum_{p-1 \ge q_1 > \cdots > q_r \ge 0} U_r(q_1, \ldots, q_r) U_{p-r}(\overline{q}_1, \ldots, \overline{q}_{p-r}) G_Q(y) G_{\overline{Q}}(z) \qquad (3.10)$$

where $\overline{q}_1 > \cdots > \overline{q}_{p-r}$ are the members of $\{0,\ldots,p-1\} - \{q_1,\ldots,q_r\}$, $Q = \sum_{k=1}^{r} q_k + r$, $\overline{Q} = p(p+1)/2 - Q$, $G_\nu(\,\cdot\,)$ with $\nu \neq 0$ is the distribution function of the chi-squared distribution with ν degrees of freedom and $G_0(\,\cdot\,) = I(0 \leq \,\cdot\,)$.

Since (3.9) is symmetric in y and z, the limiting marginal distributions of $-2\log\Lambda_{01}$ and $-2\log\Lambda_{12}$ under H_0 are equivalent, which is a mixture of chi-squared distributions.

On the null (least favorable) distribution of $\Lambda_{01}^{(r)}$, Theorem 3.2 below is obtained.

Theorem 3.2 As $M, N \to \infty$ with $M/(M+N) \to \rho_0$ $(0 \leq \rho_0 \leq 1)$,

$$\lim_{M,N\to\infty} \sup_{H_0^{(r)}} \mathrm{P}(-2\log\Lambda_{01}^{(r)} \leq y)$$
$$= d(p_1) \sum_{r'=0}^{p_1} \sum_{p_1-1 \geq q_1 > \cdots > q_{r'} \geq 0} U_{r'}(q_1,\ldots,q_{r'}) U_{p_1-r'}(\overline{q}_1,\ldots,\overline{q}_{p_1-r'}) G_{Q'}(y) \tag{3.11}$$

with $p_1 = p - r$ and $Q' = \sum_{k=1}^{r'} q_k + r'$.

(Proof) Note that

$$\sup_{H_0^{(r)}} \mathrm{P}(-2\log\Lambda_{01}^{(r)} \leq y) = \mathrm{P}(\sum_{i=1}^{p_1} f_{MN}^{+}(l_i^*) \leq y),$$

where l_i^* is defined in Theorem 2.2 (ii). The limiting joint density of

$$b_i^* = \sqrt{\frac{MN}{2(M+N)}}(l_i^* - 1), \quad 1 \leq i \leq p_1,$$

is

$$d(p_1)\exp\{-\frac{1}{2}\sum_{i=1}^{p_1} b_i^{*2}\}\prod_{i<j}(b_i^* - b_j^*).$$

On the other hand

$$\lim_{M,N\to\infty} \sum_{i=1}^{p_1} f_{MN}^{+}(l_i^*) = \sum_{i=1}^{p_1}(b_i^* \vee 0)^2.$$

The rest of the proof is similar to that of Theorem 3.1. □

Remark 3.1 In Section 2 of Chapter I we have already proved that when $\delta_1 \geq \cdots \geq \delta_r > \delta_{r+1} = \cdots = \delta_p = 1$ the limiting null distribution of $\Lambda_{01}^{(r)}$ is (2.5) of Chapter I. This means that

$$\lim_{M,N\to\infty} \mathrm{P}(-2\log\Lambda_{01}^{(r)} \leq y) = \text{R.H.S. of (3.11)}$$

holds when rank $(\boldsymbol{\Phi} - \boldsymbol{\Psi}) = r$.

3.2. Calculation of $\boldsymbol{U}_k$'s.

In the previous subsection our problem has been reduced to calculation of integral $U_k(q_1, \ldots, q_k)$ of (3.8). The method used by Pillai (1956) is exploited here.

Theorem 3.3 $U_k(q_1, \ldots, q_k)$ can be evaluated by the following recurrence formula:

$$\begin{aligned}
&U_k(q_1, \ldots, q_k) \\
=&(-1)^{k-1} U_{k-1}(q_2, \ldots, q_k) I(q_1 = 1) \\
&+ (q_1 - 1) U_k(q_1 - 2, q_2, \ldots, q_k) \\
&+ 2 \sum_{j=2}^{k} (-1)^j \frac{1}{2^{\frac{1}{2}(q_1 + q_j)}} U_1(q_1 + q_j - 1) U_{k-2}(q_2, \ldots, q_{j-1}, q_{j+1}, \ldots, q_k)
\end{aligned} \tag{3.12}$$

$(k \geq 2,\ q_1 \geq 1)$ and

$$\begin{aligned}
U_1(q) =& I(q = 1) + (q - 1) U_1(q - 2) \qquad (q \geq 1), \\
U_1(0) =& \sqrt{\pi/2}.
\end{aligned}$$

(Proof) By expanding the first column of

$$D_k(b_1, \ldots, b_k; q_1, \ldots, q_k) = \det \begin{pmatrix} {b_1}^{q_1} & \cdots & {b_1}^{q_k} \\ \vdots & & \vdots \\ {b_k}^{q_1} & \cdots & {b_k}^{q_k} \end{pmatrix}_{k \times k},$$

we get

$$U_k(q_1, \ldots, q_k) = \sum_{i=1}^{k} (-1)^{i-1} \int_0^\infty \mathrm{e}^{-\frac{1}{2} {b_i}^2} {b_i}^{q_1} db_i F_i(b_i) \tag{3.13}$$

with

$$\begin{aligned}
F_i(b_i) = \int_{b_i}^\infty db_{i-1} \cdots \int_{b_2}^\infty db_1 \int_0^{b_i} db_{i+1} \cdots \int_0^{b_{k-1}} db_k \\
\times \exp\{-\frac{1}{2} \sum_{j \neq i} {b_i}^2\} D_{k-1}(b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_k; q_2, \ldots, q_k).
\end{aligned}$$

By integration by parts, we have

$$\begin{aligned}
&\int_0^\infty \mathrm{e}^{-\frac{1}{2} {b_i}^2} {b_i}^{q_1} db_i F_i(b_i) = -\int_0^\infty (\mathrm{e}^{-\frac{1}{2} {b_i}^2})' {b_i}^{q_1 - 1} F_i(b_i) db_i \\
&= -\mathrm{e}^{-\frac{1}{2} {b_i}^2} {b_i}^{q_1 - 1} F_i(b_i) \Big|_0^\infty + (q_1 - 1) \int_0^\infty \mathrm{e}^{-\frac{1}{2} {b_i}^2} {b_i}^{q_1 - 2} F_i(b_i) db_i \\
&\quad + \int_0^\infty \mathrm{e}^{-\frac{1}{2} {b_i}^2} {b_i}^{q_1 - 1} F_i'(b_i) db_i \\
&= A_i + B_i + C_i, \quad \text{say.}
\end{aligned} \tag{3.14}$$

Noting that $F_i(0) = U_{k-1}(q_2, \ldots, q_k)I(i = k)$, we have

$$\sum_{i=1}^{k}(-1)^{i-1}A_i = (-1)^{k-1}U_{k-1}(q_2, \ldots, q_k)I(q_1 = 1), \tag{3.15}$$

and

$$\sum_{i=1}^{k}(-1)^{i-1}B_i = (q_1 - 1)U_k(q_1 - 2, q_2, \ldots, q_k). \tag{3.16}$$

Next we have

$$\begin{aligned} C_i = &- \int_0^\infty \mathrm{e}^{-\frac{1}{2}{b_i}^2} {b_i}^{q_1-1} db_i \int_{b_i}^\infty db_{i-2} \cdots \int_{b_2}^\infty db_1 \int_0^{b_i} db_{i+1} \cdots \int_0^{b_{k-1}} db_k \\ &\times \exp\{-\frac{1}{2}\sum_{j\neq i-1} {b_j}^2\} D_{k-1}(b_1, \ldots, b_{i-2}, b_i, \ldots, b_k; q_2, \ldots, q_k) \qquad (i \neq 1) \\ &+ \int_0^\infty \mathrm{e}^{-\frac{1}{2}{b_i}^2} {b_i}^{q_1-1} db_i \int_{b_i}^\infty db_{i-1} \cdots \int_{b_2}^\infty db_1 \int_0^{b_i} db_{i+2} \cdots \int_0^{b_{k-1}} db_k \\ &\times \exp\{-\frac{1}{2}\sum_{j\neq i+1} {b_j}^2\} D_{k-1}(b_1, \ldots, b_i, b_{i+2}, \ldots, b_k; q_2, \ldots, q_k) \qquad (i \neq k). \end{aligned}$$

Note that the first term vanishes when $i = 1$ and the second term vanishes when $i = k$. By expanding $D_{k-1}(b_1, \ldots, b_{i-2}, b_i, \ldots, b_k; q_2, \ldots, q_k)$ and $D_{k-1}(b_1, \ldots, b_i, b_{i+2}, \ldots, b_k; q_2, \ldots, q_k)$ in the row which contains b_i, we get

$$\begin{aligned} C_i = &- \sum_{j=2}^{k}(-1)^{i+j} \int_0^\infty \mathrm{e}^{-{b_i}^2} {b_i}^{q_1+q_j-1} db_i \\ &\times \int\!\!\cdots\!\!\int_{\substack{b_1 > \cdots > b_{i-2} > b_i \\ b_i > b_{i+1} > \cdots > b_k > 0}} db_1 \ldots db_{i-2} db_{i+1} \ldots db_k \exp\{-\frac{1}{2}\sum_{j\neq i-1,i} {b_j}^2\} \\ &\times D_{k-2}(b_1, \ldots, b_{i-2}, b_{i+1}, \ldots, b_k; q_2, \ldots, q_{j-1}, q_{j+1}, \ldots, q_k) \\ &+ \sum_{j=2}^{k}(-1)^{i+j+1} \int_0^\infty \mathrm{e}^{-{b_i}^2} {b_i}^{q_1+q_j-1} db_i \\ &\times \int\!\!\cdots\!\!\int_{\substack{b_1 > \cdots > b_{i-1} > b_i \\ b_i > b_{i+2} > \cdots > b_k > 0}} db_1 \ldots db_{i-1} db_{i+2} \ldots db_k \exp\{-\frac{1}{2}\sum_{j\neq i,i+1} {b_j}^2\} \\ &\times D_{k-2}(b_1, \ldots, b_{i-1}, b_{i+2}, \ldots, b_k; q_2, \ldots, q_{j-1}, q_{j+1}, \ldots, q_k). \end{aligned}$$

Noting that, for fixed b,

$$\sum_{i=2}^{k} \int\!\!\cdots\!\!\int_{\substack{b_1 > \cdots > b_{i-2} > b \\ b > b_{i-1} > \cdots > b_{k-2} > 0}} = \sum_{i=1}^{k-1} \int\!\!\cdots\!\!\int_{\substack{b_1 > \cdots > b_{i-1} > b \\ b > b_i > \cdots > b_{k-2} > 0}} = \int\!\!\cdots\!\!\int_{b_1 > \cdots > b_{k-2} > 0},$$

we see that

$$\sum_{i=1}^{k}(-1)^{i-1}C_i = 2\sum_{j=2}^{k}(-1)^j \int_0^\infty \mathrm{e}^{-b_i{}^2} b_i{}^{q_1+q_j-1} db_i \ U_{k-2}(q_2,\ldots,q_{j-1},q_{j+1},\ldots,q_k). \tag{3.17}$$

From (3.13)-(3.17), we derive (3.12). The case of $k=1$ is easy. □

We give the explicit expressions of (3.9) for $1 \le p \le 5$ in Appendix A.

Remark 3.2 Since $U_k(q_1,\ldots,q_k)$ is obviously a skew-symmetric function of $q_1,\ldots,q_k$, we can restrict ourselves to $q_1 > \cdots > q_k$, and the second term of the R.H.S. of (3.12) can be replaced by:

$$\begin{array}{ll} (q_1-1)U_k(q_1-2,q_2,\ldots,q_k) & \text{if } \ q_1-2>q_2, \\ -(q_1-1)U_k(q_2,q_1-2,q_3,\ldots,q_k) & \text{if } \ q_2>q_1-2>q_3, \\ 0 & \text{otherwise.} \end{array}$$

3.3. An expression in terms of Pfaffian.

In the previous subsection we derive a recurrence formula for U_k's which enable us to obtain (3.9) and R.H.S. of (3.10) numerically. We shall give an expression of the limiting characteristic function (3.9) in terms of Pfaffian defined below.

Definition 3.1 (Pfaffian) Let $\boldsymbol{A} = (a_{ij})$ be a $p \times p$ (p:even) skew-symmetric matrix. The *Pfaffian* of $\boldsymbol{A}$ is defined by

$$\mathrm{pf}\boldsymbol{A} = \sum_P \pm a_{i_1 i_2} a_{i_3 i_4} \cdots a_{i_{p-1} i_p},$$

where the summation is taken over all permutations

$$P = \begin{pmatrix} 1 & \cdots & p \\ i_1 & \cdots & i_p \end{pmatrix}$$

with the restrictions

$$i_1 < i_2, i_3 < i_4, \ldots, i_{p-1} < i_p, i_1 < i_3 < \cdots < i_{p-1},$$

and the sign is that of P.

It is well known that $\det\boldsymbol{A} = (\mathrm{pf}\boldsymbol{A})^2$, and that for $p \times p$ matrix $\boldsymbol{B}$,

$$\mathrm{pf}(\boldsymbol{B}\boldsymbol{A}\boldsymbol{B}') = \mathrm{pf}\boldsymbol{A} \cdot \det\boldsymbol{B}. \tag{3.18}$$

See, for example, Mehta (1989). We give another identity which does not appear in the literature. The proof can be easily obtained by Definition 2.1, and is omitted.

Lemma 3.2 For two $p \times p$ (p:even) skew-symmetric matrices $\boldsymbol{A} = (a_{ij})$ and $\boldsymbol{B} = (b_{ij})$,

$$\mathrm{pf}(\boldsymbol{A}+\boldsymbol{B}) = \sum_{\substack{r=0 \\ r:\mathrm{even}}}^{p} \sum_{\lambda} (-1)^{\sum_{i=1}^{r}(i+\lambda_i)} \mathrm{pf}\,(\,\boldsymbol{A}[\boldsymbol{\lambda},\boldsymbol{\lambda}]\,)\,\mathrm{pf}\,(\,\boldsymbol{B}[\overline{\boldsymbol{\lambda}},\overline{\boldsymbol{\lambda}}]\,) \tag{3.19}$$

holds, where $\boldsymbol{A}[\boldsymbol{\lambda},\boldsymbol{\lambda}] = (a_{\lambda_i\lambda_j})_{1\le i,j\le r}$ a $r \times r$ matrix, $\boldsymbol{B}[\overline{\boldsymbol{\lambda}},\overline{\boldsymbol{\lambda}}] = (b_{\overline{\lambda}_i\overline{\lambda}_j})_{1\le i,j\le p-r}$ a $(p-r)\times(p-r)$ matrix, the summation $\sum_{\lambda}$ is taken over the same combinations as that in (3.7).

The following lemma gives another method of evaluating U_k's. The proof is an application of Mehta (1960, Eq.(10)) or Krishnaiah and Chang (1971, Lemma 2.1).

Lemma 3.3

$$\begin{aligned} U_k(q_1,\ldots,q_k) &= \mathrm{pf}\,(\,U_2(q_i,q_j)\,)_{1\le i,j\le k} & \text{if } k \text{ is even,} \\ &= \mathrm{pf}\begin{pmatrix} (U_2(q_i,q_j))_{1\le i,j\le k} & (U_1(q_i))_{1\le i\le k} \\ -(U_1(q_j))_{1\le j\le k} & 0 \end{pmatrix} & \text{if } k \text{ is odd.} \end{aligned}$$

(Proof)
Proof for k even: By integrations with respects to $b_1, b_3, \ldots, b_{k-1}$, we have

$$\begin{aligned} U_k(q_1,\ldots,q_k) &= \int\!\cdots\!\int_{b_2>b_4>\cdots>b_k>0} \exp\{-\frac{1}{2}\sum_{i:\mathrm{even}} {b_i}^2\} db_2 \ldots db_k \\ &\quad \times \det\begin{pmatrix} \int_{b_2}^{\infty} \mathrm{e}^{-\frac{1}{2}{b_1}^2} {b_1}^{q_1} db_1 & \cdots & \int_{b_2}^{\infty} \mathrm{e}^{-\frac{1}{2}{b_1}^2} {b_1}^{q_k} db_1 \\ {b_2}^{q_1} & \cdots & {b_2}^{q_k} \\ \int_{b_4}^{b_2} \mathrm{e}^{-\frac{1}{2}{b_3}^2} {b_3}^{q_1} db_3 & \cdots & \int_{b_4}^{b_2} \mathrm{e}^{-\frac{1}{2}{b_3}^2} {b_3}^{q_k} db_3 \\ \vdots & & \vdots \\ {b_k}^{q_1} & \cdots & {b_k}^{q_k} \end{pmatrix}_{k\times k} \\ &= \int\!\cdots\!\int_{b_2>b_4>\cdots>b_k>0} \exp\{-\frac{1}{2}\sum_{i:\mathrm{even}} {b_i}^2\} db_2 \ldots db_k \\ &\quad \times \det\begin{pmatrix} \int_{b_2}^{\infty} \mathrm{e}^{-\frac{1}{2}{b_1}^2} {b_1}^{q_1} db_1 & \cdots & \int_{b_2}^{\infty} \mathrm{e}^{-\frac{1}{2}{b_1}^2} {b_1}^{q_k} db_1 \\ {b_2}^{q_1} & \cdots & {b_2}^{q_k} \\ \int_{b_4}^{\infty} \mathrm{e}^{-\frac{1}{2}{b_3}^2} {b_3}^{q_1} db_3 & \cdots & \int_{b_4}^{\infty} \mathrm{e}^{-\frac{1}{2}{b_3}^2} {b_3}^{q_k} db_3 \\ \vdots & & \vdots \\ {b_k}^{q_1} & \cdots & {b_k}^{q_k} \end{pmatrix}_{k\times k} \end{aligned} \tag{3.20}$$

Since the determinant in the R.H.S. of (3.20) is a symmetric function of $b_2, b_4, \ldots, b_k$, the integral $\int \cdots \int_{b_2>b_4>\cdots>b_k>0}$ can be replaced by $(1/t!) \int_{b_2=0}^{\infty} \int_{b_4=0}^{\infty} \cdots \int_{b_k=0}^{\infty}$ with $t = k/2$. Dividing k rows into t pairs of 2 rows, and applying (generalized) Laplace expansion, the determinant in the R.H.S. of (3.20) is written as

$$\sum_P{}' \pm \int_{b_2}^{\infty} \mathrm{e}^{-\frac{1}{2}{b_1}^2} db_1 \{ {b_1}^{q_{i_1}} {b_2}^{q_{i_2}} - {b_1}^{q_{i_2}} {b_2}^{q_{i_1}} \} \times \cdots$$
$$\times \int_{b_k}^{\infty} \mathrm{e}^{-\frac{1}{2}{b_{k-1}}^2} db_{k-1} \{ {b_{k-1}}^{q_{i_{k-1}}} {b_k}^{q_{i_k}} - {b_{k-1}}^{q_{i_k}} {b_k}^{q_{i_{k-1}}} \},$$

where the summation $\sum_P'$ is taken over all permutations

$$P = \begin{pmatrix} 1 & \cdots & k \\ i_1 & \cdots & i_k \end{pmatrix} \tag{3.21}$$

with the restrictions

$$i_1 < i_2, i_3 < i_4, \ldots, i_{k-1} < i_k$$

and the sign is that of P. Then we have that

$$\begin{aligned} U_k(q_1, \ldots, q_k) &= \frac{1}{t!} \sum_P{}' \pm U_2(q_{i_1}, q_{i_2}) U_2(q_{i_3}, q_{i_4}) \cdots U_2(q_{i_{k-1}}, q_{i_k}) \\ &= \sum_P \pm U_2(q_{i_1}, q_{i_2}) U_2(q_{i_3}, q_{i_4}) \cdots U_2(q_{i_{k-1}}, q_{i_k}) \end{aligned}$$

where the summation $\sum_P$ is taken over all permutations P of (3.21) with the restrictions

$$i_1 < i_2, i_3 < i_4, \ldots, i_{k-1} < i_k, i_1 < i_3 < \cdots < i_{k-1}.$$

This completes the proof (for k even).

Proof for k odd: As in the even case, it holds that

$$\begin{aligned} U_k(q_1, \ldots, q_k) = & \int \cdots \int_{b_2>b_4>\cdots>b_{k-1}>0} \exp\{-\frac{1}{2} \sum_{i:\text{even}} {b_i}^2\} db_2 \ldots db_{k-1} \\ & \times \det \begin{pmatrix} \int_{b_2}^{\infty} \mathrm{e}^{-\frac{1}{2}{b_1}^2} {b_1}^{q_1} db_1 & \cdots & \int_{b_2}^{\infty} \mathrm{e}^{-\frac{1}{2}{b_1}^2} {b_1}^{q_k} db_1 \\ {b_2}^{q_1} & \cdots & {b_2}^{q_k} \\ \int_{b_4}^{\infty} \mathrm{e}^{-\frac{1}{2}{b_3}^2} {b_3}^{q_1} db_3 & \cdots & \int_{b_4}^{\infty} \mathrm{e}^{-\frac{1}{2}{b_3}^2} {b_3}^{q_k} db_3 \\ \vdots & & \vdots \\ {b_{k-1}}^{q_1} & \cdots & {b_{k-1}}^{q_k} \\ \int_0^{\infty} \mathrm{e}^{-\frac{1}{2}{b_k}^2} {b_k}^{q_1} db_k & \cdots & \int_0^{\infty} \mathrm{e}^{-\frac{1}{2}{b_k}^2} {b_k}^{q_k} db_k \end{pmatrix}_{k \times k} \end{aligned} \tag{3.22}$$

The integral $\int \cdots \int_{b_2>b_4>\cdots>b_{k-1}>0}$ can be also replaced by $(1/t!) \int_{b_2=0}^{\infty} \cdots \int_{b_{k-1}=0}^{\infty}$ with $t = (k-1)/2$, since the determinant in the R.H.S. of (3.22) is a symmetric function of $b_2, b_4, \ldots, b_{k-1}$. Dividing k rows of this determinant into t pairs of 2 rows and 1 row, and applying (generalized) Laplace expansion, we see that

$$\begin{aligned} U_k(q_1, \ldots, q_k) &= \frac{1}{t!}\sum_P{}' \pm U_2(q_{i_1}, q_{i_2})U_2(q_{i_3}, q_{i_4}) \cdots U_2(q_{i_{k-2}}, q_{i_{k-1}})U_1(q_{i_k}) \\ &= \sum_P \pm U_2(q_{i_1}, q_{i_2})U_2(q_{i_3}, q_{i_4}) \cdots U_2(q_{i_{k-2}}, q_{i_{k-1}})U_1(q_{i_k}), \end{aligned}$$

where the summations $\sum_P'$ and $\sum_P$ are taken over all permutations P of (3.21) with the restrictions

$$i_1 < i_2, i_3 < i_4, \ldots, i_{k-2} < i_{k-1},\ i_k : \text{free}$$

and

$$i_1 < i_2, i_3 < i_4, \ldots, i_{k-2} < i_{k-1}, i_1 < i_3 < \cdots < i_{k-2},\ i_k : \text{free},$$

respectively. This completes the proof (for k odd). □

Remark 3.3 Note that

$$U_1(q) = 2^{(q-1)/2}\Gamma(\frac{q+1}{2}),$$

and that $U_2(q_1, q_2)$ $(q_1 > q_2 \geq 0)$ is evaluated with the recurrence formula of Theorem 3.3 as

$$\begin{aligned} U_2(q_1, q_2) =& \Gamma(\frac{q_1+q_2}{2}) \\ &+ \begin{cases} (q_1-1)U_2(q_1-2, q_2) & \text{if } q_1 > q_2+2 \\ 0 & \text{if } q_1 = q_2+2 \\ -(q_1-1)U_2(q_2, q_1-2) & \text{if } q_1 = q_2+1 \geq 2 \\ -\sqrt{\pi/2} & \text{if } q_1 = q_2+1 = 1. \end{cases} \end{aligned}$$

Using the lemmas provided above we get the result.

Theorem 3.4 The characteristic function $\phi_0(s,t)$ of (3.9) is expressed as

$$\begin{aligned} & d(p)\,\mathrm{pf}\begin{pmatrix} \theta^2\boldsymbol{D}(\theta)\boldsymbol{E}\boldsymbol{D}(\theta) - \varphi^2\boldsymbol{D}(-\varphi)\boldsymbol{E}\boldsymbol{D}(-\varphi) & \varphi\boldsymbol{D}(-\varphi)\boldsymbol{f} & \theta\boldsymbol{D}(\theta)\boldsymbol{f} \\ -\varphi\boldsymbol{f}'\boldsymbol{D}(-\varphi) & 0 & 1 \\ -\theta\boldsymbol{f}'\boldsymbol{D}(\theta) & -1 & 0 \end{pmatrix}_{(p+2)\times(p+2)} \\ &= d(p)\,\mathrm{pf}\begin{pmatrix} \{\theta^{2p-i-j+2} - (-\varphi)^{2p-i-j+2}\}e_{ij} & -(-\varphi)^{p-i+1}f_i & \theta^{p-i+1}f_i \\ (-\varphi)^{p-j+1}f_j & 0 & 1 \\ -\theta^{p-j+1}f_j & -1 & 0 \end{pmatrix} \qquad (3.23) \end{aligned}$$

if p is even,

$$d(p)\,\mathrm{pf}\begin{pmatrix} \theta^2\boldsymbol{D}(\theta)\boldsymbol{E}\boldsymbol{D}(\theta)-\varphi^2\boldsymbol{D}(-\varphi)\boldsymbol{E}\boldsymbol{D}(-\varphi) & \{\theta\boldsymbol{D}(\theta)+\varphi\boldsymbol{D}(-\varphi)\}\boldsymbol{f} \\ -\boldsymbol{f}'\{\theta\boldsymbol{D}(\theta)+\varphi\boldsymbol{D}(-\varphi)\} & 0 \end{pmatrix}_{(p+1)\times(p+1)}$$
$$= d(p)\,\mathrm{pf}\begin{pmatrix} \{\theta^{2p-i-j+2}-(-\varphi)^{2p-i-j+2}\}e_{ij} & \{\theta^{p-i+1}-(-\varphi)^{p-i+1}\}f_i \\ -\{\theta^{p-j+1}-(-\varphi)^{p-j+1}\}f_j & 0 \end{pmatrix} \tag{3.24}$$

if p is odd, where $\boldsymbol{E}$ is a $p\times p$ matrix with (i,j)th element $e_{ij}=U_2(p-i,p-j)$, $\boldsymbol{f}$ is a $p\times 1$ vector with ith element $f_i=U_1(p-i)$, and $\boldsymbol{D}(\xi)=\mathrm{diag}(\xi^{p-i})_{1\le i\le p}$ is a $p\times p$ diagonal matrix.

(Proof)

Proof for p even: By the definition of the Pfaffian in Definition 3.1 as well as (3.18) and (3.19), the Pfaffian in (3.23) is expanded as

$$\begin{aligned}
&\mathrm{pf}\{\theta^2\boldsymbol{D}(\theta)\boldsymbol{E}\boldsymbol{D}(\theta)-\varphi^2\boldsymbol{D}(-\varphi)\boldsymbol{E}\boldsymbol{D}(-\varphi)\} \\
&\quad+\mathrm{pf}\left\{\begin{pmatrix} \theta^2\boldsymbol{D}(\theta)\boldsymbol{E}\boldsymbol{D}(\theta) & \mathbf{0}_p & \theta\boldsymbol{D}(\theta)\boldsymbol{f} \\ \mathbf{0}_p{}' & 0 & 0 \\ -\theta\boldsymbol{f}'\boldsymbol{D}(\theta) & 0 & 0 \end{pmatrix}\right. \\
&\qquad\left.-\begin{pmatrix} \varphi^2\boldsymbol{D}(-\varphi)\boldsymbol{E}\boldsymbol{D}(-\varphi) & -\varphi\boldsymbol{D}(-\varphi)\boldsymbol{f} & \mathbf{0}_p \\ \varphi\boldsymbol{f}'\boldsymbol{D}(-\varphi) & 0 & 0 \\ \mathbf{0}_p{}' & 0 & 0 \end{pmatrix}\right\} \\
&=\sum_{r:\mathrm{even}}\sum_{\lambda}\mathrm{pf}\,(\,\boldsymbol{E}[\boldsymbol{\lambda},\boldsymbol{\lambda}]\,)\,\mathrm{pf}\,(\,\boldsymbol{E}[\overline{\boldsymbol{\lambda}},\overline{\boldsymbol{\lambda}}]\,)\,\theta^Q\varphi^{\overline{Q}} \\
&\quad+\sum_{r:\mathrm{odd}}\sum_{\lambda}\mathrm{pf}\begin{pmatrix} \boldsymbol{E}[\boldsymbol{\lambda},\boldsymbol{\lambda}] & \boldsymbol{f}[\boldsymbol{\lambda}] \\ -\boldsymbol{f}[\boldsymbol{\lambda}]' & 0 \end{pmatrix}\mathrm{pf}\begin{pmatrix} \boldsymbol{E}[\overline{\boldsymbol{\lambda}},\overline{\boldsymbol{\lambda}}] & \boldsymbol{f}[\overline{\boldsymbol{\lambda}}] \\ -\boldsymbol{f}[\overline{\boldsymbol{\lambda}}]' & 0 \end{pmatrix}\theta^Q\varphi^{\overline{Q}},
\end{aligned} \tag{3.25}$$

with $\mathbf{0}_p=(0,\dots,0)'$ $p\times 1$ zero vector, $Q=\sum_{i=1}^r(p-\lambda_i)+r$, $\overline{Q}=p(p+1)/2-Q$, $\boldsymbol{f}[\boldsymbol{\lambda}]=(f_{\lambda_i})_{1\le i\le r}$ an $r\times 1$ vector, $\boldsymbol{f}[\overline{\boldsymbol{\lambda}}]=(f_{\overline{\lambda}_i})_{1\le i\le p-r}$ a $(p-r)\times 1$ vector. Combining (3.25) and Lemma 3.3 we see that (3.23) is equal to (3.9).

Proof for p odd: As in the even case the Pfaffian in (3.24) is expanded as

$$\begin{aligned}
&\mathrm{pf}\left\{\begin{pmatrix} \theta^2\boldsymbol{D}(\theta)\boldsymbol{E}\boldsymbol{D}(\theta) & \theta\boldsymbol{D}(\theta)\boldsymbol{f} \\ -\theta\boldsymbol{f}'\boldsymbol{D}(\theta) & 0 \end{pmatrix}-\begin{pmatrix} \varphi^2\boldsymbol{D}(-\varphi)\boldsymbol{E}\boldsymbol{D}(-\varphi) & -\varphi\boldsymbol{D}(-\varphi)\boldsymbol{f} \\ \varphi\boldsymbol{f}'\boldsymbol{D}(-\varphi) & 0 \end{pmatrix}\right\} \\
&=\sum_{r:\mathrm{odd}}\sum_{\lambda}\mathrm{pf}\begin{pmatrix} \boldsymbol{E}[\boldsymbol{\lambda},\boldsymbol{\lambda}] & \boldsymbol{f}[\boldsymbol{\lambda}] \\ -\boldsymbol{f}[\boldsymbol{\lambda}]' & 0 \end{pmatrix}\mathrm{pf}\,(\,\boldsymbol{E}[\overline{\boldsymbol{\lambda}},\overline{\boldsymbol{\lambda}}]\,)\,\theta^Q\varphi^{\overline{Q}} \\
&\quad+\sum_{r:\mathrm{even}}\sum_{\lambda}\mathrm{pf}\,(\,\boldsymbol{E}[\boldsymbol{\lambda},\boldsymbol{\lambda}]\,)\,\mathrm{pf}\begin{pmatrix} \boldsymbol{E}[\overline{\boldsymbol{\lambda}},\overline{\boldsymbol{\lambda}}] & \boldsymbol{f}[\overline{\boldsymbol{\lambda}}] \\ -\boldsymbol{f}[\overline{\boldsymbol{\lambda}}]' & 0 \end{pmatrix}\theta^Q\varphi^{\overline{Q}}.
\end{aligned} \tag{3.26}$$

Combining (3.26) and Lemma 3.3 we see that (3.24) is equal to (3.9). The proof is completed. □

4. Asymptotic expansion of the null distribution of LRT.

The formal asymptotic expansions of the null distributions of the likelihood ratio statistics shall be derived here. The linkage factor $\prod_{i<j}(l_i - l_j)$ can be represented as $D_p(l_1-1,\ldots,l_p-1;p-1,\ldots,0)$. By the Laplace expansion of the linkage factor in (3.1), the joint characteristic function $\phi(s,t) = \mathrm{E}_{H_0}{\Lambda_{01}}^{-2is}{\Lambda_{12}}^{-2it}$ is given by

$$\phi(s,t) = c(p;M,N)\sum_{r=0}^{p}\sum_{p-1\geq q_1>\cdots>r_r\geq 0} V_r(\theta;q_1,\ldots,q_r)\tilde{V}_{p-r}(\varphi;q_{r+1},\ldots,q_p) \quad (4.1)$$

where $\theta = (1-2is)^{-\frac{1}{2}}$, $\varphi = (1-2it)^{-\frac{1}{2}}$,

$$V_r(\theta;q_1,\ldots,q_r) = \int\cdots\int_{l_1>\cdots>l_r>1}\prod_{i=1}^{r}\{\frac{{l_i}^{\rho}}{\rho l_i+1-\rho}\}^{\frac{1}{2}(M+N)\theta^{-2}}{l_i}^{-\frac{1}{2}(p+1)} \times D_r(l_1-1,\ldots,l_r-1;q_1,\ldots,q_r)dl_1\ldots dl_r$$

and

$$\begin{aligned}\tilde{V}_{p-r}(\theta;q_{r+1},\ldots,q_p) =& (-1)^{r(r+1)/2+\sum_{i=1}^{r}(p-q_i)} \\ &\times \int\cdots\int_{1>l_{r+1}>\cdots>l_p>0}\prod_{i=r+1}^{p}\{\frac{{l_i}^{\rho}}{\rho l_i+1-\rho}\}^{\frac{1}{2}(M+N)\varphi^{-2}}{l_i}^{-\frac{1}{2}(p+1)} \\ &\times D_{p-r}(l_{r+1}-1,\ldots,l_p-1;q_{r+1},\ldots,q_p)dl_{r+1}\ldots dl_p.\end{aligned}$$

Define $b_i,\ 1\leq i\leq p$, by (3.3). Assume that

$$\rho = \frac{M}{M+N} = \rho_0 + \mathrm{O}\{(\frac{MN}{M+N})^{-1}\} = \rho_0 + \mathrm{O}(\frac{1}{\rho N}).$$

By expanding around $b_i = 0$ we have

$$\begin{aligned}&\prod_{i=1}^{r}\{\frac{{l_i}^{\rho}}{\rho l_i+1-\rho}\}^{\frac{1}{2}(M+N)\theta^{-2}}{l_i}^{-\frac{1}{2}(p+1)} \\ &= \exp\{-\frac{\theta^{-2}}{2}\sum_i {b_i}^2\} \\ &\quad\times [1+\frac{1}{\sqrt{\rho N}}\{\frac{\sqrt{2}(1+\rho_0)}{3}\theta^{-2}\sum_i {b_i}^3 - \frac{p+1}{\sqrt{2}}\sum_i b_i\} \\ &\quad + \frac{1}{\rho N}\{-\frac{1+\rho_0+{\rho_0}^2}{2}\theta^{-2}\sum_i {b_i}^4 + \frac{p+1}{2}\sum_i {b_i}^2 + \frac{(1+\rho_0)^2}{9}\theta^{-4}(\sum_i {b_i}^3)^2 \\ &\qquad - \frac{(1+\rho_0)(p+1)}{3}\theta^{-2}(\sum_i {b_i}^3)(\sum_i b_i) + \frac{(p+1)^2}{4}(\sum_i b_i)^2\} \\ &\quad + \mathrm{O}\{\frac{1}{(\rho N)^{3/2}}\}].\end{aligned}$$

Using the relation

$$(\sum_{i=1}^{k} {b_i}^m) D_k(b_1,\dots,b_k;q_1,\dots,q_k) = \sum_{i=1}^{k} D_k(b_1,\dots,b_k;q_1,\dots,q_{i-1},q_i+m,q_{i+1},\dots,q_k),$$

V_r is evaluated as

$$\{\frac{MN}{2(M+N)}\}^{Q/2} V_r(\theta;q_1,\dots,q_r) = U_r\theta^Q + \frac{1}{\sqrt{\rho N}}\overline{U}_r\theta^{Q+1} + \frac{1}{\rho N}\overline{\overline{U}}_r\theta^{Q+2} + \mathrm{O}\{\frac{1}{(\rho N)^{3/2}}\}, \tag{4.2}$$

where $Q = \sum_{i=1}^{r} q_i + r$,

$$\overline{U}_k = \frac{\sqrt{2}(1+\rho_0)}{3} U_k^{(3)} - \frac{p+1}{\sqrt{2}} U_k^{(1)}, \tag{4.3}$$

$$\overline{\overline{U}}_k = -\frac{1+\rho_0+{\rho_0}^2}{2} U_k^{(4)} + \frac{p+1}{2} U_k^{(2)} + \frac{(1+\rho_0)^2}{9} U_k^{(3,3)} - \frac{(1+\rho_0)(p+1)}{3} U_k^{(3,1)} + \frac{(p+1)^2}{4} U_k^{(1,1)} \tag{4.4}$$

with

$$U_k^{(m)} = \sum_{i=1}^{k} U_k(q_1,\dots,q_{i-1},q_i+m,q_{i+1},\dots,q_k),$$

$$U_k^{(m,n)} = \sum_{i=1}^{k} U_k(q_1,\dots,q_{i-1},q_i+m+n,q_{i+1},\dots,q_k) + \sum_{i\neq j} U_k(q_1,\dots,q_i+m,\dots,q_j+n,\dots,q_k).$$

Because the contribution of the integral over the region

$$\{0 > b_{r+1} > \cdots > b_p > -\infty\} - \{0 > b_{r+1} > \cdots > b_p > -\sqrt{MN/2(M+N)}\}$$

is $\mathrm{o}\{(\rho N)^{-m}\}$ for any m, $\tilde{V}_{p-r}$ is also evaluated as

$$\{\frac{MN}{2(M+N)}\}^{\overline{Q}/2} \tilde{V}_{p-r}(\varphi;q_{r+1},\dots,q_p) = U_{p-r}\varphi^{\overline{Q}} - \frac{1}{\sqrt{\rho N}}\overline{U}_{p-r}\varphi^{\overline{Q}+1} + \frac{1}{\rho N}\overline{\overline{U}}_{p-r}\varphi^{\overline{Q}+2} + \mathrm{O}\{\frac{1}{(\rho N)^{3/2}}\}, \tag{4.5}$$

where $\overline{Q} = p(p+1)/2 - Q$. Applying Stirling's formula to the constant factor $c(p; M, N)$ of (3.2), we have

$$\begin{aligned} &\{\frac{2(M+N)}{MN}\}^{p(p+1)/4} c(p; M, N) \\ &= d(p)\{1 - \frac{1}{24\rho N} p(2p^2 + 3p - 1)(1 - \rho_0 + {\rho_0}^2) + \mathrm{O}\{\frac{1}{(\rho N)^2}\}\}. \end{aligned} \tag{4.6}$$

By inverting the asymptotic expansion of $\phi(s, t)$ derived from (4.1)-(4.6), we have the following theorem.

Theorem 4.1 The asymptotic expansion of joint distribution function of $-2 \log \Lambda_{01}$ and $-2 \log \Lambda_{12}$ under H_0 is given by

$$\begin{aligned} &\mathrm{P}_{H_0}(-2 \log \Lambda_{01} \leq y, -2 \log \Lambda_{12} \leq z) \\ =& d(p)\big[\sum U_r U_{p-r} G_Q(y) G_{\overline{Q}}(z) \\ &+ \frac{1}{\sqrt{\rho N}} \sum \{\overline{U}_r U_{p-r} G_{Q+1}(y) G_{\overline{Q}}(z) - U_r \overline{U}_{p-r} G_Q(y) G_{\overline{Q}+1}(z)\} \\ &+ \frac{1}{\rho N} \sum \{\overline{\overline{U}}_r U_{p-r} G_{Q+2}(y) G_{\overline{Q}}(z) - \overline{U}_r \overline{U}_{p-r} G_{Q+1}(y) G_{\overline{Q}+1}(z) \\ &\quad + U_r \overline{\overline{U}}_{p-r} G_Q(y) G_{\overline{Q}+2}(z) - \frac{1}{24} p(2p^2 + 3p - 1)(1 - \rho_0 + {\rho_0}^2) G_Q(y) G_{\overline{Q}}(z)\}\big] \\ &+ \mathrm{O}\{\frac{1}{(\rho N)^{3/2}}\}. \end{aligned}$$

Here the arguments of U_r, $\overline{U}_r$, $\overline{\overline{U}}_r$ (i.e. $q_1, \ldots, q_r$) and U_{p-r}, $\overline{U}_{p-r}$, $\overline{\overline{U}}_{p-r}$ (i.e. $\overline{q}_1, \ldots, \overline{q}_{p-r}$) are omitted for simplicity; put $U_0 = 1$, $\overline{U}_0 = 0$ and $\overline{\overline{U}}_0 = 0$ formally; the summation is over $0 \leq r \leq p$ and $p - 1 \geq q_1 > \cdots > q_r \geq 0$.

The example for $p = 2$ is given in Appendix A. The asymptotic expansion of $\sup_{H_0^{(r)}} \mathrm{P}(-2 \log \Lambda_{01}^{(r)} \leq y)$ can be derived similarly.

5. Significance points and biases.

By Theorem 2.2 and the results of Section 3, we can give the limiting significance points and biases as $M, N \to \infty$ with $M/(M+N) \to \rho_0$.

Table 5.1 shows the limiting α-significance points $c_{01}(\alpha; p, M, N)$ and $c_{12}(\alpha; p, M, N)$, that is $d(\alpha; p)$ satisfying

$$\lim_{M,N \to \infty} \mathrm{P}_{H_0}(-2 \log \Lambda_{01} > d(\alpha; p)) = \mathrm{P}(\sum_{i=1}^{p} (b_i \vee 0)^2 > d(\alpha; p)) = \alpha,$$

Table 5.1

Limiting significance points of T_{01} and T_{12}

$1-\alpha \setminus p$	2	3	4	5	6	7	8	9	10
0.010	0.0000	0.0000	0.0986	0.6832	1.7747	3.3713	5.4720	8.0757	11.1817
0.025	0.0000	9.1e-7	0.3178	1.1788	2.5559	4.4424	6.8352	9.7322	13.1324
0.050	0.0000	0.0508	0.6207	1.7319	3.3617	5.5021	8.1492	11.3010	14.9559
0.100	0.0000	0.2423	1.1146	2.5241	4.4511	6.8881	9.8313	13.2787	17.2291
0.250	0.0957	0.9190	2.3181	4.2400	6.6723	9.6108	13.0531	16.9980	21.4446
0.500	0.7717	2.2605	4.2581	6.7572	9.7567	13.2563	17.2561	21.7559	26.7557
0.750	2.1535	4.2858	6.8788	9.9539	13.5198	17.5803	22.1373	27.1920	32.7450
0.900	4.0457	6.7324	9.8503	13.4372	17.5088	22.0709	27.1272	32.6794	38.7288
0.950	5.4845	8.4904	11.9156	15.8047	20.1758	25.0360	30.3894	36.2381	42.5835
0.975	6.9229	10.1978	13.8848	18.0330	22.6618	27.7789	33.3888	39.4936	46.0949
0.990	8.8211	12.3994	16.3852	20.8310	25.7566	31.1703	37.0766	43.4779	50.3756

Table 5.2

Limiting biases of $T_{01}^{(p-1)}$ and T_{12}

($\alpha = 5 \times 10^{-2}$)

p	$T_{01}^{(p-1)}$	T_{12}
2	8.64×10^{-4}	9.59×10^{-3}
3	5.72×10^{-6}	1.79×10^{-3}
4	1.54×10^{-8}	2.78×10^{-4}
5	1.75×10^{-11}	3.51×10^{-5}

where $b_1 > \cdots > b_p$ are distributed with density (3.4). Put $d(\alpha;p)=0$ when

$$\mathrm{P}(\sum_{i=1}^{p}(b_i \vee 0)^2 > 0) < \alpha.$$

Table 5.2 shows the limiting biases of the tests $T_{01}^{(p-1)}$ and T_{12} when $\alpha = 0.05$. The first column of Table 5.2 gives

$$\lim_{M,N\to\infty} \inf_{H_1 - H_0^{(p-1)}} \beta_{01}^{(p-1)}(\boldsymbol{\delta}) = \mathrm{P}((b_p \vee 0)^2 > d(\alpha;1)).$$

The method of calculating the distribution function of b_p is given in Appendix B. The second column of Table 5.2 is

$$\lim_{M,N\to\infty} \inf_{H_2 - H_1} \beta_{12}(\boldsymbol{\delta}) = \mathrm{P}((b^{**} \wedge 0)^2 > d(\alpha;p)),$$

where b^{**} is distributed according to $N(0,1)$. Table 5.2 shows that the degree of biases of two tests are very high.

6. Power comparisons.

In this section we study a Monte Carlo simulation to compare the powers of several tests including the LRT. The null hypothesis is the equality of two covariance matrices $\boldsymbol{\Phi} = \boldsymbol{\Psi}$, and the alternative is the local hypothesis that

$$\boldsymbol{\Phi} = \boldsymbol{\Psi}^{1/2}\left(\boldsymbol{I}_p + \sqrt{\frac{2(M+N)}{MN}}\boldsymbol{\Delta}\right)\boldsymbol{\Psi}^{1/2}$$

with $\boldsymbol{\Delta} = \mathrm{diag}(\delta_i)_{p\times p}$, $\delta_i \geq 0$. We compare the limiting powers ($M, N \to \infty$ with $M/(M+N) \to \rho_0$) of four test criteria:

- One-sided likelihood ratio test based on the statistic Λ_{01} [ONE];
- Two-sided likelihood ratio test based on the statistic Λ_{02} [TWO];
- Roy's test based on the largest root of $\boldsymbol{H}\boldsymbol{G}^{-1}$ [ROY]; and
- Locally most powerful test of Giri (1968) based on the statistic $\mathrm{tr}\boldsymbol{H}(\boldsymbol{H}+\boldsymbol{G})^{-1}$ [LMP].

The limiting power functions of the four tests are

$$\Pr(\sum_{i=0}^{p}(b_i \vee 0)^2 > c), \quad \Pr(\sum_{i=0}^{p} {b_i}^2 > c'), \quad \Pr(b_1 > c''), \quad \text{and} \quad \Pr(\sum_{i=0}^{p} b_i > c'''),$$

respectively, where $b_1 > \cdots > b_p$ are the latent roots of $p \times p$ symmetric random matrix $\boldsymbol{A}$ with normal density

$$\frac{1}{2^{p/2}\pi^{p^2/2}}\mathrm{tr}(\boldsymbol{A} - \boldsymbol{\Delta})^2.$$

The average powers of the four tests over 100000 replications are given in Table 6.1 (for $p = 4$, size= 5%) and Table 6.2 (for $p = 8$, size= 5%). The results indicate that: ONE has high power when rank $\boldsymbol{\Delta}$ is small; LMP has high power when rank $\boldsymbol{\Delta}$ is large; the character of ROY is similar to ONE, however, ONE is more powerful than ROY even when rank $\boldsymbol{\Delta} = 1$; TWO is inferior to the one-sided tests ONE, ROY and LMP.

Table 6.1

Power (%, $p = 4$, size=5%)

δ_1	δ_2	δ_3	δ_4	ONE	TWO	ROY	LMP
0.0	0.0	0.0	0.0	5	5	5	5
1.0	0.0	0.0	0.0	<u>13</u>	8	12	<u>13</u>
2.0	0.0	0.0	0.0	<u>34</u>	22	32	26
4.0	0.0	0.0	0.0	90	79	<u>91</u>	64
1.0	1.0	0.0	0.0	24	12	20	<u>26</u>
2.0	2.0	0.0	0.0	<u>66</u>	43	56	64
4.0	4.0	0.0	0.0	<u>100</u>	99	99	99
1.0	1.0	1.0	0.0	36	17	28	<u>45</u>
2.0	2.0	2.0	0.0	87	64	73	<u>91</u>
1.0	1.0	1.0	1.0	49	22	37	<u>64</u>
1.0	0.5	0.0	0.0	17	9	15	<u>19</u>
2.0	1.0	0.0	0.0	<u>46</u>	27	40	45
4.0	2.0	0.0	0.0	<u>97</u>	89	94	91
1.5	1.0	0.5	0.0	39	19	31	<u>45</u>
3.0	2.0	1.0	0.0	<u>91</u>	72	82	<u>91</u>

Underline denotes largest value in each row.

Table 6.2

Power (%, $p = 8$, size=5%)

δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	ONE	TWO	ROY	LMP
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5	5	5	5
2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	<u>21</u>	12	19	17
4.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	71	51	<u>74</u>	41
2.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	<u>46</u>	23	35	41
4.0	4.0	0.0	0.0	0.0	0.0	0.0	0.0	<u>98</u>	89	95	88
1.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	24	10	17	<u>28</u>
2.0	2.0	2.0	0.0	0.0	0.0	0.0	0.0	<u>69</u>	37	50	68
4.0	4.0	4.0	0.0	0.0	0.0	0.0	0.0	<u>100</u>	99	99	99
1.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	44	15	27	<u>55</u>
2.0	2.0	2.0	2.0	2.0	0.0	0.0	0.0	94	64	73	<u>97</u>
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	71	24	43	<u>88</u>
4.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	<u>86</u>	64	81	68
4.0	3.0	2.0	1.0	0.0	0.0	0.0	0.0	<u>98</u>	86	92	97

Underline denotes largest value in each row.

Appendix A. Examples of the limiting distributions.

The limiting characteristic functions $\phi_0(s,t)$ for $1 \le p \le 5$ are presented as follows. ($\theta = (1-2is)^{-\frac{1}{2}}$, $\varphi = (1-2it)^{-\frac{1}{2}}$)

$p =1:$

$$\frac{1}{2}(\theta+\varphi)$$

$p =2:$

$$(-\frac{1}{2\sqrt{2}}+\frac{1}{2})(\theta^3+\varphi^3)+\frac{1}{2\sqrt{2}}(\theta^2\varphi+\theta\varphi^2)$$

$p =3:$

$$(-\frac{1}{\sqrt{2}\pi}+\frac{1}{4})(\theta^6+\varphi^6)+(\frac{1}{2\sqrt{2}}-\frac{1}{4})(\theta^5\varphi+\theta\varphi^5)+\frac{1}{\sqrt{2}\pi}(\theta^4\varphi^2+\theta^2\varphi^4)$$
$$+(-\frac{1}{\sqrt{2}}+1)\theta^3\varphi^3$$

$p =4:$

$$(-\frac{1}{2\pi}-\frac{1}{8\sqrt{2}}+\frac{1}{4})(\theta^{10}+\varphi^{10})+(\frac{3}{8\sqrt{2}}-\frac{1}{4})(\theta^9\varphi+\theta\varphi^9)$$
$$+(\frac{1}{\pi}-\frac{3}{8\sqrt{2}})(\theta^8\varphi^2+\theta^2\varphi^8)+(-\frac{17}{8\sqrt{2}}+\frac{13}{8})(\theta^7\varphi^3+\theta^3\varphi^7)$$
$$+(-\frac{1}{2\pi}+\frac{1}{2\sqrt{2}})(\theta^6\varphi^4+\theta^4\varphi^6)+(\frac{7}{2\sqrt{2}}-\frac{9}{4})\theta^5\varphi^5$$

$p =5:$

$$(-\frac{1}{12\sqrt{2}\pi}-\frac{1}{3\pi}+\frac{1}{8})(\theta^{15}+\varphi^{15})+(-\frac{2}{3\pi}+\frac{1}{8\sqrt{2}}+\frac{1}{8})(\theta^{14}\varphi+\theta\varphi^{14})$$
$$+(-\frac{11}{12\sqrt{2}\pi}+\frac{2}{3\pi})(\theta^{13}\varphi^2+\theta^2\varphi^{13})$$
$$+(-\frac{8}{3\sqrt{2}\pi}+\frac{8}{3\pi}-\frac{1}{2\sqrt{2}}+\frac{1}{8})(\theta^{12}\varphi^3+\theta^3\varphi^{12})$$
$$+(-\frac{1}{4\sqrt{2}\pi}+\frac{1}{3\pi})(\theta^{11}\varphi^4+\theta^4\varphi^{11})+(\frac{6}{\sqrt{2}\pi}-\frac{5}{\pi}+\frac{13}{8\sqrt{2}}-\frac{13}{16})(\theta^{10}\varphi^5+\theta^5\varphi^{10})$$
$$+(\frac{17}{4\sqrt{2}\pi}-\frac{3}{\pi}+\frac{1}{\sqrt{2}}-\frac{9}{16})(\theta^9\varphi^6+\theta^6\varphi^9)$$
$$+(-\frac{19}{3\sqrt{2}\pi}+\frac{16}{3\pi}-\frac{9}{4\sqrt{2}}+\frac{3}{2})(\theta^8\varphi^7+\theta^7\varphi^8)$$

In particular the asymptotic expansion of characteristic function $\phi(s,t)$ up to $\mathrm{O}(1/\rho N)$ for $p=2$ is given by:

$$
\begin{aligned}
&(-\frac{1}{2\sqrt{2}}+\frac{1}{2})(\theta^3+\varphi^3)+\frac{1}{2\sqrt{2}}(\theta^2\varphi+\theta\varphi^2)\\
&+\frac{2\rho_0-1}{\sqrt{\rho N}}\{\frac{1}{2\sqrt{2\pi}}(\theta^4-\varphi^4)+\frac{\sqrt{\pi}}{4\sqrt{2}}(\theta^3\varphi-\theta\varphi^3)\}\\
&+\frac{\rho_0(1-\rho_0)}{\rho N}\{(\frac{1}{12\sqrt{2}}-\frac{13}{24})(\theta^5+\varphi^5)-\frac{2}{3\sqrt{2}}(\theta^4\varphi+\theta\varphi^4)\\
&\quad+\frac{7}{12\sqrt{2}}(\theta^3\varphi^2+\theta^2\varphi^3)+(-\frac{13}{24\sqrt{2}}+\frac{13}{24})(\theta^3+\varphi^3)+\frac{13}{24\sqrt{2}}(\theta^2\varphi+\theta\varphi^2)\}\\
&+\frac{1}{\rho N}\{(-\frac{11}{24\sqrt{2}}+\frac{13}{24})(\theta^5+\varphi^5)+\frac{5}{12\sqrt{2}}(\theta^4\varphi+\theta\varphi^4)+\frac{1}{24\sqrt{2}}(\theta^3\varphi^2+\theta^2\varphi^3)\\
&\quad+(\frac{13}{24\sqrt{2}}-\frac{13}{24})(\theta^3+\varphi^3)-\frac{13}{24\sqrt{2}}(\theta^2\varphi+\theta\varphi^2)\}\\
&+\mathrm{O}\{\frac{1}{(\rho N)^{3/2}}\}
\end{aligned}
$$

Appendix B. Distributions of the maximum and minimum roots.

Theorem B.1 Let b_1 and b_p be the maximum and minimum roots of the $p\times p$ symmetric random matrix $\boldsymbol{A}$ with density (3.5). Then the distributions of b_1 and b_p are given by

$$\mathrm{P}(b_1\le x)=1-\mathrm{P}(b_p\le -x)=d(p)V_p(x;p-1,\dots,0),$$

where $d(p)$ is defined in (3.4), and $V_p(x;q_1,\dots,q_p)$ can be evaluated by the following recurrence formula:

$$
\begin{aligned}
&V_k(x;q_1,\dots,q_k)\\
=&-\mathrm{e}^{-\frac{1}{2}x^2}x^{q_1-1}V_{k-1}(x;q_2,\dots,q_k)\\
&+(q_1-1)V_k(x;q_1-2,q_2,\dots,q_k)\\
&+2\sum_{j=2}^{k}(-1)^j\frac{1}{2^{\frac{1}{2}(q_1+q_j)}}V_1(\sqrt{2}x;q_1+q_j-1)V_{k-2}(x;q_2,\dots,q_{j-1},q_{j+1},\dots,q_k)
\end{aligned}
$$

$(k\ge 2,\ q_1\ge 1)$ and

$$
\begin{aligned}
V_1(x;q)&=-\mathrm{e}^{-\frac{1}{2}x^2}x^{q-1}+(q-1)V_1(x;q-2) \qquad (q\ge 1),\\
V_1(x;0)&=\sqrt{2\pi}\Phi(x).
\end{aligned}
$$

($\Phi(\,\cdot\,)$ is the distribution function of $N(0,1)$)

(Proof) Note that

$$\mathrm{P}(b_1 \leq x) = d(p) \int_{x>b_1>\cdots>b_p} \cdots \int \exp\{-\frac{1}{2}\sum_{i=1}^{p} {b_i}^2\} \prod_{i<j}(b_i - b_j)db_1 \ldots db_p.$$

By the method parallel to the proof of Theorem 3.3, these relations can be derived.
□

The distribution function $\mathrm{P}(b_1 \leq x)$ for $1 \leq p \leq 5$ are provided as follows. ($\Phi(\,\cdot\,)$, $\phi(\,\cdot\,)$ are the distribution function and the density function of $N(0,1)$)

$$\begin{aligned}
p =& 1 : \\
& \Phi(x) \\
p =& 2 : \\
& \Phi(\sqrt{2}x) - \sqrt{\pi}\phi(x)\Phi(x) \\
p =& 3 : \\
& \Phi(x)\Phi(\sqrt{2}x) - 2x\phi(x)\Phi(\sqrt{2}x) - \frac{1}{\sqrt{\pi}}\phi(\sqrt{3}x) \\
p =& 4 : \\
& \Phi(\sqrt{2}x)^2 - \frac{\sqrt{\pi}}{2}(2x^2+1)\phi(x)\Phi(x)\Phi(\sqrt{2}x) \\
& - \sqrt{2}x\phi(\sqrt{2}x)\Phi(\sqrt{2}x) - \frac{1}{2}x\phi(\sqrt{3}x)\Phi(x) - \frac{1}{\sqrt{2\pi}}\phi(2x) \\
p =& 5 : \\
& \Phi(x)\Phi(\sqrt{2}x)^2 - \frac{4}{3}x^3\phi(x)\Phi(\sqrt{2}x)^2 \\
& - \frac{1}{3\sqrt{2}}(2x^3+9x)\phi(\sqrt{2}x)\Phi(x)\Phi(\sqrt{2}x) \\
& - \frac{1}{6\sqrt{\pi}}(10x^2+1)\phi(\sqrt{3}x)\Phi(\sqrt{2}x) \\
& - \frac{1}{3\sqrt{2\pi}}(x^2+4)\phi(2x)\Phi(x) - \frac{1}{2\pi}x\phi(\sqrt{5}x)
\end{aligned}$$

The formulae of $p \leq 3$ can be found in Muirhead (1982, p.424).

III. One-sided test for the equality of two covariance matrices concerning the complex multivariate normal population.

1. Introduction.

Let $\boldsymbol{H}$ and $\boldsymbol{G}$ be $p \times p$ random matrices which are independently distributed according to the complex Wishart distributions $CW_p(M, \boldsymbol{\Phi})$ and $CW_p(N, \boldsymbol{\Psi})$, respectively, where $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ are assumed to be positive definite and $M \geq p$, $N \geq p$. Consider the hierarchical hypotheses $H_0 \subset H_0^{(r)} \subset H_1 \subset H_2$ with

$$H_0 : \boldsymbol{\Phi} = \boldsymbol{\Psi}, \qquad H_0^{(r)} : \boldsymbol{\Phi} \geq \boldsymbol{\Psi},\ \mathrm{rank}\,(\boldsymbol{\Phi} - \boldsymbol{\Psi}) \leq r \quad (0 < r < p),$$
$$H_1 : \boldsymbol{\Phi} \geq \boldsymbol{\Psi}, \qquad H_2 : \boldsymbol{\Phi}, \boldsymbol{\Psi} \ \text{ are unrestricted.}$$

Here $\boldsymbol{A} \geq \boldsymbol{B}$ denotes Löwner order meaning that $\boldsymbol{A} - \boldsymbol{B}$ is nonnegative definite. In this chapter the likelihood ratio tests (LRT's) for the following hypotheses:

(i) T_{01} : LRT for testing H_0 against $H_1 - H_0$,

(ii) $T_{01}^{(r)}$: LRT for testing $H_0^{(r)}$ against $H_1 - H_0^{(r)}$, and

(iii) T_{12} : LRT for testing H_1 against $H_2 - H_1$

are discussed. The main purpose is to derive the limiting null distributions of test statistics of these LRT's.

These testing problems appear in the multiple time series model with replications:

$$\boldsymbol{x}_j(t) = \boldsymbol{v}(t) + \boldsymbol{u}_j(t), \quad t = 1, \ldots, T, \quad j = 1, \ldots, k, \tag{1.1}$$

where $\boldsymbol{x}_j(t)$ is a $p \times 1$ observed vector, $\boldsymbol{v}(t)$ is a $p \times 1$ unobserved stationary Gaussian signal with zero mean and continuous spectral density matrix $\boldsymbol{f}_v(\omega)$, and $\boldsymbol{u}_j(t)$ is a $p \times 1$ unobserved stationary Gaussian noise with zero mean and continuous spectral density matrix $\boldsymbol{f}_u(\omega)$ distributed independently of $\boldsymbol{v}(t)$ and $\boldsymbol{u}_l(t)$ $(l \neq j)$. Assume that $\boldsymbol{f}_u(\omega)$ is positive definite for each ω. Let $h = h(T)$ be an integer such that $\omega_h = 2\pi(h-1)/T \to \omega_0 \neq 0 \pmod{\pi}$ as $T \to \infty$. Define

$$\boldsymbol{H} = k \sum_{l=-n}^{n} \overline{\boldsymbol{X}}(\omega_{h+l}) \overline{\boldsymbol{X}}(\omega_{h+l})^* \tag{1.2}$$

and

$$\boldsymbol{G} = \sum_{l=-n}^{n} \sum_{j=1}^{k} \{\boldsymbol{X}_j(\omega_{h+l}) - \overline{\boldsymbol{X}}(\omega_{h+l})\}\{\boldsymbol{X}_j(\omega_{h+l}) - \overline{\boldsymbol{X}}(\omega_{h+l})\}^* \tag{1.3}$$

with

$$\boldsymbol{X}_j(\omega) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \mathrm{e}^{it\omega} \boldsymbol{x}_j(t), \quad \overline{\boldsymbol{X}}(\omega) = \frac{1}{k} \sum_{j=1}^{k} \boldsymbol{X}_j(\omega),$$

where '$*$' denotes conjugate transpose. It is easy to see that $\boldsymbol{H}$ (1.2) and $\boldsymbol{G}$ (1.3) are distributed independently. Moreover, using the usual chi-squared approximation of the smoothed periodogram (e.g. Brillinger, 1981, Section 5.4; Brockwell and Davis, 1991, Section 10.5), we see that $\boldsymbol{H}$ (1.2) and $\boldsymbol{G}$ (1.3) are approximately distributed according to the complex Wishart distributions $CW_m(M, \boldsymbol{\Phi})$ and $CW_m(N, \boldsymbol{\Psi})$, respectively, where $\boldsymbol{\Phi} = \boldsymbol{f}_v(\omega_0) + k\boldsymbol{f}_u(\omega_0)$, $M = 2n+1$, $\boldsymbol{\Psi} = \boldsymbol{f}_u(\omega_0)$ and $N = (2n+1)(k-1)$. Then, under this approximation, testing the null hypothesis of no signal at the frequency ω_0: $\boldsymbol{f}_u(\omega_0) = \boldsymbol{O}$ reduces to testing H_0 based on $\boldsymbol{H}$ (1.2) and $\boldsymbol{G}$ (1.3) against the one-sided alternative hypothesis $H_1 - H_0$. We are also interested in testing the null hypothesis that the signal vectors are linear combinations of r or less factors, which reduces to $H_0^{(r)}$, and the alternative should be $H_1 - H_0^{(r)}$. Testing the goodness of fit of the model (1.1) gives another type of restricted inference, testing H_1 against the alternative $H_2 - H_1$.

In the previous chapter we treated the similar testing problems for the multivariate variance components model, where $\boldsymbol{H}$ and $\boldsymbol{G}$ are distributed according to the real Wishart distributions. The likelihood ratio test statistics and some of their properties given in Section 2 closely parallel the results of the real case, however, the limiting null characteristic function of the likelihood ratio criterion given in Section 3 can be represented in terms of determinant (not Pfaffian). Significance points of these test statistics are tabulated in the same section. Appendix A illustrates examples of the limiting null distributions.

2. LRT statistics and their properties.

We give the test statistics and their least favorable distributions to calculate the significance points here. Let Λ_{01}, $\Lambda_{01}^{(r)}$ and Λ_{12} be the likelihood ratio criteria for T_{01}, $T_{01}^{(r)}$ and T_{12}, respectively. The similar discussion to the real case of Anderson, et al. (1986) yields that

$$\begin{aligned} \Lambda_{01}^{(r)} &= \prod_{i=r+1}^{R} \{\frac{l_i^{\rho}}{\rho l_i + 1 - \rho}\}^{M+N} && \text{if } R \geq r+1, \\ &= 1 && \text{if } R \leq r, \\ \Lambda_{01} &= \Lambda_{01}^{(0)}, \end{aligned}$$

and

$$\Lambda_{12} = \prod_{i=R+1}^{p} \{\frac{l_i^{\rho}}{\rho l_i + 1 - \rho}\}^{M+N} \qquad \text{if } R \le p-1,$$
$$= 1 \qquad \text{if } R = p,$$

with $\rho = M/(M+N)$, $l_1 > \cdots > l_p$ the latent roots of $(N/M)\boldsymbol{HG}^{-1}$, and R the number of $l_i > 1$. Since these statistics are functions of l_i's, their distributions depend only on the latent roots $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_p)'$, $\delta_1 \ge \cdots \ge \delta_p$, of $\boldsymbol{\Phi\Psi}^{-1}$.

We can summarize the properties of the likelihood ratio tests as follows. The proofs are the same as that of the real case of Chapter II.

Theorem 2.1
(i) The power function of T_{01}, $\beta_{01}(\boldsymbol{\delta}) = \mathrm{P}_{\delta}(-2\log\Lambda_{01} > c)$ with $c > 0$, is monotonically increasing in the components of $\boldsymbol{\delta}$. T_{01} is unbiased.
(ii) The power function of $T_{01}^{(r)}$ $(0 < r < p)$, $\beta_{01}^{(r)}(\boldsymbol{\delta}) = \mathrm{P}_{\delta}(-2\log\Lambda_{01}^{(r)} > c)$ with $c > 0$, is monotonically increasing in the components of $\boldsymbol{\delta}$. The least favorable distribution of $T_{01}^{(r)}$ is given at $\delta_1, \ldots, \delta_r \uparrow +\infty$ and $\delta_{r+1} = \cdots = \delta_p = 1$. $T_{01}^{(r)}$ is biased.
(iii) The power function of T_{12}, $\beta_{12}(\boldsymbol{\delta}) = \mathrm{P}_{\delta}(-2\log\Lambda_{12} > c)$ with $c > 0$, is monotonically decreasing in the components of $\boldsymbol{\delta}$. The least favorable distribution of T_{12} is given at $\delta_1 = \cdots = \delta_p = 1$, i.e. when H_0 holds. T_{12} is biased.

3. Limiting null distribution of LRT.

Firstly we treat T_{01} and T_{12}. To obtain the significance points of T_{01} and T_{12}, we have to derive the distribution of Λ_{01} and Λ_{12} under H_0 because the least favorable distribution of T_{12} is given under H_0. This section gives the limiting joint distribution of Λ_{01} and Λ_{12} under H_0.

Let $\phi(s,t)$ denote the joint characteristic function of $-2\log\Lambda_{01}$ and $-2\log\Lambda_{12}$ under H_0, i.e. $\phi(s,t) = \mathrm{E}_{H_0}{\Lambda_{01}}^{-2is}{\Lambda_{12}}^{-2it}$. The joint density function of $l_1 > \cdots > l_p > 0$ is given by

$$\tilde{c}(p; M, N) \prod_{i=1}^{p} l_i^{M-p-1} (\frac{M}{M+N} l_i + \frac{N}{M+N})^{-(M+N)} \prod_{i<j} (l_i - l_j)^2,$$

where

$$\tilde{c}(p; M, N) = \frac{\pi^{p(p-1)} \tilde{\Gamma}_p(M+N)}{\tilde{\Gamma}_p(M)\tilde{\Gamma}_p(N)\tilde{\Gamma}_p(p)} \cdot \frac{M^{pM} N^{pN}}{(M+N)^{p(M+N)}}$$

with $\tilde{\Gamma}_p(a) = \pi^{p(p-1)/2} \prod_{i=1}^{p} \Gamma(a-i+1)$. Define

$$b_i = \frac{1}{2}\sqrt{\frac{MN}{M+N}}(l_i - 1), \quad 1 \le i \le p.$$

By letting $M, N \to \infty$ with $M/(M+N) \to \rho_0$ $(0 \le \rho_0 \le 1)$ and b_i fixed (i.e. $l_i \to 1$), the limiting joint density of $\boldsymbol{b} = (b_1, \ldots, b_p)$, $b_1 > \cdots > b_p$, is

$$\varphi_0(\boldsymbol{b}) = \tilde{d}(p) \exp\{-\frac{1}{2}\sum_{i=1}^{p} {b_i}^2\} \prod_{i<j}(b_i - b_j)^2 \tag{3.1}$$

with

$$\tilde{d}(p) = \frac{\pi^{p(p-2)/2}}{2^{p/2}\tilde{\Gamma}_p(p)} = \frac{1}{2^{p/2}\pi^{p/2}\prod_{i=1}^{p}\Gamma(i)}.$$

$\varphi_0(\boldsymbol{b})$ in (3.1) turns out to be the density of the latent roots of a $p \times p$ Hermitian random matrix $\boldsymbol{A}$ with complex normal density

$$\frac{1}{2^{p/2}\pi^{p^2/2}} \exp\{-\frac{1}{2}\mathrm{tr}\boldsymbol{A}\boldsymbol{A}^*\}.$$

On the other hand it is easy to show that

$$\lim_{M,N\to\infty} -2\log\Lambda_{01} = \sum_{i=1}^{p}(b_i \vee 0)^2, \quad \lim_{M,N\to\infty} -2\log\Lambda_{12} = \sum_{i=1}^{p}(b_i \wedge 0)^2,$$

where $x \vee y$ and $x \wedge y$ are the maximum and the minimum of x and y, respectively. Therefore by Lemma 3.1 of Chapter II we get the characteristic function of the limiting null distribution as

$$\begin{aligned}\phi_0(s,t) &= \lim_{M,N\to\infty} \phi(s,t)\\ &= \int_{b_1>\cdots>b_p} \exp\{is\sum_{i=1}^{p}(b_i \vee 0)^2 + it\sum_{i=1}^{p}(b_i \wedge 0)^2\}\varphi_0(\boldsymbol{b})d\boldsymbol{b}\\ &= \sum_{r=0}^{p}\int_{B_r\times\overline{B}_{p-r}} \exp\{is\sum_{i=1}^{r}{b_i}^2 + it\sum_{i=r+1}^{p}{b_i}^2\}\varphi_0(\boldsymbol{b})d\boldsymbol{b}\end{aligned} \tag{3.2}$$

with $d\boldsymbol{b} = \prod_{i=1}^{p} db_i$,

$$B_r = \{(b_1, \ldots, b_r) \mid b_1 > \cdots > b_r > 0\},$$
$$\overline{B}_{p-r} = \{(b_{r+1}, \ldots, b_p) \mid 0 > b_{r+1} > \cdots > b_p\}.$$

By the Laplace expansion of the linkage factor $\prod_{i<j}(b_i - b_j)^2$, which is the Vandermonde determinant squared $\{\det({b_i}^{p-j})_{1\le i,j\le p}\}^2$, (3.2) is

$$\begin{aligned}&\tilde{d}(p)\sum_{r=0}^{p}\sum_{\lambda,\mu}(-1)^{\sum_{i=1}^{r}(\lambda_i+\mu_i)}\\ &\times\int_{B_r}\exp\{-\frac{1}{2\theta^2}\sum_{i=1}^{r}{b_i}^2\}\det({b_i}^{p-\lambda_j})_{1\le i,j\le r}\det({b_i}^{p-\mu_j})_{1\le i,j\le r}db_1\ldots db_r\\ &\times\int_{\overline{B}_{p-r}}\exp\{-\frac{1}{2\varphi^2}\sum_{i=r+1}^{p}{b_i}^2\}\det({b_{i+r}}^{p-\overline{\lambda}_j})_{1\le i,j\le p-r}\det({b_{i+r}}^{p-\overline{\mu}_j})_{1\le i,j\le p-r}db_{r+1}\ldots db_p\end{aligned} \tag{3.3}$$

where $\theta = (1-2is)^{-\frac{1}{2}}$ and $\varphi = (1-2it)^{-\frac{1}{2}}$, $\sum_{\lambda,\mu}$ is summation over all combinations of $\lambda_1 < \cdots < \lambda_r$, $\overline{\lambda}_1 < \cdots < \overline{\lambda}_{p-r}$, $\mu_1 < \cdots < \mu_r$, $\overline{\mu}_1 < \cdots < \overline{\mu}_{p-r}$ such that

$$\{\lambda_1, \ldots, \lambda_r, \overline{\lambda}_1, \ldots, \overline{\lambda}_{p-r}\} = \{\mu_1, \ldots, \mu_r, \overline{\mu}_1, \ldots, \overline{\mu}_{p-r}\} = \{1, \ldots, p\}.$$

Using the determinental Cauchy-Binet formula ((2.1) of Krishnaiah (1976), (2.12) of Karlin and Rinott (1988)), (3.3) reduces to

$$\tilde{d}(p) \sum_{r=0}^{p} \sum_{\lambda,\mu} (-1)^{\sum_{i=1}^{r}(\lambda_i+\mu_i)} \det(\boldsymbol{G}[\boldsymbol{\lambda}, \boldsymbol{\mu}]) \det(\boldsymbol{G}[\overline{\boldsymbol{\lambda}}, \overline{\boldsymbol{\mu}}]) \theta^{Q} \varphi^{\overline{Q}} \tag{3.4}$$

with $Q = \sum_{i=1}^{r}(\lambda_i + \mu_i) - r$, $\overline{Q} = p^2 - Q$, $\boldsymbol{G}[\boldsymbol{\lambda}, \boldsymbol{\mu}] = (g_{\lambda_i \mu_j})_{1 \le i,j \le r}$ an $r \times r$ matrix, $\boldsymbol{G}[\overline{\boldsymbol{\lambda}}, \overline{\boldsymbol{\mu}}] = (g_{\overline{\lambda}_i \overline{\mu}_j})_{1 \le i,j \le p-r}$ a $(p-r) \times (p-r)$ matrix with

$$g_{ij} = \int_0^\infty \mathrm{e}^{-\frac{1}{2}b^2} b^{2p-i-j} db = 2^{(2p-i-j-1)/2} \Gamma\left(\frac{2p-i-j+1}{2}\right).$$

Here (3.4) is a characteristic function of a mixture of the bivariate chi-squared distributions with Q and $\overline{Q}$ $(0 \le Q, \overline{Q} \le p^2)$ degrees of freedom. Noting the identity that, for two $p \times p$ matrices $\boldsymbol{A}$ and $\boldsymbol{B}$,

$$\det(\boldsymbol{A} + \boldsymbol{B}) = \sum_{r=0}^{p} \sum_{\lambda,\mu} (-1)^{\sum_{i=1}^{r}(\lambda_i+\mu_i)} \det(\boldsymbol{A}[\boldsymbol{\lambda}, \boldsymbol{\mu}]) \det(\boldsymbol{B}[\overline{\boldsymbol{\lambda}}, \overline{\boldsymbol{\mu}}]),$$

which corresponds to Lemma 3.2 of Chapter II, (3.4) is represented as

$$\begin{aligned} &\tilde{d}(p) \det(\theta \boldsymbol{D}(\theta) \boldsymbol{G} \boldsymbol{D}(\theta) + \varphi \boldsymbol{D}(-\varphi) \boldsymbol{G} \boldsymbol{D}(-\varphi)) \\ &\quad = \tilde{d}(p) \det(\{\theta^{i+j-1} - (-\varphi)^{i+j-1}\} g_{ij})_{1 \le i,j \le p} \end{aligned} \tag{3.5}$$

with $\boldsymbol{G} = (g_{ij})_{1 \le i,j \le p}$ and $\boldsymbol{D}(\xi) = \mathrm{diag}(\xi^{i-1})_{1 \le i \le p}$. By inverting the characteristic function (3.5), we get the following formula.

Theorem 3.1 As $M, N \to \infty$ with $M/(M+N) \to \rho_0$ $(0 \le \rho_0 \le 1)$, the limiting joint distribution function of $-2\log\Lambda_{01}$ and $-2\log\Lambda_{12}$ under H_0 is given by

$$\lim_{M,N \to \infty} \mathrm{P}_{H_0}(-2\log\Lambda_{01} \le y, -2\log\Lambda_{12} \le z) = \sum_{Q=0}^{p^2} w(Q; p^2) G_Q(y) G_{\overline{Q}}(z),$$

where $w(Q; p^2)$ is the coefficient of the term of $\theta^Q \varphi^{\overline{Q}}$ in (3.4) or (3.5), $G_\nu(\,\cdot\,)$ with $\nu \ne 0$ is the distribution function of the chi-squared distribution with ν degrees of freedom and $G_0(\,\cdot\,) = I(0 \le \,\cdot\,)$.

Appendix A gives the characteristic functions of the limiting null distribution (3.5) for $1 \leq p \leq 5$ obtained with REDUCE3, a software for algebraic computation. Table 3.1 shows the limiting α-significance points of T_{01} and T_{12}, i.e., $d(\alpha; p)$ satisfying

$$\lim_{M,N\to\infty} \mathrm{P}_{H_0}(-2\log\Lambda_{01} > d(\alpha;p)) = \mathrm{P}(\sum_{i=1}^{p}(b_i \vee 0)^2 > d(\alpha;p)) = \alpha,$$

where $b_1 > \cdots > b_p$ are distributed with density (3.1). Put $d(\alpha; p) = 0$ when

$$\mathrm{P}(\sum_{i=1}^{p}(b_i \vee 0)^2 > 0) < \alpha.$$

Next we can obtain the limiting null (least favorable) distribution of $\Lambda_{01}^{(r)}$. The proof is given in the same manner as Theorem 3.2 of Chapter II.

Theorem 3.2 As $M, N \to \infty$ with $M/(M+N) \to \rho_0$ $(0 \leq \rho_0 \leq 1)$,

$$\lim_{M,N\to\infty} \sup_{H_0^{(r)}} \mathrm{P}(-2\log\Lambda_{01}^{(r)} \leq y) = \sum_{Q=0}^{{p_1}^2} w(Q; {p_1}^2) G_Q(y)$$

with $p_1 = p - r$.

Table 3.1

Limiting significance points of T_{01} and T_{12}

$1-\alpha \setminus p$	2	3	4	5	6	7	8	9	10
0.010	0.0000	0.0220	0.8116	2.6165	5.4321	9.2526	14.0770	19.9041	26.7328
0.025	0.0000	0.1707	1.3574	3.5714	6.7989	11.0354	16.2759	22.5207	29.7670
0.050	0.0000	0.4166	1.9565	4.5305	8.1168	12.7141	18.3146	24.9203	32.5271
0.100	0.0000	0.8455	2.8049	5.7998	9.8042	14.8198	20.8376	27.8606	35.8841
0.250	0.2996	1.9389	4.6204	8.3267	13.0369	18.7548	25.4734	33.1955	41.9176
0.500	1.2526	3.7558	7.2540	11.7547	17.2543	23.7545	31.2543	39.7544	49.2543
0.750	2.9080	6.2558	10.5696	15.8628	22.1515	29.4335	37.7145	46.9923	57.2700
0.900	5.0344	9.1224	14.1599	20.1616	27.1565	35.1399	44.1219	54.0985	65.0748
0.950	6.6066	11.1271	16.5909	23.0134	30.4280	38.8297	48.2295	58.6233	70.0168
0.975	8.1564	13.0450	18.8739	25.6589	33.4355	42.1986	51.9597	62.7148	74.4692
0.990	10.1791	15.4872	21.7347	28.9381	37.1329	46.3148	56.4943	67.6683	79.8415

Appendix A. Examples of the limiting distributions.

The limiting characteristic functions $\phi_0(s,t)$ for $1 \le p \le 5$ are presented as follows. ($\theta = (1-2is)^{-\frac{1}{2}}$, $\varphi = (1-2it)^{-\frac{1}{2}}$)

$p = 1:$

$$\frac{1}{2}(\theta + \varphi)$$

$p = 2:$

$$(\frac{1}{4} - \frac{1}{2\pi})(\theta^4 + \varphi^4) + \frac{1}{4}(\theta^3\varphi + \theta\varphi^3) + \frac{1}{\pi}\theta^2\varphi^2$$

$p = 3:$

$$(\frac{1}{8} - \frac{3}{8\pi})(\theta^9 + \varphi^9) + (\frac{3}{16} - \frac{1}{2\pi})(\theta^8\varphi + \theta\varphi^8) + \frac{1}{4\pi}(\theta^7\varphi^2 + \theta^2\varphi^7)$$
$$+ \frac{1}{2\pi}(\theta^6\varphi^3 + \theta^3\varphi^6) + (\frac{3}{16} + \frac{1}{8\pi})(\theta^5\varphi^4 + \theta^4\varphi^5)$$

$p = 4:$

$$(\frac{1}{16} - \frac{29}{96\pi} + \frac{1}{3\pi^2})(\theta^{16} + \varphi^{16}) + (\frac{3}{32} - \frac{7}{24\pi})(\theta^{15}\varphi + \theta\varphi^{15})$$
$$+ (\frac{7}{16\pi} - \frac{4}{3\pi^2})(\theta^{14}\varphi^2 + \theta^2\varphi^{14}) + (\frac{3}{32} - \frac{1}{4\pi})(\theta^{13}\varphi^3 + \theta^3\varphi^{13})$$
$$+ (\frac{15}{64} - \frac{47}{32\pi} + \frac{8}{3\pi^2})(\theta^{12}\varphi^4 + \theta^4\varphi^{12}) + (-\frac{3}{32} + \frac{17}{32\pi})(\theta^{11}\varphi^5 + \theta^5\varphi^{11})$$
$$+ (-\frac{3}{8} + \frac{17}{6\pi} - \frac{4}{\pi^2})(\theta^{10}\varphi^6 + \theta^6\varphi^{10}) + (\frac{5}{32} + \frac{1}{96\pi})(\theta^9\varphi^7 + \theta^7\varphi^9)$$
$$+ (\frac{21}{32} - \frac{3}{\pi} + \frac{14}{3\pi^2})\theta^8\varphi^8$$

$p = 5:$

$$(\frac{1}{32} - \frac{145}{768\pi} + \frac{41}{144\pi^2})(\theta^{25} + \varphi^{25}) + (\frac{15}{256} - \frac{125}{384\pi} + \frac{4}{9\pi^2})(\theta^{24}\varphi + \theta\varphi^{24})$$
$$+ (\frac{17}{128\pi} - \frac{5}{12\pi^2})(\theta^{23}\varphi^2 + \theta^2\varphi^{23}) + (\frac{109}{384\pi} - \frac{8}{9\pi^2})(\theta^{22}\varphi^3 + \theta^3\varphi^{22})$$
$$+ (\frac{75}{512} - \frac{135}{256\pi} + \frac{2}{9\pi^2})(\theta^{21}\varphi^4 + \theta^4\varphi^{21}) + (\frac{15}{128} - \frac{237}{512\pi} + \frac{1}{3\pi^2})(\theta^{20}\varphi^5 + \theta^5\varphi^{20})$$
$$+ (-\frac{15}{64} + \frac{271}{192\pi} - \frac{73}{36\pi^2})(\theta^{19}\varphi^6 + \theta^6\varphi^{19})$$
$$+ (-\frac{5}{32} + \frac{563}{768\pi} - \frac{5}{9\pi^2})(\theta^{18}\varphi^7 + \theta^7\varphi^{18}) + (\frac{345}{512} - \frac{271}{64\pi} + \frac{169}{24\pi^2})(\theta^{17}\varphi^8 + \theta^8\varphi^{17})$$
$$+ (\frac{245}{512} - \frac{1949}{768\pi} + \frac{187}{48\pi^2})(\theta^{16}\varphi^9 + \theta^9\varphi^{16})$$
$$+ (-\frac{75}{64} + \frac{259}{32\pi} - \frac{155}{12\pi^2})(\theta^{15}\varphi^{10} + \theta^{10}\varphi^{15})$$
$$+ (-\frac{15}{16} + \frac{1675}{256\pi} - \frac{121}{12\pi^2})(\theta^{14}\varphi^{11} + \theta^{11}\varphi^{14})$$
$$+ (\frac{765}{512} - \frac{13703}{1536\pi} + \frac{44}{3\pi^2})(\theta^{13}\varphi^{12} + \theta^{12}\varphi^{13})$$

IV. Tests for covariance structure in random coefficient regression model.

1. Introduction.

Let $\boldsymbol{W}$ be a $p \times p$ random matrix distributed according to the Wishart distribution $W_p(n, \boldsymbol{\Phi})$ with $\boldsymbol{\Phi}$ positive definite and $n \geq p$. Let $(\nu n/\sigma^2)g$ be a random variable distributed according to the chi-squared distribution $\chi^2(\nu n)$, $\nu > 0$. Define hierarchical hypotheses $H_0 \subset H_1 \subset H_2$ as

$$H_0 : \boldsymbol{\Phi} = \sigma^2 \boldsymbol{I}_p, \quad H_1 : \boldsymbol{\Phi} \geq \sigma^2 \boldsymbol{I}_p, \quad \text{and} \quad H_2 : \boldsymbol{\Phi}, \sigma^2 \ \text{ unrestricted},$$

where $\boldsymbol{A} \geq \boldsymbol{B}$ means that $\boldsymbol{A} - \boldsymbol{B}$ is nonnegative definite. We treat two likelihood ratio tests (LRT's) based on the observed values $\boldsymbol{W}$ and g denoted by

$$T_{01} : \text{LRT for testing } H_0 \text{ against } H_1 - H_0,$$

and

$$T_{12} : \text{LRT for testing } H_1 \text{ against } H_2 - H_1.$$

Here the notations are due to Robertson, et al. (1988, Chapter 2) who discussed the testing problems for the hierarchical hypotheses of the multivariate normal means. In this chapter two basic properties, the unbiasedness of the LRT T_{01}, and the monotonicity of power function of the LRT T_{12} are proved. As n goes to infinity, limiting null distributions of these two LRT statistics are derived as mixtures of chi-squared distributions. Moreover for a general class of tests for testing H_0 against $H_1 - H_0$ including the LRT T_{01} the local unbiasedness is proved.

The hypothesis H_1 is equivalent to

$$\boldsymbol{\Phi} = \boldsymbol{\Theta} + \sigma^2 \boldsymbol{I}_p \tag{1.1}$$

for some $p \times p$ nonnegative matrix $\boldsymbol{\Theta}$. This covariance structure (1.1) was introduced by Rao (1965) who discuss multivariate regression models with random coefficients (random effects model). When a Wishart matrix with a parameter $\boldsymbol{\Phi}$ in (1.1) and an unbiased estimate of σ^2 are observed independently, testing the hypothesis that $\boldsymbol{\Theta} = \boldsymbol{O}$ reduces to T_{01}, and T_{12} is a test of goodness of fit of the covariance structure (1.1). There are a lot of studies on testing problems for random effects models, however, few paper treated the likelihood ratio test for the covariance structure (1.1) exactly. One exception is Anderson, et al. (1986, Section 3) who derived the likelihood ratio criterion of T_{01} in terms of Scheffé's mixed model. Our discussions are based on the result of Anderson, et al. (1986).

Outline of this chapter is as follows. In Section 2 the unbiasedness of T_{01} and the monotonicity of power function of T_{12} are proved. Section 3 gives the limiting null distributions of the test statistics of T_{01} and T_{12} as mixtures of chi-squared distributions. In Section 4 a general class of the tests for H_0 against $H_1 - H_0$ including the LRT T_{01} are proved to be locally unbiased using FKG inequality technique. Applications of T_{01} and T_{12} to random effects models are stated in Section 5. Appendix A illustrates examples of the limiting null distributions for $p = 2, 3$. In Appendix B we summarize the FKG inequality which is used in the proof in Section 4. Here a new sufficient condition for the FKG condition is proved.

2. Properties of power functions of LRT.

2.1 LRT statistics.

In this section we shall prove the unbiasedness of the LRT T_{01} and the monotonicity of power function of the LRT T_{12}.

To begin with, we display the LRT statistics Λ_{01} and Λ_{12} of T_{01} and T_{12}, respectively. The LRT statistic Λ_{01} was basically given by Anderson, et al. (1986) as

$$\begin{aligned}\Lambda_{01} &= \prod_{i=1}^{m^*} t_i{}^{n/2} \frac{s_{m^*}{}^{n(\nu+p-m^*)/2}}{s_0{}^{n(\nu+p)/2}} && \text{if } m^* \geq 1,\\ &= 1 && \text{if } m^* = 0, \qquad (2.1)\end{aligned}$$

where $t_1 > \cdots > t_p$ are the latent roots of $(1/n)\boldsymbol{W}$,

$$s_i = \frac{\nu g + \sum_{j=i+1}^{p} t_j}{\nu + p - i}, \quad 0 \leq i \leq p,$$

and m^* a random integer such that

$$t_{m^*} \geq s_{m^*} \quad \text{and} \quad t_{m^*+1} < s_{m^*+1}. \qquad (2.2)$$

The random integer m^* can be determined uniquely, because

$$t_{i+1} \geq s_{i+1} \quad \Rightarrow \quad t_i \geq s_i$$

or equivalently

$$t_{i+1} < s_{i+1} \quad \Leftarrow \quad t_i < s_i \qquad (2.3)$$

hold from the relation

$$t_i - s_i = (t_i - t_{i+1}) + \frac{\nu + p - i - 1}{\nu + p - i}(t_{i+1} - s_{i+1}).$$

Note that our definition of m^* in (2.2) is equivalent to (3.13) of Anderson, et al. (1986). Since the likelihood ratio criterion Λ_{02} for testing H_0 against $H_2 - H_0$ is

$$\Lambda_{02} = \prod_{i=1}^{p} {t_i}^{n/2} \frac{{s_p}^{n\nu/2}}{{s_0}^{n(\nu+p)/2}},$$

we have

$$\begin{aligned} \Lambda_{12} = \frac{\Lambda_{02}}{\Lambda_{01}} &= \prod_{i=m^*+1}^{p} {t_i}^{n/2} \frac{{s_p}^{n\nu/2}}{{s_{m^*}}^{n(\nu+p-m^*)/2}} && \text{if } m^* \leq p-1, \\ &= 1 && \text{if } m^* = p. \end{aligned} \tag{2.4}$$

Figure 2.1 shows the acceptance regions in $\mathbf{R}^2$ of T_{01} and T_{12} defined by $\{(t_1, t_2) \,|\, \Lambda_{01} \leq c\}$ and $\{(t_1, t_2) \,|\, \Lambda_{12} \leq c\}$, respectively, for fixed g. Note that Λ_{01} is not monotone in t_1.

2.2 Monotonicity of power function of the LRT T_{12}.

We prove that the power function of the LRT T_{12} is strictly monotone in the latent roots of $(1/\sigma^2)\boldsymbol{\Phi}$, and show that the least favorable distribution of Λ_{12} is given when H_0 holds. The following lemma is used in proving Theorems 2.1 and 2.2.

Lemma 2.1 The LRT statistic Λ_{12} is an increasing function of $\bar{t}_i = t_i/g$, $1 \leq i \leq p$. In particular Λ_{12} is strictly increasing in each argument of $\bar{t}_i$, $1 \leq i \leq p$, on the set $\{m^* = 0\}$.

(Proof) Put

$$\bar{s}_i = \frac{s_i}{g} = \frac{\nu + \sum_{j=i+1}^{p} \bar{t}_j}{\nu + p - i}. \tag{2.5}$$

On the set $\{m^* = m\}$ we have

$$-2 \log \Lambda_{12} = n\{- \sum_{i=m+1}^{p} \log \bar{t}_i - \nu \log \bar{s}_p + (\nu + p - m) \log \bar{s}_m\},$$

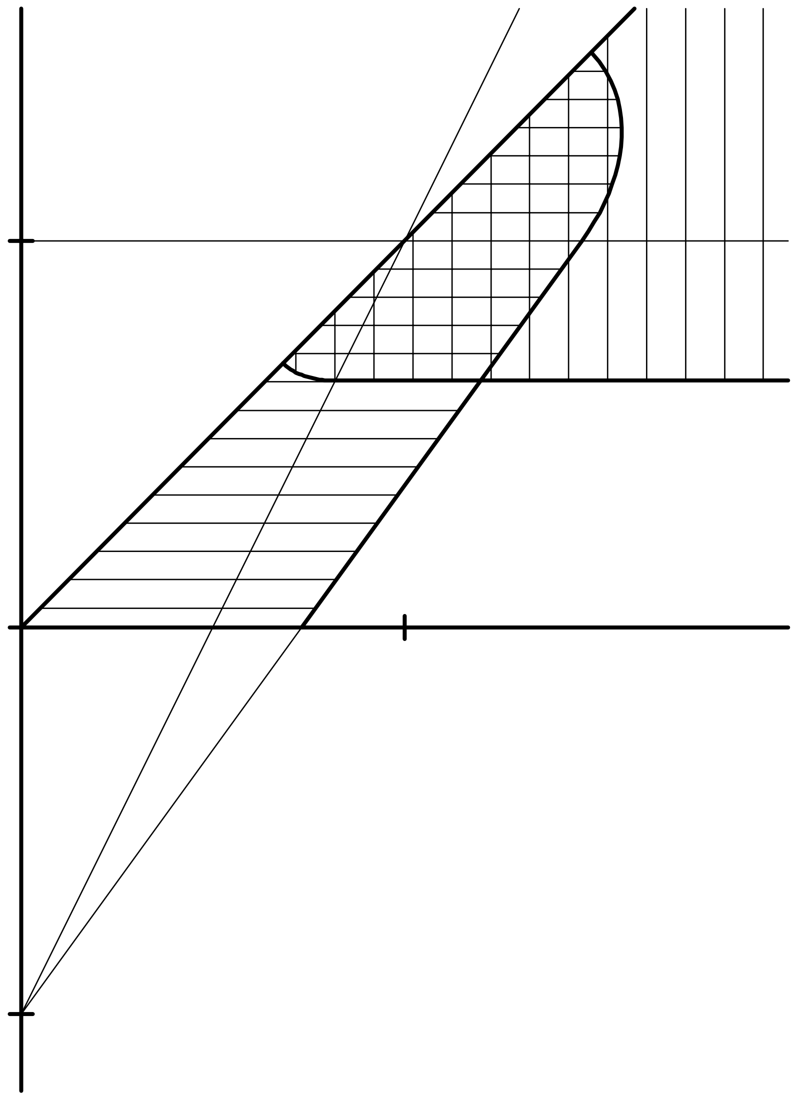
which depends only on $\bar{t}_{m+1}, \ldots, \bar{t}_p$. And for $j \geq m+1$ it holds

$$\begin{aligned} \frac{1}{n} \frac{\partial}{\partial \bar{t}_j}(-2 \log \Lambda_{12}) &= -\frac{1}{\bar{t}_j} + \frac{1}{\bar{s}_m} \\ &= \frac{(\nu + p - j)(\bar{t}_j - \bar{s}_j) - \sum_{i=m+1}^{j} (\bar{t}_i - \bar{t}_j)}{\bar{t}_j(\nu + \sum_{i=m+1}^{p} \bar{t}_i)}, \end{aligned} \tag{2.6}$$

which is not positive since $\bar{t}_j - \bar{s}_j \leq 0$ and $\bar{t}_i - \bar{t}_j \geq 0$ for $m+1 \leq i \leq j$. In particular (2.6) is negative on the set $\{m^* = 0\} = \{t_1 < s_1\}$ since $\bar{t}_j - \bar{s}_j < 0$, $1 \leq j \leq p$,

Figure 2.1

Acceptance regions of T_{01} and T_{12}



holds from (2.3). Because Λ_{12} is a continuous function of $\bar{t}_1, \ldots, \bar{t}_p$, the proof is completed. □

Theorem 2.1 The power function β_{12} of the LRT T_{12}, $\mathrm{P}(\Lambda_{12} < c)$, $0 < c \leq 1$, is a function of the latent roots $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_p)$, $\delta_1 \geq \cdots \geq \delta_p$, of $(1/\sigma^2)\boldsymbol{\Phi}$. $\beta_{12}(\boldsymbol{\delta})$ is strictly decreasing in each component of $\boldsymbol{\delta}$. The least favorable distribution of T_{12} is given at $\boldsymbol{\delta} = (1, \ldots, 1)$ $(= \boldsymbol{e}$, say), i.e. H_0 is true. T_{12} is biased.

(Proof) We follow the same argument of Theorem 2 of Anderson and Das Gupta (1964).

We can regard Λ_{12} in (2.4) as a function of $\overline{\boldsymbol{W}} = (1/g)\boldsymbol{W}$ and put $f^{-}(\overline{\boldsymbol{W}}) = \Lambda_{12}$. It is easy to see that the distribution of the latent roots of $\overline{\boldsymbol{W}}$ depends only on the latent roots of $(1/\sigma^2)\boldsymbol{\Phi}$. The probability measure is denoted by $\mathrm{P}(\,\cdot\,; \boldsymbol{\delta})$. For two $p \times 1$ positive vectors $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ such that $\boldsymbol{d}_2 \geq \boldsymbol{d}_1$, i.e. all components of $\boldsymbol{d}_2 - \boldsymbol{d}_1$ are nonnegative, we have that

$$f^{-}(\overline{\boldsymbol{W}}) \leq f^{-}(\boldsymbol{D}^{1/2}\overline{\boldsymbol{W}}\boldsymbol{D}^{1/2}) \tag{2.7}$$

with $\boldsymbol{D} = \mathrm{diag}(\boldsymbol{d}_1)^{-1}\mathrm{diag}(\boldsymbol{d}_2) \geq \boldsymbol{I}_p$ because of Lemma 2.1 and the fact that

$$\lambda_i(\overline{\boldsymbol{W}}) \leq \lambda_i(\boldsymbol{D}^{1/2}\overline{\boldsymbol{W}}\boldsymbol{D}^{1/2}), \quad 1 \leq i \leq p, \tag{2.8}$$

where $\lambda_i(\,\cdot\,)$ means the ith largest latent root. Therefore

$$\begin{aligned} &\mathrm{P}(f^{-}(\overline{\boldsymbol{W}}) < c\,; \boldsymbol{d}_1) - \mathrm{P}(f^{-}(\overline{\boldsymbol{W}}) < c\,; \boldsymbol{d}_2) \\ &\quad = \mathrm{P}(f^{-}(\overline{\boldsymbol{W}}) < c\,; \boldsymbol{d}_1) - \mathrm{P}(f^{-}(\boldsymbol{D}^{1/2}\overline{\boldsymbol{W}}\boldsymbol{D}^{1/2}) < c\,; \boldsymbol{d}_1) \\ &\quad = \mathrm{P}(f^{-}(\overline{\boldsymbol{W}}) < c \leq f^{-}(\boldsymbol{D}^{1/2}\overline{\boldsymbol{W}}\boldsymbol{D}^{1/2})\,; \boldsymbol{d}_1) \geq 0. \end{aligned} \tag{2.9}$$

If $\boldsymbol{d}_1 \neq \boldsymbol{d}_2$ (i.e. $\boldsymbol{D} \neq \boldsymbol{I}_p$), the inequality of (2.9) holds strictly, because at least one inequality of (2.8) holds strictly for positive definite $\overline{\boldsymbol{W}}$ and hence

$$\text{L.H.S. of (2.9)} \geq \mathrm{P}(f^{-}(\overline{\boldsymbol{W}}) < c \leq f^{-}(\boldsymbol{D}^{1/2}\overline{\boldsymbol{W}}\boldsymbol{D}^{1/2}), m^* = 0\,; \boldsymbol{d}_1) > 0$$

for $0 < c \leq 1$. This means the power function $\beta_{12}(\boldsymbol{\delta}) = \mathrm{P}(f^{-}(\overline{\boldsymbol{W}}) < c\,; \boldsymbol{\delta})$ is strictly decreasing in each component of $\boldsymbol{\delta}$.

$\sup_{H_1} \beta_{12}(\boldsymbol{\delta})$ is attained when $\boldsymbol{\delta} = \boldsymbol{e}$, which means that the least favorable distribution of T_{12} is given when H_0 is true.

T_{12} is biased because $\inf_{H_2 - H_1} \beta_{12}(\boldsymbol{\delta})$, which attains when $\delta_1, \ldots, \delta_{p-1} \uparrow +\infty$ and $\delta_p \uparrow 1$, is strictly less than $\alpha = \sup_{H_1} \beta_{12}(\boldsymbol{\delta}) = \beta_{12}(\boldsymbol{e})$. The proof is completed. □

2.3 Unbiasedness of the LRT $\boldsymbol{T}_{01}$.

From Lemma 2.1 we prove the unbiasedness of the LRT T_{01} with the aid of the arguments by Sugiura and Nagao (1968) or Perlman (1980).

Theorem 2.2 The power function β_{01} of the LRT T_{01}, $\mathrm{P}(\Lambda_{01} < c)$, $0 < c \leq 1$, is a function of the latent roots $\boldsymbol{\delta}$ of $(1/\sigma^2)\boldsymbol{\Phi}$. T_{01} is strictly unbiased, i.e. $\beta_{01}(\boldsymbol{\delta}) \geq \beta_{01}(\boldsymbol{e})$ for $\boldsymbol{\delta} \geq \boldsymbol{e}$, and inequality holds strictly if $\boldsymbol{\delta} \neq \boldsymbol{e}$.

(Proof) Similar to Λ_{12} we put $f^+(\overline{\boldsymbol{W}}) = \Lambda_{01}$ in (2.1) as a function of $\overline{\boldsymbol{W}}$. Let $f(\overline{\boldsymbol{W}}\,;\boldsymbol{\delta})$ be the density function of $\overline{\boldsymbol{W}}$. By making the change of variables from $(\boldsymbol{W}, g)$ to $(\overline{\boldsymbol{W}}, g)$, and integrating the joint density of $(\overline{\boldsymbol{W}}, g)$ with respect to g, we can write

$$\begin{aligned} f(\overline{\boldsymbol{W}}\,;\boldsymbol{\delta}) =& K'|\boldsymbol{D}|^{-n/2}|\overline{\boldsymbol{W}}|^{(n-p-1)/2}(1+\frac{1}{n\nu}\mathrm{tr}\boldsymbol{D}^{-1}\overline{\boldsymbol{W}})^{-n(\nu+p)/2} \\ =& K''f^+(\boldsymbol{D}^{-1/2}\overline{\boldsymbol{W}}\boldsymbol{D}^{-1/2})f^-(\boldsymbol{D}^{-1/2}\overline{\boldsymbol{W}}\boldsymbol{D}^{-1/2})|\overline{\boldsymbol{W}}|^{-(p+1)/2}, \end{aligned} \tag{2.10}$$

where $\boldsymbol{D} = \mathrm{diag}(\boldsymbol{\delta})$, K' and K'' are normalizing constants. The power function is written as

$$\beta_{01}(\boldsymbol{\delta}) = 1 - \mathrm{P}(f^+(\overline{\boldsymbol{W}}) \geq c\,;\boldsymbol{\delta}) = 1 - \mathrm{P}(f^+(\boldsymbol{D}^{1/2}\overline{\boldsymbol{W}}\boldsymbol{D}^{1/2}) \geq c\,;\boldsymbol{e}).$$

Put

$$\begin{aligned} \omega_0 =& \{\overline{\boldsymbol{W}} \mid f^+(\overline{\boldsymbol{W}}) \geq c, \overline{\boldsymbol{W}} \geq \boldsymbol{O}\}, \\ \omega_1 =& \{\overline{\boldsymbol{W}} \mid f^+(\boldsymbol{D}^{1/2}\overline{\boldsymbol{W}}\boldsymbol{D}^{1/2}) \geq c, \overline{\boldsymbol{W}} \geq \boldsymbol{O}\} \end{aligned}$$

and

$$d\mu(\overline{\boldsymbol{W}}) = \frac{(d\overline{\boldsymbol{W}})}{|\overline{\boldsymbol{W}}|^{\frac{1}{2}(p+1)}}$$

the invariant measure. Then for $\boldsymbol{\delta} \geq \boldsymbol{e}$ (i.e. $\boldsymbol{D} \geq \boldsymbol{I}_p$) we have

$$\begin{aligned} &\beta_{01}(\boldsymbol{\delta}) - \beta_{01}(\boldsymbol{e}) \\ &\quad = K''\{\int_{\omega_0} - \int_{\omega_1}\}f^+(\overline{\boldsymbol{W}})f^-(\overline{\boldsymbol{W}})d\mu(\overline{\boldsymbol{W}}) \\ &\quad = K''\{\int_{\omega_0-\omega_0\cap\omega_1} - \int_{\omega_1-\omega_0\cap\omega_1}\}f^+(\overline{\boldsymbol{W}})f^-(\overline{\boldsymbol{W}})d\mu(\overline{\boldsymbol{W}}) \\ &\quad \geq cK''\{\int_{\omega_0-\omega_0\cap\omega_1} - \int_{\omega_1-\omega_0\cap\omega_1}\}f^-(\overline{\boldsymbol{W}})d\mu(\overline{\boldsymbol{W}}) \\ &\quad = cK''\{\int_{\omega_0} - \int_{\omega_1}\}f^-(\overline{\boldsymbol{W}})d\mu(\overline{\boldsymbol{W}}) \\ &\quad = cK''\int_{\omega_0}\{f^-(\overline{\boldsymbol{W}}) - f^-(\boldsymbol{D}^{-1/2}\overline{\boldsymbol{W}}\boldsymbol{D}^{-1/2})\}d\mu(\overline{\boldsymbol{W}}), \end{aligned} \tag{2.11}$$

since

$$\int_{\omega_0\cap\omega_1} f^-(\overline{\boldsymbol{W}})d\mu(\overline{\boldsymbol{W}}) \le \frac{1}{c}\int f^+(\overline{\boldsymbol{W}})f^-(\overline{\boldsymbol{W}})d\mu(\overline{\boldsymbol{W}}) < \infty.$$

The R.H.S. of (2.11) is nonnegative because (2.7) yields

$$f^-(\overline{\boldsymbol{W}}) - f^-(\boldsymbol{D}^{-1/2}\overline{\boldsymbol{W}}\boldsymbol{D}^{-1/2}) \ge 0 \tag{2.12}$$

for $\boldsymbol{D} \ge \boldsymbol{I}_p$. Furthermore R.H.S. of (2.11) is positive when $\boldsymbol{\delta} \ne \boldsymbol{e}$, because the inequality of (2.12) holds strictly on the set $\{m^* = 0\}$ and the measure $\mu(\omega_0 \cap \{m^* = 0\})$ is not zero. □

3. Limiting null distribution of the test statistics of LRT.

To obtain the significance points of the LRT's T_{01} and T_{12}, we have to derive the distribution of Λ_{01} and Λ_{12} under H_0 because the least favorable distribution of Λ_{12} is given under H_0. This section gives the limiting joint distribution of $-2\log\Lambda_{01}$ and $-2\log\Lambda_{12}$ under H_0 as n goes to infinity. We use Lemma 3.1 of Chapter II to obtain it.

Before deriving the distributions, we rewrite Λ_{01} (2.1) and Λ_{12} (2.4) in terms of

$$u_i = \frac{t_i}{s_i} = \frac{\bar{t}_i}{\bar{s}_i}, \quad 1 \le i \le p.$$

Since (2.5) yields

$$\bar{s}_i = \frac{\bar{t}_{i+1}}{\nu+p-i} + \frac{\nu+p-i-1}{\nu+p-i}\bar{s}_{i+1},$$

we have

$$\frac{\bar{s}_i}{\bar{s}_{i+1}} = 1 + \frac{u_{i+1}-1}{\nu+p-i}. \tag{3.1}$$

Then we get

$$\begin{aligned}
-2\log\Lambda_{01} =& n\sum_{i=1}^{m^*}\{-\log\frac{\bar{t}_i}{\bar{s}_i} + (\nu+p-i+1)\log\frac{\bar{s}_{i-1}}{\bar{s}_i}\} \\
=& \sum_{i=1}^{p} I(u_i > 1)g_i(u_i)
\end{aligned}$$

where

$$g_i(u) = n\{-\log u + (\nu+p-i+1)\log(1+\frac{u-1}{\nu+p-i+1})\}$$

and $I(\,\cdot\,)$ indicator function. Similarly we get

$$-2\log\Lambda_{12} = \sum_{i=1}^{p} I(u_i < 1)g_i(u_i).$$

The function $g_i(u)$ is decreasing when $u < 1$, and increasing when $u > 1$, and hence the LRT statistics $-2\log\Lambda_{01}$ and $-2\log\Lambda_{12}$ are nondecreasing and nonincreasing functions of u_i, $1 \le i \le p$, respectively.

Since $\bar{t}_1 > \cdots > \bar{t}_p$ are $(1/n)$ times the latent roots of $\overline{\boldsymbol{W}} = (1/g)\boldsymbol{W}$ whose density is of the form of (2.10), joint null density of $\bar{t}_1 > \cdots > \bar{t}_p$ is given by

$$K(n,p,\nu)\prod_{i=1}^{p}\bar{t}_i^{\,(n-p-1)/2}(1+\frac{\sum_{i=1}^{p}\bar{t}_i}{\nu})^{-n(\nu+p)/2}\prod_{i<j}(\bar{t}_i-\bar{t}_j)_+, \tag{3.2}$$

where

$$K(n,p,\nu) = \frac{\pi^{p^2/2}\Gamma(\frac{n(\nu+p)}{2})}{\Gamma_p(\frac{p}{2})\Gamma_p(\frac{n}{2})\Gamma(\frac{n\nu}{2})\nu^{np/2}} \tag{3.3}$$

the normalizing constant, and $(x)_+ = x$ if $x > 0$, 0 otherwise. Here from (3.1) and the fact $\bar{s}_p = 1$ it holds

$$\bar{s}_i = \prod_{j=i+1}^{p}(1+\frac{u_j-1}{\nu+p-j+1})$$

and

$$\bar{t}_i = u_i\prod_{j=i+1}^{p}(1+\frac{u_j-1}{\nu+p-j+1}),$$

then the Jacobian of the transformation is given by

$$\left|\frac{\partial(\bar{t}_1,\ldots,\bar{t}_p)}{\partial(u_1,\ldots,u_p)}\right| = \prod_{i=1}^{p}\bar{s}_i. \tag{3.4}$$

From (3.2) and (3.4) the joint null density φ_0 of $\boldsymbol{u} = (u_1,\ldots,u_p)$ is written as

$$\varphi_0(\boldsymbol{u}) = K(n,p,\nu)\prod_{i=1}^{p}\bar{t}_i^{\,(n-p-1)/2}(1+\frac{\sum_{i=1}^{p}\bar{t}_i}{\nu})^{-n(\nu+p)/2}\prod_{i<j}(\bar{t}_i-\bar{t}_j)_+\prod_{i=1}^{p}\bar{s}_i. \tag{3.5}$$

Define

$$b_i = (n/2)^{1/2}\gamma_i^{-1}(u_i-1), \quad \gamma_i = \left(\frac{\nu+p-i+1}{\nu+p-i}\right)^{1/2}, \quad 1 \le i \le p.$$

Then the joint null density of $\boldsymbol{b} = (b_1,\ldots,b_p)$ is

$$\varphi_0(\boldsymbol{u}(\boldsymbol{b}))\cdot(n/2)^{-p/2}(\frac{\nu}{\nu+p})^{-1/2}. \tag{3.6}$$

As $n \to \infty$ with b_i fixed (i.e. $u_i \to 1$), one can see that $\bar{t}_i \to 1$, $\bar{s}_i \to 1$, $g_i(u_i) \to {b_i}^2$,

$$(\frac{\nu}{\nu+p})^{-n(\nu+p)/2} \prod_{i=1}^{p} {\bar{t}_i}^{n/2} (1 + \frac{\sum_{i=1}^{p} \bar{t}_i}{\nu})^{-n(\nu+p)/2} = \exp\{-\frac{1}{2}\sum_{i=1}^{p} g_i(u_i)\}$$
$$\to \exp\{-\frac{1}{2}\sum_{i=1}^{p} {b_i}^2\}$$

and

$$(n/2)^{1/2}(\bar{t}_i - \bar{t}_{i+1}) \to \gamma_i b_i - {\gamma_{i+1}}^{-1} b_{i+1}.$$

By Stirling's formula it holds

$$K(n,p,\nu)(n/2)^{-p(p+1)/4}(\frac{\nu}{\nu+p})^{n(\nu+p)/2-1/2} \to d(p) \quad \text{as } n \to \infty$$

with

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

As a result the joint density function of $\boldsymbol{b}$ of (3.6) is shown to converge to

$$d(p)\exp\{-\frac{1}{2}\sum_{i=1}^{p} {b_i}^2\} \prod_{i<j} \{\sum_{k=i}^{j-1} (\gamma_k b_k - {\gamma_{k+1}}^{-1} b_{k+1})\}_+ \tag{3.7}$$

for each $\boldsymbol{b} \in \mathbf{R}^p$. On the other hand we can see that

$$-2\log\Lambda_{01} = \sum_{i=1}^{p} I(u_i > 1) g_i(u_i) \to \sum_{i=1}^{p} (b_i \vee 0)^2$$

and

$$-2\log\Lambda_{12} = \sum_{i=1}^{p} I(u_i < 1) g_i(u_i) \to \sum_{i=1}^{p} (b_i \wedge 0)^2$$

for each $\boldsymbol{b} \in \mathbf{R}^p$, where $x \vee y$ and $x \wedge y$ are the maximum and the minimum of x and y, respectively. Following lemma shows that (3.7) is also a density function.

Lemma 3.1 (3.7) is a joint density of

$$b_i = {\gamma_i}^{-1}\{(a_i - h) - \frac{1}{\nu+p-i}\sum_{j=i+1}^{p}(a_j - h)\}, \quad 1 \le i \le p, \tag{3.8}$$

where $a_1 > \cdots > a_p$ are the latent roots of $p \times p$ symmetric random matrix $\boldsymbol{A}$ with the normal density

$$\frac{1}{2^{p/2}\pi^{p(p+1)/4}} \exp\{-\frac{1}{2}\mathrm{tr}\boldsymbol{A}^2\}$$

and h is a random variable distributed independently according to $N(0, 1/\nu)$.

(Proof) The joint density of $a_1 > \cdots > a_p$ is

$$d(p) \exp\{-\frac{1}{2}\sum_{i=1}^{p} {a_i}^2\} \prod_{i<j}(a_i - a_j)_+.$$

(Anderson, 1984b, Theorem 13.3.5.) Making the change of variables from $(a_1, \ldots, a_p, h)$ to $(f_1, \ldots, f_p, h)$ by $f_i = a_i - h$, and integrating the joint density with respect to h, we get the marginal density of $f_1 > \cdots > f_p$ as

$$d(p)(\frac{\nu}{\nu+p})^{1/2} \exp\Big\{-\frac{1}{2}\Big(\sum_{i=1}^{p} {f_i}^2 - \frac{(\sum_{i=1}^{p} f_i)^2}{\nu+p}\Big)\Big\} \prod_{i<j}(f_i - f_j)_+. \tag{3.9}$$

Put $e_i = \sum_{j=i+1}^{p} f_j/(\nu + p - i)$, $0 \le i \le p-1$, and $e_p = 0$. Noting that $b_i = {\gamma_i}^{-1}(f_i - e_i)$ and

$$f_i - e_i = (\nu + p - i + 1)(e_{i-1} - e_i), \tag{3.10}$$

we can see that

$${b_i}^2 = {f_i}^2 - (\nu + p - i + 1){e_{i-1}}^2 + (\nu + p - i){e_i}^2,$$

and hence

$$\sum_{i=1}^{p} {b_i}^2 = \sum_{i=1}^{p} {f_i}^2 - (\nu+p){e_0}^2 = \sum_{i=1}^{p} {f_i}^2 - \frac{(\sum_{i=1}^{p} f_i)^2}{\nu+p}. \tag{3.11}$$

From (3.10) it holds

$$e_i = \sum_{j=i+1}^{p} (e_{j-1} - e_j) = \sum_{j=i+1}^{p} \frac{\gamma_j}{\nu + p - j + 1} b_j$$

and

$$f_i = \gamma_i b_i + e_i = \gamma_i b_i + \sum_{j=i+1}^{p} \frac{\gamma_j}{\nu + p - j + 1} b_j.$$

Then we have

$$f_i - f_{i+1} = \gamma_i b_i - {\gamma_{i+1}}^{-1} b_{i+1}. \tag{3.12}$$

From (3.9), (3.11), (3.12) and the fact $\prod_i db_i = \{\nu/(\nu+p)\}^{1/2} \prod_i df_i$, we see that (3.7) is the joint density of $(b_1, \ldots, b_p)$ defined in (3.8). The proof is completed. □

By applying Lemma 3.1 of Chapter II, we obtain the following theorem:

Theorem 3.1 As n goes to infinity, $(-2\log\Lambda_{01}, -2\log\Lambda_{12})$ converges to $(\sum_{i=1}^{p}(b_i \vee 0)^2, \sum_{i=1}^{p}(b_i \wedge 0)^2)$ in distribution under H_0, where $(b_1, \ldots, b_p)$ are distributed with density (3.7).

The joint characteristic function of the limiting distribution is given by

$$\begin{aligned}\phi(s,t) =&\mathrm{E}[\exp\{is\sum_{i=1}^{p}(b_i\vee 0)^2+it\sum_{i=1}^{p}(b_i\wedge 0)^2\}]\\ =&\sum_{m=0}^{p}\mathrm{E}[I(m^*=m)\exp\{is\sum_{i=1}^{m}{b_i}^2+it\sum_{i=m+1}^{p}{b_i}^2\}]\\ =&d(p)\sum_{m=0}^{p}\int_{B_m\times\overline{B}_{p-m}}\exp\{-\frac{1}{2}(1-2is)\sum_{i=1}^{m}{b_i}^2-\frac{1}{2}(1-2it)\sum_{i=m+1}^{p}{b_i}^2\}\\ &\times\Delta(b_1,\dots,b_p)db_1\cdots db_p, \end{aligned} \tag{3.13}$$

where m^* is a random integer such that $b_{m^*}\geq 0 > b_{m^*+1}$,

$$\Delta(b_1,\dots,b_p)=\prod_{i<j}\{\sum_{k=i}^{j-1}(\gamma_k b_k-{\gamma_{k+1}}^{-1}b_{k+1})\} \tag{3.14}$$

and

$$B_m=\{(b_1,\dots,b_m)\mid \gamma_k b_k-{\gamma_{k+1}}^{-1}b_{k+1}>0,\ 1\leq k\leq m-1,\ \ b_m>0\},$$

$$\overline{B}_{p-m}=\{(b_{m+1},\dots,b_p)\mid \gamma_k b_k-{\gamma_{k+1}}^{-1}b_{k+1}>0,\ m+1\leq k\leq p-1,\ \ b_{m+1}<0\}.$$

Putting

$$c_i={c_i}^{(m)}(b_i,\dots,b_m)=\sum_{k=i}^{m-1}(\gamma_k b_k-{\gamma_{k+1}}^{-1}b_{k+1})+\gamma_m b_m,\quad 1\leq i\leq m,$$

$$\begin{aligned} d_j={d_j}^{(m)}(b_{m+1},\dots,b_{p-j+1})={\gamma_{m+1}}^{-1}(-b_{m+1})+\sum_{k=m+1}^{p-j}(\gamma_k b_k-{\gamma_{k+1}}^{-1}b_{k+1}),\\ 1\leq j\leq p-m,\end{aligned}$$

we have that (3.14) is

$$\prod_{1\leq i<j\leq m}(c_i-c_j)\times\prod_{1\leq i<j\leq p-m}(d_i-d_j)\times\prod_{i=1}^{m}\prod_{j=1}^{p-m}(c_i+d_j). \tag{3.15}$$

By Macdonald (1979, p.37, Example 5) (3.15) can be expressed as

$$\sum_{q}\det({c_i}^{q_j})_{1\leq i,j\leq m}\det({d_i}^{\overline{q}_j})_{1\leq i,j\leq p-m}, \tag{3.16}$$

where the summation $\sum_q$ is over all combinations $p-1\geq q_1>\cdots>q_m\geq 0$, $p-1\geq\overline{q}_1>\cdots>\overline{q}_{p-m}\geq 0$ such that

$$\{q_1,\dots,q_m,\overline{q}_1,\dots,\overline{q}_{p-m}\}=\{0,1,\dots,p-1\}.$$

Using (3.16), (3.13) reduces to

$$\phi(s,t) = d(p)\sum_{m=0}^{p}\sum_{q} U_m\overline{U}_{p-m}(1-2is)^{-Q/2}(1-2it)^{-\overline{Q}/2}, \tag{3.17}$$

where $Q = \sum_{i=1}^{m} q_i + m$, $\overline{Q} = \sum_{i=1}^{p-m} \overline{q}_i + p - m = p(p+1)/2 - Q$,

$$U_m = U_m(q_1,\ldots,q_m) = \int_{B_m} \exp\{-\frac{1}{2}\sum_{i=1}^{m} {b_i}^2\}\det({c_i}^{q_j})_{1\le i,j\le m} db_1\cdots db_m \tag{3.18}$$

for $m \ge 1$ and

$$\begin{aligned}\overline{U}_{p-m} =& \overline{U}_{p-m}(\overline{q}_1,\ldots,\overline{q}_m)\\ =& \int_{\overline{B}_{p-m}} \exp\{-\frac{1}{2}\sum_{i=1}^{p-m} {b_{i+m}}^2\}\det({d_i}^{\overline{q}_j})_{1\le i,j\le p-m} db_{m+1}\cdots db_p\end{aligned} \tag{3.19}$$

for $p - m \ge 1$, $U_0 = \overline{U}_0 = 1$. By inverting the characteristic function (3.17), we get the following formula.

Theorem 3.2 As n goes to infinity, the limiting joint distribution function of $-2\log\Lambda_{01}$ and $-2\log\Lambda_{12}$ under H_0 is given by

$$\begin{aligned}&\lim_{n\to\infty} \mathrm{P}_{H_0}(-2\log\Lambda_{01}\le y, -2\log\Lambda_{12}\le z)\\ &\quad = d(p)\sum_{m=0}^{p}\sum_{q} U_m\overline{U}_{p-m}G_Q(y)G_{\overline{Q}}(z)\end{aligned}$$

where $G_\nu(\,\cdot\,)$, $\nu > 0$, is the distribution function of the chi-squared distribution with ν degrees of freedom and $G_0(\,\cdot\,) = I(0 \le\,\cdot\,)$.

Differently from the model discussed in Chapter II it is not easy to evaluate the integrals (3.18) and (3.19) when p is not small. Appendix A gives the limiting characteristic functions (3.17) for $p = 2, 3$.

4. Local unbiasedness of a general class of tests for H_0 against $H_1 - H_0$.

In Section 3 the LRT statistic Λ_{01} was shown to be a nondecreasing function of u_i, $1 \le i \le p$. Here we consider a a more general class $\mathcal{C}$ of the critical function $\phi : (0,\infty)^p \to [0,1]$ which is nondecreasing in each argument of $\boldsymbol{u} \in \{\varphi_0 > 0\}$, and not constant on $\{\varphi_0 > 0\}$. The tests based on the statistics u_1, $\sum_{i=1}^{p} \overline{t}_i$ (F-test) are also in the class $\mathcal{C}$. In this section we prove that the test in the class $\mathcal{C}$ is locally unbiased as the test for H_0 against $H_1 - H_0$ using the FKG inequality which is summarized in Appendix B.

Before proving the theorem we provide a following lemma.

Lemma 4.1 The density φ_0 in (3.5) on $(0,\infty)^p$ satisfies (B.1) (FKG condition).

(Proof) It is sufficient to prove (a) and (b) of Lemma B.1.
Proof of (a) : Note that $\varphi_0 > 0 \Leftrightarrow \bar{t}_i - \bar{t}_{i+1} > 0$, $1 \le i \le p-1 \Leftrightarrow u_i - g_{i+1}(u_{i+1}) > 0$, $1 \le i \le p-1$ with

$$g_{i+1}(u) = u\left(1 + \frac{u-1}{\nu+p-i}\right)^{-1}.$$

Because the function $g_{i+1}(u)$ is an increasing function on $u > 0$, $\{\varphi_0 > 0\}$ is of the form of (B.4).
Proof of (b) : On the set $\{\varphi_0 > 0\}$, $\log \varphi_0$ can be expressed as $\sum_{i=1}^{p} \eta_i(u_i) + L(\boldsymbol{u})$ for some functions η_i and $L(\boldsymbol{u}) = \sum_{i<j} \log(\bar{t}_i - \bar{t}_j)$. Therefore we only have to show that $\partial^2 L/\partial u_l \partial u_m \ge 0$, $l < m$, on $\{\varphi_0 > 0\}$. Note first that

$$\begin{aligned} \frac{\partial \bar{t}_i}{\partial u_j} &= \frac{\bar{t}_i}{u_j + \nu + p - j} && \text{if } i < j, \\ &= \frac{\bar{t}_i}{u_i} && \text{if } i = j, \\ &= 0 && \text{otherwise.} \end{aligned}$$

Using the above relations some straightforward calculations yield

$$\frac{\partial^2 L}{\partial u_l \partial u_m} = \sum_{i<l} \frac{\bar{t}_i}{u_l + \nu + p - l} E_{im} + \frac{\bar{t}_l}{u_l} E_{lm}, \quad l < m,$$

with

$$E_{im} = \left(\frac{\bar{t}_m}{u_m} - \frac{\bar{t}_m}{u_m + \nu + p - m}\right)\frac{1}{(\bar{t}_i - \bar{t}_m)^2} - \frac{1}{u_m + \nu + p - m}\sum_{j>m} \frac{\bar{t}_j}{(\bar{t}_i - \bar{t}_j)^2}$$

for $i \le l$. Noting that for $i < m$,

$$\sum_{j>m} \frac{\bar{t}_j}{(\bar{t}_i - \bar{t}_j)^2} < \frac{1}{(\bar{t}_i - \bar{t}_m)^2} \sum_{j>m} \bar{t}_j = \frac{1}{(\bar{t}_i - \bar{t}_m)^2}\{(\nu + p - m)\frac{\bar{t}_m}{u_m} - \nu\},$$

then

$$E_{im} > \frac{\nu}{(\bar{t}_i - \bar{t}_m)^2 (u_m + \nu + p - m)} > 0,$$

which implies $\partial^2 L/\partial u_l \partial u_m > 0$. The proof is completed. □

Remark 4.1 The limiting joint density of (3.7) is also proved to satisfy (B.1) by the same way.

Using Lemma 4.1 we can prove the following result.

Theorem 4.1 Let ϕ be a critical function in the class $\mathcal{C}$, and let $\beta(\boldsymbol{\delta})$ be the power function with $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_p)$ the latent roots of $(1/\sigma^2)\boldsymbol{\Phi}$. Then

$$\frac{\partial}{\partial \lambda}\beta(\boldsymbol{e} + \lambda \boldsymbol{c})\bigg|_{\lambda=0} > 0 \qquad \text{if } \sum_{i=1}^{p} c_i > 0,$$

with $\boldsymbol{e} = (1, \ldots, 1)$ and $\boldsymbol{c} = (c_1, \ldots, c_p)$. This means that the test ϕ is locally unbiased as a test for H_0 $(\delta_i \equiv 1)$ against the alternative $\sum_{i=1}^{p} \delta_i > p$ (which includes $H_1 - H_0$).

(Proof) The non-null joint density φ of $\boldsymbol{u} = (u_1, \ldots, u_p)$ is given by

$$\begin{aligned}\varphi(\boldsymbol{u}; \boldsymbol{\delta}) =& K(n, p, \nu)|\boldsymbol{D}|^{-n/2} \prod_{i=1}^{p} \bar{t}_i^{\,(n-p-1)/2} \prod_{i<j}(\bar{t}_i - \bar{t}_j)_+ \prod_{i=1}^{p} \bar{s}_i \\ &\times \int_{O(p)} (d\boldsymbol{H})(1 + \frac{1}{\nu}\mathrm{tr}\boldsymbol{D}^{-1}\boldsymbol{H}\overline{\boldsymbol{T}}\boldsymbol{H}')^{-n(\nu+p)/2}\end{aligned}$$

where $\overline{\boldsymbol{T}} = \mathrm{diag}(\bar{t}_1, \ldots, \bar{t}_p)$ $p \times p$ diagonal matrix, $\boldsymbol{D} = \mathrm{diag}(\delta_1, \ldots, \delta_p)$ $p \times p$ diagonal matrix, $(d\boldsymbol{H})$ the normalized invariant measure over $O(p)$ such that $\int_{O(p)}(d\boldsymbol{H}) = 1$, $K(n, p, \nu)$ the normalizing constant given in (3.3). Then

$$\begin{aligned}\frac{\partial}{\partial \delta_i}\varphi(\boldsymbol{u}; \boldsymbol{\delta})\bigg|_{\boldsymbol{\delta}=\boldsymbol{e}} =& \left\{\frac{n(\nu+p)}{2\nu}(1 + \frac{1}{\nu}\sum_j \bar{t}_j)^{-1}\int_{O(p)}(d\boldsymbol{H})\sum_j h_{ij}{}^2\bar{t}_j - \frac{n}{2}\right\}\varphi_0(\boldsymbol{u}) \\ =& h(\boldsymbol{u})\varphi_0(\boldsymbol{u}),\end{aligned}$$

where h_{ij} is (i, j)-th element of $\boldsymbol{H}$ and

$$h(\boldsymbol{u}) = \frac{n}{2}\left(\frac{\nu+p}{p}\frac{\sum \bar{t}_j}{\nu + \sum \bar{t}_j} - 1\right).$$

Here $h(\boldsymbol{u})$ is strictly increasing in each argument of $\boldsymbol{u} \in \{\varphi_0 > 0\}$ and

$$\int h(\boldsymbol{u})\varphi_0(\boldsymbol{u})d\boldsymbol{u} = 0.$$

By virtue of Lemma 4.1, we apply Theorem B.1 (FKG inequality) as

$$\begin{aligned}\frac{\partial}{\partial \lambda}\beta(\boldsymbol{e} + \lambda\boldsymbol{c})\bigg|_{\lambda=0} &= (\sum c_i)\int \phi(\boldsymbol{u})h(\boldsymbol{u})\varphi_0(\boldsymbol{u})d\boldsymbol{u} \\ &> (\sum c_i)\int \phi(\boldsymbol{u})\varphi_0(\boldsymbol{u})d\boldsymbol{u} \cdot \int h(\boldsymbol{u})\varphi_0(\boldsymbol{u})d\boldsymbol{u}. \\ &= 0.\end{aligned}$$

The proof is completed. □

5. Applications to testing problems in random coefficient regression model.

In this section we state applications of the LRT's T_{01} and T_{12} to the following random coefficient regression model proposed by Rao (1965):

$$\boldsymbol{y}_i = \boldsymbol{F}\boldsymbol{u}_i + \boldsymbol{e}_i, \quad i = 1, \ldots, N, \tag{5.1}$$

where $\boldsymbol{y}_i$ is a $k \times 1$ observed vector, $\boldsymbol{F}$ a $k \times p$ ($k > p$) design matrix with rank p, $\boldsymbol{u}_i$ a $p \times 1$ unobserved random effect vector, and $\boldsymbol{e}_i$ a $k \times 1$ unobserved random effect measurement error. $\boldsymbol{u}_i$ and $\boldsymbol{e}_i$ are independently distributed according to the normal distributions $N_p(\boldsymbol{\zeta}, \boldsymbol{\Lambda})$ and $N_k(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_k)$, respectively. This is the balanced case because the design matrix $\boldsymbol{F}$ does not depend on i. If $\boldsymbol{F}' = (\boldsymbol{I}_p \ldots \boldsymbol{I}_p)_{p \times pr}$ ($k = pr$), (5.1) reduces to

$$\boldsymbol{y}_{ij} = \boldsymbol{u}_i + \boldsymbol{e}_{ij}, \quad i = 1, \ldots, N, \quad j = 1, \ldots, r,$$

where $\boldsymbol{y}_{ij}$ and $\boldsymbol{e}_{ij}$ are $p \times 1$ vector such that $\boldsymbol{y}_i{}' = (\boldsymbol{y}_{i1}{}' \ldots \boldsymbol{y}_{ir}{}')$, $\boldsymbol{e}_i{}' = (\boldsymbol{e}_{i1}{}' \ldots \boldsymbol{e}_{ir}{}')$. This is Scheffé's mixed model with replications (Scheffé, 1959, Chapter 8) discussed in Section 3 of Anderson, et al. (1986). For other applications of the model (5.1), see Johansen (1984, Chapter 4), Lange and Laird (1989), Crowder and Hand (1990, Chapter 6), etc.

We start with transforming (5.1) into the canonical form. Let $\boldsymbol{H} = (\boldsymbol{H}_1\, \boldsymbol{H}_2)$ be a $k \times k$ orthogonal matrix such that $\boldsymbol{H}_1$ is $k \times p$, $\boldsymbol{H}_2$ is $k \times (k-p)$, and $\boldsymbol{H}_2{}'\boldsymbol{F} = \boldsymbol{O}$. Putting $\boldsymbol{z}_i = \boldsymbol{H}'\boldsymbol{y}_i$, $i = 1, \ldots, N$, $\boldsymbol{\xi}_1 = \boldsymbol{G}\boldsymbol{\zeta}$, $\boldsymbol{\Theta} = \boldsymbol{G}\boldsymbol{\Lambda}\boldsymbol{G}'$ with $\boldsymbol{G} = \boldsymbol{H}_1{}'\boldsymbol{F}$, we have the canonical model

$$M_1 : \quad \boldsymbol{z}_i \sim N_k\left(\begin{pmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Theta} + \sigma^2 \boldsymbol{I}_p & \boldsymbol{O} \\ \boldsymbol{O} & \sigma^2 \boldsymbol{I}_{k-p} \end{pmatrix} \right), \quad i = 1, \ldots, N, \quad \text{i.i.d.}$$

Then, as a test for the hypothesis that the coefficients $\boldsymbol{u}_i$ are constant, i.e. the hypothesis $\boldsymbol{\Lambda} = \boldsymbol{O}$, we can use the LRT T_{01} for testing $\boldsymbol{\Theta} = \boldsymbol{O}$ based on the statistic Λ_{01} (2.1) with

$$\boldsymbol{W} = \boldsymbol{S}_{11} \quad \text{and} \quad g = \frac{1}{N(k-p)} \operatorname{tr}(\boldsymbol{S}_{22} + \boldsymbol{T}_{22}), \tag{5.2}$$

where

$$\boldsymbol{S} = \begin{pmatrix} \boldsymbol{S}_{11} & \boldsymbol{S}_{12} \\ \boldsymbol{S}_{21} & \boldsymbol{S}_{22} \end{pmatrix} = \sum_{i=1}^{N} (\boldsymbol{z}_i - \overline{\boldsymbol{z}})(\boldsymbol{z}_i - \overline{\boldsymbol{z}})',$$

$$\boldsymbol{T} = \begin{pmatrix} \boldsymbol{T}_{11} & \boldsymbol{T}_{12} \\ \boldsymbol{T}_{21} & \boldsymbol{T}_{22} \end{pmatrix} = N \overline{\boldsymbol{z}}\, \overline{\boldsymbol{z}}', \quad \overline{\boldsymbol{z}} = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{z}_i.$$

Next we consider the more general model than M_1 as

$$M_3: \quad \boldsymbol{z}_i \sim N_k\left(\begin{pmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_2 \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}\right), \quad i = 1, \ldots, N, \quad \text{i.i.d.}$$

and obtain a test for goodness of fit of M_1 against M_3 (Johansen, 1984, Section 4). We settle an intermediate model between the models M_1 and M_3 as

$$M_2: \quad \boldsymbol{z}_i \sim N_k\left(\begin{pmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{O} \\ \boldsymbol{O} & \sigma^2 \boldsymbol{I}_{k-p} \end{pmatrix}\right), \quad i = 1, \ldots, N, \quad \text{i.i.d.}$$

and divide this testing problem into two parts: testing goodness of fit of M_1 against M_2, and testing goodness of fit of M_2 against M_3. For testing M_1 against M_2 we can use the LRT T_{12} based on the statistic Λ_{12} with $\boldsymbol{W}$ and g in (5.2). For testing M_2 against M_3 likelihood ratio criterion can be easily obtained as

$$\Lambda_{23} = \left\{ \frac{|\boldsymbol{S}|}{|\boldsymbol{S}_{11}||\boldsymbol{S}_{22}|} \times \frac{|\boldsymbol{S}_{22}|}{|\boldsymbol{S}_{22} + \boldsymbol{T}_{22}|} \times \frac{|\boldsymbol{S}_{22} + \boldsymbol{T}_{22}|}{\{\frac{1}{k-p}\text{tr}(\boldsymbol{S}_{22} + \boldsymbol{T}_{22})\}^{k-p}} \right\}^{N/2}.$$

Note that three components of Λ_{23} combined by the symbols '$\times$' are familiar likelihood ratio criteria for testing the hypotheses $\boldsymbol{\Sigma}_{12} = \boldsymbol{O}$, $\boldsymbol{\xi}_2 = \boldsymbol{0}$ and $\boldsymbol{\Sigma}_{22} = \sigma^2 \boldsymbol{I}_{k-p}$, respectively.

Concerning the null distribution of Λ_{12} and Λ_{23}, it holds that:

Lemma 5.1 Under the null model M_1, Λ_{12} and three components of Λ_{23} are mutually independent.

(Proof) Under the model M_1, four matrices $\boldsymbol{S}_{11}$, $\boldsymbol{S}_{22\cdot1} = \boldsymbol{S}_{22} - \boldsymbol{S}_{21}{\boldsymbol{S}_{11}}^{-1}\boldsymbol{S}_{12}$, $\boldsymbol{B} = \boldsymbol{S}_{21}{\boldsymbol{S}_{11}}^{-1}\boldsymbol{S}_{12}$ and $\boldsymbol{T}_{22}$ are distributed independently according to $W_p(N-1, \boldsymbol{\Theta} + \sigma^2 \boldsymbol{I}_p)$, $W_{k-p}(N-p-1, \sigma^2 \boldsymbol{I}_{k-p})$, $W_{k-p}(p, \sigma^2 \boldsymbol{I}_{k-p})$, and $W_{k-p}(1, \sigma^2 \boldsymbol{I}_{k-p})$, respectively. Noting that Λ_{12} is a function of $\boldsymbol{S}_{11}$ and $\text{tr}(\boldsymbol{S}_{22\cdot1} + \boldsymbol{B} + \boldsymbol{T}_{22})$, and that

$$\Lambda_{23} = \left\{ \frac{|\boldsymbol{S}_{22\cdot1}|}{|\boldsymbol{S}_{22\cdot1} + \boldsymbol{B}|} \times \frac{|\boldsymbol{S}_{22\cdot1} + \boldsymbol{B}|}{|\boldsymbol{S}_{22\cdot1} + \boldsymbol{B} + \boldsymbol{T}_{22}|} \times \frac{|\boldsymbol{S}_{22\cdot1} + \boldsymbol{B} + \boldsymbol{T}_{22}|}{\{\frac{1}{k-p}\text{tr}(\boldsymbol{S}_{22\cdot1} + \boldsymbol{B} + \boldsymbol{T}_{22})\}^{k-p}} \right\}^{N/2},$$

we can obtain the proof easily. □

Then by choosing α' and α'' such that $1 - \alpha = (1 - \alpha')(1 - \alpha'')$, we can construct a level-α test for goodness of fit of the model M_1 as a step-down testing procedure combining: a level-α' test based on Λ_{23} (or its three components); and a level-α'' test based on Λ_{12}. For step-down testing procedures, see Anderson (1984b, Section 9.6).

6. Power comparisons.

In this section we study a Monte Carlo simulation to compare the powers of several tests including the LRT. The null hypothesis is $\boldsymbol{\Phi} = \sigma^2 \boldsymbol{I}_p$, and the alternative is local hypothesis that

$$\boldsymbol{\Phi} = \sigma^2 \left(\boldsymbol{I}_p + \sqrt{\frac{2}{n}} \boldsymbol{\Delta} \right)$$

with $\boldsymbol{\Delta} = \mathrm{diag}(\delta_i)_{p \times p}$, $\delta_i \geq 0$. We compare the limiting powers ($n \to \infty$) of four test criteria:

- · One-sided likelihood ratio test based on the statistic Λ_{01} [ONE];
- · Two-sided likelihood ratio test based on the statistic $\Lambda_{02} = \Lambda_{01}\Lambda_{12}$ [TWO];
- · Test based on the statistic u_1 [MAX];
- · Test based on the F-statistic $\mathrm{tr}\boldsymbol{H}/g$ [F].

The limiting power functions of the four tests are

$$\Pr(\sum_{i=0}^{p} (b_i \vee 0)^2 > c), \quad \Pr(\sum_{i=0}^{p} {b_i}^2 > c'), \quad \Pr(b_1 > c''), \quad \text{and} \quad \Pr(\sum_{i=0}^{p} b_i > c'''),$$

respectively, where

$$b_i = \sqrt{\frac{\nu + p - i}{\nu + p - i + 1}} \{(a_i - h) - \frac{1}{\nu + p - i} \sum_{j=i+1}^{p} (a_j - h)\}, \quad 1 \leq i \leq p;$$

$a_1 > \cdots > a_p$ are the latent roots of $p \times p$ symmetric random matrix $\boldsymbol{A}$ with normal density

$$\frac{1}{2^{p/2} \pi^{p^2/2}} \mathrm{tr}(\boldsymbol{A} - \boldsymbol{\Delta})^2,$$

and h is a random variable independently distributed according to $N(0, 1/\nu)$.

The average powers of the four tests over 100000 replications are given in Table 6.1 (for $p = 4$, $\nu = 1$, size= 5%) and Table 6.2 (for $p = 8$, $\nu = 1$, size= 5%). The results indicate that the performance of ONE, TWO, MAX and F are very similar to ONE, TWO, ROY and LMP of Section 6 in Chapter II, respectively.

Table 6.1

Power (%, $p = 4$, $\nu = 1$, size=5%)

δ_1	δ_2	δ_3	δ_4	ONE	TWO	MAX	F
0.0	0.0	0.0	0.0	5	5	5	5
1.0	0.0	0.0	0.0	<u>9</u>	7	8	8
2.0	0.0	0.0	0.0	21	18	<u>22</u>	11
4.0	0.0	0.0	0.0	73	67	<u>78</u>	23
1.0	1.0	0.0	0.0	<u>12</u>	9	10	11
2.0	2.0	0.0	0.0	<u>33</u>	25	24	23
4.0	4.0	0.0	0.0	<u>92</u>	87	80	55
1.0	1.0	1.0	0.0	14	9	10	<u>16</u>
2.0	2.0	2.0	0.0	37	26	21	<u>38</u>
1.0	1.0	1.0	1.0	14	7	9	<u>23</u>
1.0	0.5	0.0	0.0	<u>10</u>	8	8	9
2.0	1.0	0.0	0.0	<u>23</u>	17	20	16
4.0	2.0	0.0	0.0	<u>76</u>	67	69	38
1.5	1.0	0.5	0.0	<u>16</u>	11	12	<u>16</u>
3.0	2.0	1.0	0.0	<u>49</u>	37	35	38

Underline denotes largest value in each row.

Table 6.2

Power (%, $p = 8$, $\nu = 1$, size=5%)

δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	ONE	TWO	MAX	F
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5	5	5	5
2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	13	11	<u>14</u>	8
4.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	49	44	<u>62</u>	12
2.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	<u>22</u>	18	18	12
4.0	4.0	0.0	0.0	0.0	0.0	0.0	0.0	<u>82</u>	76	75	24
1.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	<u>10</u>	8	8	<u>10</u>
2.0	2.0	2.0	0.0	0.0	0.0	0.0	0.0	<u>29</u>	23	19	17
4.0	4.0	4.0	0.0	0.0	0.0	0.0	0.0	<u>93</u>	89	75	41
1.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	12	9	8	<u>14</u>
2.0	2.0	2.0	2.0	2.0	0.0	0.0	0.0	<u>37</u>	26	18	32
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	12	6	7	<u>24</u>
4.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	57	50	<u>58</u>	17
4.0	3.0	2.0	1.0	0.0	0.0	0.0	0.0	<u>69</u>	60	53	32

Underline denotes largest value in each row.

Appendix A. Examples of the limiting distributions.

The limiting characteristic functions $\phi(s,t)$ in (3.17) for $p=2,3$ are presented as follows. ($\theta=(1-2is)^{-\frac{1}{2}}$, $\varphi=(1-2it)^{-\frac{1}{2}}$)

$$
\begin{aligned}
p=&2:\\
&\{\frac{1}{2}-\frac{1}{2\sqrt{2}}\sqrt{\frac{\nu}{\nu+1}}\}\theta^3+\frac{1}{2\sqrt{2}}\sqrt{\frac{\nu+2}{\nu+1}}\theta^2\varphi\\
&+\frac{1}{2\sqrt{2}}\sqrt{\frac{\nu}{\nu+1}}\theta\varphi^2+\{\frac{1}{2}-\frac{1}{2\sqrt{2}}\sqrt{\frac{\nu+2}{\nu+1}}\}\varphi^3\\
p=&3:\\
&[\frac{1}{2\pi}\{\sin^{-1}\sqrt{\frac{\nu+3}{3(\nu+1)}}+\sin^{-1}\sqrt{\frac{2(\nu+3)}{3(\nu+2)}}\}-\frac{1}{\sqrt{2}\pi}\frac{\sqrt{\nu(\nu+3)}}{\nu+1}]\theta^6\\
&+\{\frac{1}{2\sqrt{2}}\sqrt{\frac{\nu+1}{\nu+2}}-\frac{1}{4}\frac{\nu-1}{\nu+1}\}\theta^5\varphi+\frac{1}{\sqrt{2}\pi}\frac{\sqrt{\nu(\nu+3)^3}}{(\nu+1)(\nu+2)}\theta^4\varphi^2\\
&+\{-\frac{1}{\sqrt{2}}\frac{\nu+3/2}{\sqrt{(\nu+1)(\nu+2)}}+\frac{\nu^2+3\nu+3/2}{(\nu+1)(\nu+2)}\}\theta^3\varphi^3\\
&+\frac{1}{\sqrt{2}\pi}\frac{\sqrt{\nu^3(\nu+3)}}{(\nu+1)(\nu+2)}\theta^2\varphi^4+\{\frac{1}{2\sqrt{2}}\sqrt{\frac{\nu+2}{\nu+1}}-\frac{1}{4}\frac{\nu+4}{\nu+2}\}\theta\varphi^5\\
&+[\frac{1}{2\pi}\{\sin^{-1}\sqrt{\frac{\nu}{3(\nu+2)}}+\sin^{-1}\sqrt{\frac{2\nu}{3(\nu+1)}}\}-\frac{1}{\sqrt{2}\pi}\frac{\sqrt{\nu(\nu+3)}}{\nu+2}]\varphi^6
\end{aligned}
$$

By letting $\nu\to\infty$ they agree with the formulae given in Appendix A of Chapter II.

Appendix B. A sufficient condition for the FKG condition.

We summarize the FKG inequality and prove a sufficient condition for the FKG condition. Let φ be a density on $X\in\mathbf{R}^n$ with respect to the measure μ, where $X=\prod_{i=1}^n X_i$ with X_i an interval of $\mathbf{R}^1$, and $\mu=\prod_{i=1}^n\mu_i$ with μ_i a σ-finite measure on X_i. For two points $\boldsymbol{x}=(x_1,\dots,x_n)$, $\boldsymbol{y}=(y_1,\dots,y_n)$ in X, write $\boldsymbol{x}\le\boldsymbol{y}$ iff $x_i\le y_i$, $1\le i\le n$, and write $\boldsymbol{x}\wedge\boldsymbol{y}=(v_1,\dots,v_n)$, $\boldsymbol{x}\vee\boldsymbol{y}=(w_1,\dots,w_n)$ with $v_i=x_i\wedge y_i$, $w_i=x_i\vee y_i$. We call that the density φ satisfies the *FKG condition* (or *MTP_2 property*) when

$$\varphi(\boldsymbol{x})\varphi(\boldsymbol{y})\le\varphi(\boldsymbol{x}\wedge\boldsymbol{y})\varphi(\boldsymbol{x}\vee\boldsymbol{y})\quad\text{for all }\boldsymbol{x},\boldsymbol{y}\in X. \tag{B.1}$$

The density which satisfies the FKG condition is known to satisfy the following inequality.

Theorem B.1 Suppose that the density φ satisfies the FKG condition (B.1) and that the functions g and h on X are nondecreasing in each argument on $\{\varphi > 0\}$, i.e. for $\boldsymbol{x}, \boldsymbol{y} \in \{\varphi > 0\}$ $\boldsymbol{x} \leq \boldsymbol{y}$ implies $g(\boldsymbol{x}) \leq g(\boldsymbol{y})$ and $h(\boldsymbol{x}) \leq h(\boldsymbol{y})$. Then the *FKG inequality*

$$\int_X gh\varphi d\mu \geq \Big(\int_X g\varphi d\mu\Big)\Big(\int_X h\varphi d\mu\Big) \tag{B.2}$$

holds provided that the integrals exist. In addition, suppose that g is not constant on $\{\varphi > 0\}$ and that h is strictly increasing on $\{\varphi > 0\}$, i.e. for $\boldsymbol{x}, \boldsymbol{y} \in \{\varphi > 0\}$ $\boldsymbol{x} \leq \boldsymbol{y}$ and $\boldsymbol{x} \neq \boldsymbol{y}$ implies $h(\boldsymbol{x}) < h(\boldsymbol{y})$, then the inequality in (B.2) holds strictly.

(Proof) See, for example, Perlman and Olkin (1980, Proposition 2.4 and Remark 2.5). □

When we apply the FKG inequality, the FKG condition should be verified. One useful sufficient condition for (B.1) is that φ on X is positive and has second derivatives, and that

$$\frac{\partial^2}{\partial x_i \partial x_j} \log \varphi(x_1, \ldots, x_n) \geq 0 \quad \text{for all } i < j, \tag{B.3}$$

see Kemperman (1977, p. 329, Remark 1). The condition $\varphi > 0$ is, however, too strict to be satisfied in many applications. We therefore present a weaker sufficient condition for the FKG condition as follows.

Lemma B.1 The density φ on X satisfies the FKG condition (B.1) provided that

(a) $\{\boldsymbol{x} \in X \mid \varphi(\boldsymbol{x}) > 0\} = X \cap D$ where

$$D = \bigcap_{\alpha \in A} \{\boldsymbol{x} \mid f_\alpha(x_{i_\alpha}) - g_\alpha(x_{j_\alpha}) > 0\} \tag{B.4}$$

for some nondecreasing functions f_α, g_α on X_{i_α}, X_{j_α}, $i_\alpha < j_\alpha$, respectively, and the set of indices A;

and

(b) $\varphi(\boldsymbol{x})$ has second derivatives on $\{\varphi > 0\}$ and satisfies (B.3) for $\boldsymbol{x} \in \{\varphi > 0\}$.

(Proof) If $\varphi(\boldsymbol{x}) = 0$ or $\varphi(\boldsymbol{y}) = 0$ then (B.1) holds trivially. Let $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, \ldots, y_n)$ be fixed in $\{\varphi > 0\}$. Let further $v_i = x_i \wedge y_i$ and $w_i = x_i \vee y_i$. Without loss of generality we permute the coordinates such that

$$\boldsymbol{x} = (w_1, \ldots, w_r, v_{r+1}, \ldots, v_n), \quad \boldsymbol{y} = (v_1, \ldots, v_r, w_{r+1}, \ldots, w_n)$$

where $r = \sharp\{x_i \geq y_i\}$, and that

$$\forall i < j, \quad \exists i', j' \quad \text{s.t.} \quad v_i = x_{i'}, v_j = x_{j'} \quad \Rightarrow \quad i' < j',$$

and

$$\forall i < j, \quad \exists i', j' \quad \text{s.t.} \quad w_i = x_{i'}, w_j = x_{j'} \quad \Rightarrow \quad i' < j'.$$

On the other hand, this permutation of coordinates changes the expression of (B.4) as $D = D_1 \cap D_2 \cap D_3 \cap D_4$ with

$$D_k = \bigcap_{\alpha \in A_k} \{\boldsymbol{x} \mid f_\alpha(x_{i_\alpha}) - g_\alpha(x_{j_\alpha}) > 0\}, \quad 1 \le k \le 4,$$

where

$$A_1 = \{\alpha \in A \mid i_\alpha < j_\alpha \le r\}, \quad A_2 = \{\alpha \in A \mid r+1 \le i_\alpha < j_\alpha\},$$

$$A_3 = \{\alpha \in A \mid i_\alpha \le r, r+1 \le j_\alpha\}, \quad A_4 = \{\alpha \in A \mid j_\alpha \le r, r+1 \le i_\alpha\}.$$

Define the sequence with double suffix $\{\boldsymbol{x}_{ij} \in \mathbf{R}^n\}$ by

$$\begin{aligned} \boldsymbol{x}_{ij} =& (w_1, \ldots, w_i, v_{i+1}, \ldots, v_r, w_{r+1}, \ldots, w_{r+j}, v_{r+j+1}, \ldots, v_n), \\ & i = 0, \ldots, r, \quad j = 0, \ldots, s, \end{aligned}$$

with $s = n - r$. Note that $\boldsymbol{x} = \boldsymbol{x}_{r0}$, $\boldsymbol{y} = \boldsymbol{x}_{0s}$, $\boldsymbol{x} \wedge \boldsymbol{y} = \boldsymbol{x}_{00}$, $\boldsymbol{x} \vee \boldsymbol{y} = \boldsymbol{x}_{rs}$. Define the closed rectangle S_{ij} which has four vertices $\{\boldsymbol{x}_{i-1,j-1}, \boldsymbol{x}_{i,j-1}, \boldsymbol{x}_{i,j}, \boldsymbol{x}_{i-1,j}\}$. Suppose the statement that

$$\boldsymbol{x}, \boldsymbol{y} \in D \quad \Rightarrow \quad S_{ij} \subset D \quad \text{for all} \quad 1 \le i \le r, 1 \le j \le s, \tag{B.5}$$

which shall be proved later. Then if φ satisfies (b), by integrating (B.3) with respect to x_i and x_j over S_{ij}, we get

$$0 < \varphi(\boldsymbol{x}_{i,j-1})\varphi(\boldsymbol{x}_{i-1,j}) \le \varphi(\boldsymbol{x}_{i-1,j-1})\varphi(\boldsymbol{x}_{i,j}).$$

Therefore from the argument of Kemperman (1977, p.329, Assertion (i)) or Karlin and Rinott (1980, Proposition 2.1), we see that the FKG condition (B.1) holds.

Finally we prove (B.5). We only have to prove that if $\boldsymbol{x}, \boldsymbol{y} \in D$ then $\boldsymbol{x}^* \in D$ with

$$\begin{aligned} \boldsymbol{x}^* = & (w_1, \ldots, w_{i-1}, v_i + s(w_i - v_i), v_{i+1}, \ldots, v_r, \\ & w_{r+1}, \ldots, w_{r+j-1}, v_{r+j} + t(w_{r+j} - v_{r+j}), v_{r+j+1}, \ldots, v_n), \\ & \text{for } 0 \le s \le 1, \quad 0 \le t \le 1. \end{aligned}$$

The assumption $\boldsymbol{x}, \boldsymbol{y} \in D_1$ means

$$f_\alpha(v_i) - g_\alpha(v_j) > 0, \quad f_\alpha(w_i) - g_\alpha(w_j) > 0, \quad i < j \le r,$$

which implies $\boldsymbol{x}^* \in D_1$ because

$$f_\alpha(w_i) - g_\alpha(v_j + s(w_j - v_j)) \ge f_\alpha(w_i) - g_\alpha(w_j) > 0,$$

and

$$f_\alpha(v_i + s(w_i - v_i)) - g_\alpha(v_j) \geq f_\alpha(v_i) - g_\alpha(v_j) > 0$$

for all $0 \leq s \leq 1$. The assumption $\boldsymbol{x}, \boldsymbol{y} \in D_2$ also implies $\boldsymbol{x}^* \in D_2$. The assumption $\boldsymbol{y} \in D_3$ means

$$f_\alpha(v_i) - g_\alpha(w_j) > 0, \quad i \leq r, \quad r + 1 \leq j,$$

which implies $\boldsymbol{x}^* \in D_3$ because

$$f_\alpha(v_i + s(w_i - v_i)) - g_\alpha(v_j + t(w_j - v_j)) \geq f_\alpha(v_i) - g_\alpha(w_j) > 0$$

for all $0 \leq s, t \leq 1$. The assumption $\boldsymbol{x} \in D_4$ also implies $\boldsymbol{x}^* \in D_4$. Therefore we prove $\boldsymbol{x}^* \in D$ and (B.5) is established. The proof is completed. □

Remark B.1 If the density φ on X which satisfies (a) of Lemma B.1 is TP_2 (Karlin, 1968) in each pair of the coordinates, φ can be proved to satisfy the FKG condition by the same way.

References

Amemiya, Y., Anderson, T. W. and Lewis, P. A. W. (1990). Percentage points for a test of rank in multivariate components of variance. *Biometrika* **77**, 637-641.

Anderson, B. M., Anderson, T. W. and Olkin, I. (1986). Maximum likelihood estimators and likelihood ratio criteria in multivariate components of variance. *Ann. Statist.* **14**, 405-417.

Anderson, T. W. (1984a). Estimating linear statistical relationships. *Ann. Statist.* **12**, 1-45.

Anderson, T. W. (1984b). *An Introduction to Multivariate Statistical Analysis*, 2nd ed. John-Wiley & Sons, New York.

Anderson, T. W. (1989). The asymptotic distribution of the likelihood ratio criterion for testing rank in multivariate components of variance. *J. Multivariate Anal.* **30**, 72-79.

Anderson, T. W. and Amemiya, Y. (1991). Testing dimensionality in the multivariate analysis of variance. *Statist. Probab. Lett.* **12**, 445-463.

Anderson, T. W. and Das Gupta, S. (1964). A monotonicity property of the power functions of some tests of the equality of two covariance matrices. *Ann. Math. Statist.* **35**, 1059-1063.

Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory*, Expanded ed. Holden-Day, San Francisco.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*, 2nd ed. Springer, New York.

Calvin, J. A. and Dykstra, R. L. (1991). Maximum likelihood estimation of a set of covariance matrices under Löwner order restrictions with applications to balanced multivariate variance components models. *Ann. Statist.* **19**, 850-869.

Chernoff, H. (1954). On the distribution of the likelihood ratio. *Ann. Math. Statist.* **25**, 573-578.

Giri, N. (1968). On tests of the equality of two covariance matrices. *Ann. Math. Statist.* **39**, 275-277.

Crowder, M. J. and Hand, D. J. (1990). *Analysis of Repeated Measures.* Chapman and Hall, London.

Johansen, S. (1984). *Functional Relations, Random Coefficients, and Nonlinear Regression with Application to Kinetic Data.* Lecture Notes in Statistics **22**, Springer, New York.

Karlin, S. (1968). *Total Positivity.* Stanford University Press, Stanford.

Karlin, S. and Rinott, Y. (1980). Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. *J. Multivariate Anal.* **10**, 467-498.

Karlin, S. and Rinott, Y. (1988). A generalized Cauchy-Binet formula and applications to total positivity and majorization. *J. Multivariate Anal.* **27**, 284-299.

Kemperman, J. H. B. (1977). On the FKG-inequality for measures on a partially ordered space. *Indag. Math.* **39**, 313-331.

Klotz, J. and Putter, J. (1969). Maximum likelihood estimation of multivariate covariance components for the balanced one-way layout. *Ann. Math. Statist.* **40**, 1100-1105.

Krishnaiah, P. R. (1976). Some recent developments on complex multivariate distributions. *J. Multivariate Anal.* **6**, 1-30.

Krishnaiah, P. R. and Chang, T. C. (1971). On the exact distributions of the extreme roots of the Wishart and MANOVA matrices. *J. Multivariate Anal.* **1**, 108-117.

Kuriki, S. (1990). One-sided test for the equality of two covariance matrices. *METR* 90-15, Department of Mathematical Engineering and Information Physics, University of Tokyo. To appear in *Ann. Statist.*

Kuriki, S. (1991). Likelihood ratio tests for the multivariate random effects models. *METR* 91-09, Department of Mathematical Engineering and Information Physics, University of Tokyo. To appear in *J. Multivariate Anal.*

Kuriki, S. (1992). Limiting null distributions of one-sided likelihood ratio criteria for testing dimensionality. *METR* 92-09, Department of Mathematical Engineering and Information Physics, University of Tokyo. Submitted for publication.

Lange, N. and Laird, N. M. (1989). The effect of covariance structure on variance estimation in balanced growth-curve models with random parameters. *J. Amer. Statist. Assoc.* **84**, 241-247.

Lehmann, E. L. (1983). *Theory of Point Estimation.* John-Wiley & Sons, New York.

Lehmann, E. L. (1986). *Testing Statistical Hypotheses*, 2nd ed. John-Wiley & Sons, New York.

Macdonald, I. G. (1979). *Symmetric Functions and Hall Polynomials.* Oxford Mathematical Monographs, Clarendon Press, Oxford.

McMullen, P. (1975). Non-linear angle-sum relations for polyhedral cones and polytopes. *Math. Proc. Camb. Phil. Soc.* **78**, 247-261.

Mehta, M. L. (1960). On the statistical properties of the level-spacings in nuclear spectra. *Nuclear Physics* **18**, 395-419.

Mehta, M. L. (1989). *Matrix Theory: Selected Topics and Useful Results*. Les Editions de Physique, Les Ulis Cedex (France).

Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory.* John-Wiley & Sons, New York.

Nagao, H. (1973). On some test criteria for covariance matrix. *Ann. Statist.* **1**, 700-709.

Perlman, M. D. (1980). Unbiasedness of the likelihood ratio tests for equality of several covariance matrices and equality of several multivariate normal populations. *Ann. Statist.* **8**, 247-263.

Perlman, M. D. and Olkin, I. (1980). Unbiasedness of invariant tests for MANOVA and other multivariate problems. *Ann. Statist.* **8**, 1326-1341.

Pillai, K. C. S. (1956). Some results useful in multivariate analysis. *Ann. Math. Statist.* **27**, 1106-1114.

Rao, C. R. (1965). The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. *Biometrika* **52**, 447-458.

Rao, C. R. and Kleffe, J. (1988). *Estimation of Variance Components and Applications*. North-Holland, Amsterdam.

Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Order Restricted Statistical Inference.* John-Wiley & Sons, Chichester.

Sakata, T. (1987). Likelihood ratio test for one-sided hypothesis of covariance matrices of two normal populations. *Commun. Statist.* **A16**, 3157-3168.

Scheffé, H. (1959). *The Analysis of Variance.* John-Wiley & Sons, New York.

Schott, J. R. and Saw, J. G. (1984). A multivariate one-way classification model with random effects. *J. Multivariate Anal.* **15**, 1-12.

Searle, S. R., Casella, G. and McCulloch, C. E. (1992). *Variance Components.* John-Wiley & Sons, New York.

Self, S. G. and Liang, K-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *J. Amer. Statist. Assoc.* **82**, 605-610.

Shapiro, A. (1985). Asymptotic distribution of test statistics in the analysis of moment structures under inequality constraints. *Biometrika* **72**, 133-144.

Shapiro, A. (1987). A conjecture related to chi-bar-squared distributions. *Amer. Math. Monthly* **94**, 46-48.

Shapiro, A. (1988). Towards a unified theory of inequality constrained testing in multivariate analysis. *Internat. Statist. Rev.* **56**, 49-62.

Sugiura, N. and Nagao, H. (1968). Unbiasedness of some test criteria for the equality of one or two covariance matrices. *Ann. Math. Statist.* **39**, 1686-1692.

Wynn, H. P. (1975). Integrals for one-sided confidence bounds: A general result. *Biometrika* **62**, 393-396.

論 文 の 内 容 の 要 旨

論文題目　Likelihood Ratio Tests in Multivariate Variance Components Models
（多変量分散成分模型に関する尤度比検定の研究）

氏　　名　栗木　哲

因子（効果）を確率変数で記述する分散分析模型は変量模型，あるいは分散成分模型と呼ばれる．このモデルは長い歴史を持ち，理論，応用の両面にわたって莫大な数の研究がなされている．ところがこれらの分散成分模型（特に多変量分散成分模型）における最尤推定や尤度比検定は，その推定量や検定統計量が場合分けを含む複雑な関数となるという困難のために，現在でもそれらの導出方法が研究対象とされている段階であり，統計的性質に関する研究はほとんどなされていない．本論文は，３つの典型的な多変量分散成分模型に関する尤度比検定について，有意点計算の際に必要な検定統計量の帰無仮説の下での漸近分布を導き，また不偏性，検出力の単調性などのいくつかの性質を調べたものである．

分散成分模型は，不等式制約で制限された空間がモデルの下で許容される母数空間になる，という著しい特徴を持つ．そのため効果行列（多変量分散成分の分散行列）の最尤推定量は固有根の関数の符号による場合分けを含

んだ複雑な形で与えられる．同様の困難は効果行列の尤度比検定の際にも生じる．Chernoff (1954) は，制約された母数空間の下での尤度比検定統計量の漸近帰無分布に関する一般的な議論を展開した．Chernoff の理論は我々の扱う効果行列に関する仮説検定問題に対して適用可能であることが示されるので，Shapiro (1985, 1988) などの結果と併せて尤度比検定統計量の帰無仮説の下での漸近分布はカイ・バー・二乗分布と呼ばれるカイ二乗分布の混合分布であることが分かる．しかしその混合確率は，個々のモデルで個別に求められなければならない．それらの導出が本論文の主要な目的の一つである．

本論文のもう一つの目的は，多変量分散成分の尤度比検定の検出力関数のいくつかの性質の証明である．棄却域（あるいは同じことであるが検定関数）が，単調性を持つ場合は Anderson & Das Gupta (1964) の議論を用いて検出力関数の単調性，不偏性がほとんど自明に証明される．そうでない場合は個別の議論がなされなければならない．

論文の第Ⅰ章では，多変量分散成分模型に関する簡単な説明が与えられる．また第Ⅱ章で扱うモデルを例題として，効果行列のランクに関する片側尤度比検定問題に Chernoff の理論が適用可能であること，その結果我々の求めようとする尤度比検定統計量の帰無仮説の下での漸近分布がカイ二乗分布の混合分布であることが示される．

第Ⅱ章では，２つの分散行列の同等性などに関する片側検定を扱う．このモデルは，多変量分散成分模型の典型例である多変量一元配置変量模型において，群内，群間平方和行列を考えることに相当する．尤度比検定統計量はKlotz & Putter (1969), Anderson, et al.(1986) が与えている．最初にいくつかの検定の検出力関数の単調性や不偏性が，その検定関数の単調性から証明される．次に，尤度比検定統計量の帰無仮説の下での漸近分布がカイ二乗分布の混合分布であることが（Chernoff の理論を用いずに）直接示される．さらに，混合確率が Pillai (1954) と同様の方法で計算できることを導き，実際に分布の数表を与える．また漸近分布の特性関数を Pfaffian を用いて陽に書き下す方法を与えるが，この方法は混合確率計算の別解法を与えている．これらの結果は次元が２の場合の分布を導いた Sakata (1987), Anderson (1989) を含む一般的な結果である．以上述べた極限分布の表現の他，分布の漸近展開公式も与えられる．

第Ⅲ章では，複素正規母集団に関する２つの分散行列の同等性などに関する片側尤度比検定を扱う．この問題は第Ⅱ章のモデルを複素母集団の場合に書き直した問題で，第Ⅱ章の結果の多くが同様に成立する．ただし，尤度比検定統計量の帰無仮説の下での漸近分布の特性関数は（Pfaffian でなく）行列式を用いて陽に書かれることが示される．

第Ⅳ章では，いわゆる Random coefficient regression model の分散構造に関する尤度比検定問題を扱う．このモデルは Scheffe の混合モデルを特殊な場合として含んでいる．尤度比検定統計量は基本的には Anderson, et al.(1986) が与えている．最初に効果行列が零行列であるという仮説に関する尤度比検定の不偏性を証明するが，検定関数が単調性を持たないので，その証明は第Ⅱ，Ⅲ章のモデルほど自明ではない．ここでは，Sugiura & Nagao (1968) の方法を部分的に援用して証明を得る．また同じ検定問題に対し，尤度比検定を含むより広いクラスの検定方式に対する局所不偏性を FKG 不等式を用いて示す．そのために，FKG 不等式を適用するための新たな一つの十分条件を準備する．尤度比検定統計量の帰無仮説の下での漸近分布はやはりカイ二乗分布の混合分布である．次元が２，３の混合確率が与えられている．

尤度比検定の特徴付けを行うため，第Ⅱ，Ⅳ章のモデルにおいて，効果行列が零行列である，という仮説の検定問題に対してモンテカルロシミュレーションが行われる．その結果，効果行列のランクが小さい場合には，尤度比検定が他の検定方式に比べて高い検出力を持つことが示される．