INVARIANT TESTS AND LIKELIHOOD RATIO TESTS
FOR MULTIVARIATE ELLIPTICALLY CONTOURED DISTRIBUTIONS

TECHNICAL REPORT NO. 14

HUANG HSU

MAY 1985

U. S. ARMY RESEARCH OFFICE
CONTRACT DAAG29-82-K-0156
THEODORE W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.
INVARIANT TESTS AND LIKELIHOOD RATIO TESTS
FOR MULTIVARIATE ELLIPTICALLY CONTOURED DISTRIBUTIONS

Technical Report No. 14

Huang Hsu
Stanford University

May 1985

U. S. Army Research Office
Contract DAAG29-82-K-0156

Theodore W. Anderson, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.
THE VIEW, OPINIONS, AND/OR FINDINGS CONTAINED IN THIS REPORT ARE THOSE OF THE AUTHOR(S) AND SHOULD NOT BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, POLICY, OR DECISION, UNLESS SO DESIGNATED BY OTHER DOCUMENTATION.
ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to my advisor Professor Theodore W. Anderson for his constant guidance, continual support and encouragement that have made this report possible during the period of my studies at Stanford. I would also like to thank Professors Charles Stein and Ingram Olkin for reading this manuscript and making many helpful suggestions. The "Chinese Mafia" have been important to me. Thanks are due to Drs. Chen-Hsin Chen, Inchi Hu, Young-Lim Su, Bwo-Han Yung, Sweetson Chen, Chung-Teh Wu, Lian-Tai Chen, and Gary Chang. I am grateful to my parents and my fiancée, Men-Li, for their moral support and care they have given me all through the years. I dedicate this report to them.
Contents

Chapter 1. Introduction ........................................... 1

Chapter 2. General results for multivariate elliptically contoured distributions .... 6
  2.1 The distributions of certain specified statistics .................. 6
  2.2 Power functions of specified tests ................................ 8

Chapter 3. Multivariate regression analysis ................................ 11
  3.1 Invariant tests and null distributions .............................. 11
  3.2 Admissibility of invariant tests .................................... 15

Chapter 4. Testing hypothesis for mean vector ............................ 19
  4.1 Invariant test, null distribution, and admissibility ............... 19
  4.2 Locally most powerful invariant test ............................... 20

Chapter 5. Testing various other hypotheses ............................. 32
  5.1 Criteria for testing lack of correlations among the sets .......... 32
  5.2 Testing equality of covariance matrices and mean vectors .......... 35
  5.3 Test of hypothesis about correlation coefficient .................. 38
  5.4 Test of hypothesis about partial correlation coefficient ........... 40
  5.5 Test of hypothesis about the multiple correlation coefficient ...... 43

Chapter 6. Maximum likelihood estimates and likelihood ratio tests ........ 44
  6.1 The result for a special case ...................................... 44
  6.2 Examples ..................................................... 50
  6.3 General theorem ............................................... 58

Chapter 7. Noncentral distributions of quadratic forms .................. 62
Abstract

The usual assumption in multivariate hypothesis testing is that the sample consists of \( n \) independent, identically distributed Gaussian \( p \)-vectors. In this dissertation this assumption is weakened by considering a class of distributions for which the vector observation are not necessarily Gaussian or independent. This class consists of the elliptically symmetric laws with densities of the form \( f(X_{n	imes p}) = g(\text{tr}(X - M)'(X - M)\Lambda^{-1}) \). The following hypothesis testing problems are considered: testing for equality between the mean vector and a specified vector, lack of correlations among different sets, equality of covariance matrices and mean vectors, equality between the correlation coefficient and a specified number, and MANOVA. For each of the above hypotheses, invariant tests and their properties are developed. These include the uniformly most powerful test, the locally most powerful test, admissibility, and null and non-null distributions. Further, under the assumptions that \( g(\cdot) \) is continuous, each element of the covariance matrix of \( X \) is finite, and the null hypothesis is scalar-invariant, it is shown that the usual normal-theory likelihood ratio tests are exactly robust for the null case under this wider class (i.e. the likelihood ratio tests, sampling from this general class, are the same as the usual normal-theory likelihood ratio tests and their null distributions are the same.)
Chapter 1

INTRODUCTION

The usual assumption in multivariate hypothesis testing is that the sample consists of independent and identically distributed Gaussian vectors. In this paper we consider the more general hypothesis testing problem when the sample observations, not necessarily either Gaussian or independent, are from the family of multivariate elliptically contoured distributions.

The elliptically contoured distributions on the \( n \)-dimensional Euclidean space \( \mathbb{R}^n \) are defined as follows. If the characteristic function of an \( n \)-dimensional random vector \( \mathbf{x} \) has the form \( \exp(it'\mu)\phi(t'S^*t) \), where \( \mu : n \times 1 \), \( S^* : n \times n \), \( \text{rank}(S^*) = k \), \( S^* \succeq 0 \), and \( \phi \in \Phi_k = \{ \phi \mid \phi(\cdot) \text{ is a function such that } \phi(t_1^2 + \cdots + t_k^2) \text{ is a characteristic function on } \mathbb{R}^k \} \), we say that \( \mathbf{x} \) is distributed according to an elliptically contoured distribution with parameters \( \mu, S^*, \) and \( \phi \), and write \( \mathbf{x} \sim EC_n(\mu, S^*, \phi) \).

Elliptically contoured distributions have been extended to the case of matrices by Dawid (1977, 1978), Chmielewski (1980), and Anderson and Fang (1982b).

Let \( \mathbf{X}, \mathbf{M}, \) and \( \mathbf{T} \) be \( n \times p \) matrices. We express them in terms of elements, columns, and rows as
Chapter 1: Introduction

\[ \mathbf{X} = (z_{ij}) = (x_1, \ldots, x_p) = \begin{pmatrix} x_{(1)}' \\ \vdots \\ x_{(n)}' \end{pmatrix}, \quad x = \text{vec } \mathbf{X}', \]

\[ \mathbf{M} = (\mu_{ij}) = (\mu_1, \ldots, \mu_p) = \begin{pmatrix} \mu_{(1)}' \\ \vdots \\ \mu_{(n)}' \end{pmatrix}, \quad \mu = \text{vec } \mathbf{M}', \quad (1.1) \]

\[ \mathbf{T} = (t_{ij}) = (t_1, \ldots, t_p) = \begin{pmatrix} t_{(1)}' \\ \vdots \\ t_{(n)}' \end{pmatrix}, \quad t = \text{vec } \mathbf{T}', \]

where \( x = \text{vec } \mathbf{X}' \overset{\text{def}}{=} \left( x_{(1)}', \ldots, x_{(n)}' \right)' \) with the corresponding meanings for \( \mu \) and \( t \).

If the characteristic function of a random matrix \( \mathbf{X} \) has the form

\[ \exp \left( i \sum_{j=1}^{n} t_{(j)}' \mu(j) \right) \phi \left( t_{(1)}' \Sigma_1 t_{(1)}, \ldots, t_{(n)}' \Sigma_n t_{(n)} \right), \quad (1.2) \]

with \( \Sigma_1, \ldots, \Sigma_n \geq 0 \), we say that \( \mathbf{X} \) is distributed according to a multivariate (rows) elliptically contoured distribution (MECD) and write \( \mathbf{X} \sim \text{MECD}_{n \times p} (\mathbf{M}; \Sigma_1, \ldots, \Sigma_n; \phi) \). In this paper we consider only the subclass of MECD in which the function \( \phi \) satisfies

\[ \phi(t_1, \ldots, t_n) = \varphi(t_1 + \cdots + t_n). \quad (1.3) \]

We continue to denote MECD in this subclass by \( \text{MECD}_{n \times p} (\mathbf{M}; \Sigma_1, \ldots, \Sigma_n; \phi) \) and \( \varphi \) by \( \phi \).

When \( \mu_{(1)} = \mu_{(2)} = \cdots = \mu_{(n)} = \mu \) and \( \Sigma_1 = \cdots = \Sigma_n = \Sigma \), we write \( \mathbf{X} \sim \text{LEC}_{n \times p} (\mu, \Sigma, \phi) \), and if the density of \( \mathbf{X} \) exists (in this paper, this is the usual assumption), then it has the form

\[ |\Sigma|^{-\frac{1}{2}} n g \left( \text{tr } \Sigma^{-1} \mathbf{G} \right), \quad (1.4) \]
where \( G = \sum_{j=1}^{n} (x(j) - \mu)(x(j) - \mu)' \).

Let \( u^{(k)} \) denote a random vector which is uniformly distributed on the unit sphere in \( \mathbb{R}^q \) and \( \Omega_q (\|t\|^2) \) denote its characteristic function. Schoenberg (1933) pointed out that a characteristic function \( \phi \in \Phi_q \) if and only if

\[
\phi(t) = \int_0^\infty \Omega_q (tr^2) \, dF(r)
\]

for some distribution function \( F \) on \([0, \infty)\).

Throughout \( X \triangleq Y \) means that the random matrices \( X \) and \( Y \) are identically distributed; \( \text{rank}(A) \) denotes the rank of the matrix \( A \).

Cambanis, Huang and Simons (1981) obtained the following two properties:

(i) \( x \sim E_{C_n} (\mu, \Sigma^*, \phi) \) with \( \text{rank}(\Sigma^*) = k \) if and only if

\[
x \triangleq \mu + RD'u^{(k)},
\]

where \( R \geq 0 \) is independent of \( u^{(k)} \), \( \Sigma^* = D'D \) is a factorization of \( \Sigma^* \) (i.e., \( D \) is a \( k \times n \) matrix and \( \text{rank}(D) = k \)) and the distribution function \( F \) of \( R \) is related to \( \phi \) as in (1.5) with \( k \) substituted for \( q \). For convenience we denote this relationship by \( R \leftrightarrow \phi \in \Phi_k \). Here \( x \triangleq y \) denotes that the random vectors \( x \) and \( y \) are identically distributed.

(ii) Write \( u^{(n)} = (u_1^{(n)}', u_2^{(n)}')' \), where \( u_1^{(n)} \) is \( m \)-dimensional column vector \((1 \leq m < n)\). Then \( (u_1^{(n)}, u_2^{(m)}) \triangleq (R_{mn}u^{(m)}, (1 - R_{mn}^2)^{1/2}u^{(n-m)}) \), where \( R_{mn} \geq 0 \), \( u^{(m)} \), and \( u^{(n-m)} \) are independent, and \( R_{mn}^2 \sim \text{Beta} \left( \frac{m}{2}, \frac{n-m}{2} \right) \).

Anderson and Fang (1982b) pointed out that \( X \sim MEC_{nxp} (M; \Sigma_1, \ldots, \Sigma_n; \phi) \) if and only if \( x = \text{vec} X' \sim E_{C_{np}} (\mu, V, \phi) \) where \( V \) is a matrix with diagonal blocks \( \Sigma_i \), \( i = 1, \ldots, n \), and off-diagonal blocks \( 0 \), and \( \mu = \text{vec} M' \). So it is easy to see that if \( X \sim LEC_{nxp} (\mu, \Sigma, \phi) \) and \( \text{rank}(\Sigma) = \ell \) then

\[
X \triangleq 1_n \mu' + RU \ell D',
\]

where \( 1_n' = (1, 1, \ldots, 1) \), \( U : n \times \ell, \text{vec} U = u^{(n\ell)} \), \( D : \ell \times p, D'D = \Sigma \), and \( R \leftrightarrow \phi \in \Phi_{n\ell} \).
When $\Sigma > 0$, $X \sim \text{LEC}_{n \times p}(\mu, \Sigma, \phi)$ with $\phi(t) = \exp(-\frac{1}{2}t)$ if and only if $x_{(j)}$, $j = 1, \ldots, n$, are i.i.d. $N_p(\mu, \Sigma)$. (This implies that sampling from a multivariate normal distribution is a special case of sampling from the family of multivariate elliptically contoured distributions.)

There are three definitions to be emphasized here:

(i) The Dirichlet Distribution $D_m(p_1, \cdots, p_{m-1}; p_m)$.

If $y = (y_1, \cdots, y_m)'$ is a random vector with $\sum_{i=1}^{m} y_i = 1$ and $(y_1, \cdots, y_{m-1})'$ has the density with $p_i > 0$, $i = 1, \ldots, m$,

$$f_m(t_1, \cdots, t_{m-1}) = \begin{cases}
\Gamma\left(\sum_{i=1}^{m} p_i\right) \prod_{i=1}^{m-1} t_i^{p_i-1} \left(1 - \sum_{i=1}^{m-1} t_i\right)^{p_m-1}, & \text{if } t_i \geq 0, \\
0, & \text{otherwise},
\end{cases}
$$

we say that $(y_1, \cdots, y_{m-1})' \sim D_m(p_1, \cdots, p_{m-1}; p_m)$.

(ii) The distribution $MG_{p,k+1}(\Sigma; \frac{n_1}{2}, \ldots, \frac{n_k}{2}, \frac{n_{k+1}}{2}; \phi)$.

If $X = (X_1, \cdots, X_{k+1})' \sim \text{LEC}_{n \times p}(0, \Sigma, \phi)$ with $\Sigma > 0$, then we say that $(W_1, \cdots, W_m) \sim MG_{p,k+1}(\Sigma; \frac{n_1}{2}, \ldots, \frac{n_k}{2}, \frac{n_{k+1}}{2}; \phi)$, where $W_i = X_i'X_i$, $i = 1, \ldots, k$, $X_i$, $i = 1, \ldots, k+1$, is a $n_i \times p$ matrix, $p \leq n_i < n$, $i = 1, \ldots, k$, $n_{k+1} \geq 1$, and $\sum_{i=1}^{k+1} n_i = n$.

(iii) Majorization ($x < y$).

For $x, y \in \mathbb{R}^n$, $x < y$ ($y$ majorizes $x$), if

$$\begin{cases}
\sum_{i=1}^{n} x[i] \leq \sum_{i=1}^{n} y[i], & i = 1, \ldots, n-1; \\
\sum_{i=1}^{n} x[n] = \sum_{i=1}^{n} y[i],
\end{cases}
$$

where $x[i], i = 1, \ldots, n$, is the arrangement of $x_i$, $i = 1, \ldots, n$, such that $x_{[1]} \geq \cdots \geq x_{[n]}$.

In this paper invariant tests, sampling from the family of multivariate elliptically contoured distributions, are discussed. Chapter 2 develops some general results in finding the distributions and power functions of various test statistics. Chapter 3 discusses invariant
tests in multivariate regression models and their properties including admissibility. Chapter 4 studies an invariant test for equality between the mean vector and a specified vector, shows that the generalized $T^2$-test is the locally most powerful invariant test under some regularity conditions and is also the only admissible invariant test, and discusses the monotonicity of the power function of the generalized $T^2$-test. The invariant tests for testing various hypotheses and their properties are studied in Chapter 5. The following hypothesis tests are considered: lack of correlations among different sets, equality of covariance matrices and mean vectors, known correlation coefficient, known partial correlation coefficient, and zero multiple correlation coefficient.

Anderson and Fang (1982c) considered maximum likelihood estimation and the likelihood ratio test. In Chapter 6 we assume that $X$ has a continuous probability density function, each element of its covariance matrix is finite, and the null hypothesis is scalar-invariant. It is then shown that the usual normal-theory likelihood ratio tests are exactly robust for the null case under this wider class (i.e. the likelihood ratio tests, sampling from this general class, are the same as the usual normal-theory likelihood ratio tests; and their null distributions are the same.)

Anderson and Fang (1982a) derive the distribution of a quadratic form for the central case for the family of multivariate elliptically contoured distributions. Chapter 7 evaluates its non-central distribution.
Chapter 2

GENERAL RESULTS FOR MULTIVARIATE ELLIPTICALLY CONTOURED DISTRIBUTIONS

In this chapter two general results for multivariate elliptically contoured distributions, used in later chapters, are derived. The multivariate normal distribution is a special case of a multivariate elliptically contoured distribution, but the distribution of a certain class of specified functions of the multivariate normal distribution is the same as that of the multivariate elliptically contoured distribution. Based on this fact, if the test statistic is one of the above specified functions, the power of this test remains the same as in the multivariate normal case.

2.1. The distributions of certain specified statistics.

Anderson and Fang (1982b, 1982c) proved that the null distributions of the likelihood ratio statistics for testing lack of correlation between sets of variates, equality between the mean vector and a specified vector, equality of covariance matrices, equality of several means, and MANOVA in the family of multivariate elliptically contoured distributions have the same distributions as in the multivariate normal case. Chmielewski (1980) also showed that the null and non-null distributions of all invariant statistics for testing the sphericity hypothesis are the same as in the multivariate normal case. In this section a general result is given.
Chapter 2: General results for multivariate elliptically contoured distributions

Theorem 2.1.

Let \( X_{n \times p} \sim MEC_{n \times p}(M; \Sigma_1, \ldots, \Sigma_i, \Sigma_{i+1}, \ldots, \Sigma_q; \phi) \) with \( n_i \) \( \Sigma_i \)'s, \( \Sigma_i > 0 \), the rows of \( M \) are \( n_i \mu_i \)'s, \( i = 1, \ldots, q \), \( \sum_{i=1}^q n_i = n \), and the density of \( X \) exists.

(i) If \( M = 0 \) and if the vector \( f(X) \) of statistics satisfies the condition that \( f(X) = f(cX) \), for every \( c > 0 \), then the distribution of \( f(X) \) does not depend on \( \phi \).

(ii) If the vector \( f(X) \) of statistics satisfies the condition that \( f(X) = f(c(X + B)) \), where the rows of \( B \) are \( n_i b_i \)'s, for every \( c > 0 \) and \( b_i \in \mathbb{R}^p \), \( i = 1, \ldots, q \), then the distribution of \( f(X) \) does not depend on \( \phi \).

Proof.

(i) We can write
\[
x = \text{vec} X' \overset{d}{=} RD'u^{(np)},
\]
where \( u^{(np)} \) and \( R \) are independent, \( D_i' D_i = \Sigma_i > 0 \), and \( D \) is a matrix with \( n_1 D_1 \)'s, \ldots, \( n_q D_q \)'s as diagonal blocks and off-diagonal blocks of 0's.

Define \( f'(\text{vec} X') = f(X) \), then \( f'(x) = f'(cx) \) and \( f'(RD'u^{(np)}) \overset{d}{=} f'(cRD'u^{(np)}) \) for every \( c > 0 \). Since the density of \( X \) exists, \( \Pr(R = 0) = 0 \). So \( f'(RD'u^{(np)}) \overset{d}{=} f'(cRD'u^{(np)}) \), where \( f'(rD'u^{(np)}) \) is the conditional random vector at \( R = r > 0 \). Then by letting \( c = r^{-1} \), \( f'(rD'u^{(np)}) \overset{d}{=} f'(D'u^{(np)}) \). Since \( u^{(np)} \) and \( R \) are independent, \( f'(x) \) is independent of \( R \) and \( f(X) \) is independent of \( R \). So the distribution of \( f(X) \) does not depend on \( \phi \).

(ii) Let \( B = -M \) and \( Y = X - M \), then
\[
f(X) = f(c(X - M))
= f(cY).
\]
When \( c = 1 \), \( f(X) = f(Y) \). So it is true that \( f(Y) = f(cY) \), for any \( c > 0 \). Then from (i), the distribution of \( f(Y) \) does not depend on \( \phi \). So the distribution of \( f(X) \) does not depend on \( \phi \).
Corollary 2.1.

Let $X_{n \times p} \sim LEC_{n \times p}(\mu, \Sigma, \phi)$ with $\Sigma > 0$, and the density of $X$ exists.

(i) If $\mu = 0$ and if the vector $f(X)$ of statistics satisfies the condition that $f(X) = f(cX)$, for every $c > 0$, then the distribution of $f(X)$ does not depend on $\phi$.

(ii) If the vector $f(X)$ of statistics satisfies the condition that $f(X) = f(c(X + 1_n d'))$, for every $c > 0$, and $d \in \mathbb{R}^p$, then the distribution of $f(X)$ does not depend on $\phi$.

Proof.

It is clear by letting $q = 1, \Sigma_1 = \Sigma, \mu_1 = \mu$ in Theorem 2.1. 

The results of Anderson and Fang (1982b, 1982c) and Chmielewski (1980) are easily obtained from Corollary 2.1. There are a number of other examples in Chapters 3, 4, 5, and 6.

2.2. Power functions of specified tests.

The previous section consists of general results in distribution theory. In this section, we will attack the problem of power functions by using the above result.

Theorem 2.2.

(i) Assume $X_{n \times p} \sim MEC_{n \times p}(M; \Sigma_1, \ldots, \Sigma_1, \Sigma_2, \ldots, \Sigma_2, \ldots, \Sigma_q, \ldots, \Sigma_q; \phi)$ with $n_i \Sigma_i$'s, $\Sigma_i > 0$, the rows of $M$ are $n_i \mu_i$'s, $i = 1, \ldots, q$, and the density of $X$ exists. For testing $H_0 : (M, \Sigma_1, \ldots, \Sigma_q) \in \Omega_0$ vs $H_1 : (M, \Sigma_1, \ldots, \Sigma_q) \in \Omega \setminus \Omega_0$, if the hypotheses remain invariant under the group $G = \{g : X \mapsto c(X + B), \text{where the rows of } B \text{ are } n_i b_i \text{'s}, c > 0, b_i \in \mathbb{R}^p, i = 1, \ldots, q\}$, and $f(X)$ is an invariant test statistic, then the power function of this test with rejection region $f(X) \in S_1$ does not depend on $\phi$.

(ii) $X_{n \times p} \sim LEC_{n \times p}(\mu, \Sigma, \phi)$, with $\Sigma > 0$, and the density of $X$ exists. For testing $H_0 : (\mu, \Sigma) \in \Omega_0$ vs $H_1 : (\mu, \Sigma) \in \Omega \setminus \Omega_0$, if the hypotheses remain invariant under the group
Chapter 2: General results for multivariate elliptically contoured distributions

\( \{ g \mid g : X \mapsto c(X + d)^t, c > 0, d \in \mathbb{R}^p \} \), and \( f(X) \) is an invariant test statistic, then the power function of this test with rejection region \( f(X) \in S_1 \) does not depend on \( \phi \).

Proof.

(i) Since \( f(X) = f(c(X + B)) \), by Theorem 2.1, the distribution of \( f(X) \) does not depend on \( \phi \) under both the null and the non-null cases. Therefore the power function of this invariant test does not depend on \( \phi \).

(ii) This is a special case of (i). Using this, we can complete the proof. \( \square \)

Corollary 2.2.

(i) In Theorem 2.2 (i), if \( \Omega = B \times \Omega^*, \Omega_0 = B \times \Omega_0^*, \Omega \setminus \Omega_0 = B \times (\Omega^* \setminus \Omega_0^*) \), where \( B \) is the set of parameters \( \mu \), \( \Omega^* \) is the set of parameters \( \Sigma \), and \( \Omega_0^* \subset \Omega^* \), \( f(X) \) is an invariant test statistic of \( H_0 : (\Sigma_1, \ldots, \Sigma_q) \in \Omega_0^* \) vs \( H_1 : (\Sigma_1, \ldots, \Sigma_q) \in \Omega^* \setminus \Omega_0^* \), then the power function of this test with same rejection region \( S_1 \) does not depend on \( \phi \).

(ii) In Theorem 2.2 (ii), if \( \Omega = B \times \Omega^*, \Omega_0 = B \times \Omega_0^*, \Omega \setminus \Omega_0 = B \times (\Omega^* \setminus \Omega_0^*) \), where \( B \) is the set of parameters \( \mu \), \( \Omega^* \) is the set of parameters \( \Sigma \), and \( \Omega_0^* \subset \Omega^* \), \( f(X) \) is an invariant test statistic of \( H_0 : \Sigma \in \Omega_0^* \) vs \( H_1 : \Sigma \in \Omega^* \setminus \Omega_0^* \), then the power function of this test with same rejection region \( S_1 \) does not depend on \( \phi \).

Proof.

(i) It is clear that the hypotheses remain invariant under the group \( \{ g \mid g : X \mapsto X + B \} \) where the rows of \( B \) are \( n_i b_i^t, b_i \in \mathbb{R}^p, i = 1, \ldots, q \}. \) Since under the transformation \( X \mapsto cX, c > 0, (\Sigma_1, \ldots, \Sigma_q) \) does not change (only \( \phi \) changes) (Since \( \Sigma_i, i = 1, \ldots, q, \) are scale matrix. i.e., for any \( a > 0, X \sim MEC_{n \times p}(0; \Sigma_1, \ldots, \Sigma_1, \Sigma_2, \ldots, \Sigma_2, \ldots, \Sigma_q, \ldots, \Sigma_q; \phi), \) then \( aX \sim MEC_{n \times p}(0; \Sigma_1, \ldots, \Sigma_1, \Sigma_2, \ldots, \Sigma_2, \ldots, \Sigma_q, \ldots, \Sigma_q; \phi^*), \) see Cambanis, Huang and Simons (1981). ) So by Theorem 2.2, the power function of this test with same rejection region \( S_1 \) does not depend on \( \phi \).

(ii) Similarly, as in (i), we can easily complete the proof. \( \square \)
Section 2.2: Power functions of specified tests

We will apply this result frequently in the later chapters. In particular, under the same assumptions as Corollary 2.2, if an invariant test for the multivariate elliptically contoured distribution is UMPI (uniformly most powerful invariant test) in the multivariate normal case, then we can claim that this test is also UMPI in the multivariate elliptically contoured distribution.
Chapter 3

MULTIVARIATE REGRESSION ANALYSIS

In this chapter, some properties in a multivariate regression model are studied. The sampling is from the family of multivariate elliptically contoured distributions. Invariant tests are derived and turn out to be the same as those dealt with in the multivariate normal case. Admissibility is a desirable property. Admissibility and null distributions are studied.

3.1. Invariant tests and null distributions.

The usual assumption in the multivariate regression model is that the sample consists of independent and identically distributed normal vectors. Now we consider, the more general case, the following multivariate regression model sampling from the family of multivariate elliptically contoured distributions:

\[
\begin{align*}
Y_{n \times p} &= X_{n \times q}B_{q \times p} + E_{n \times p}, \\
E &\sim \text{LEC}_{n \times p}(0, \Sigma, \phi),
\end{align*}
\]  

(3.1)

where \( X \) is known, \( q \leq p \leq n, n - q \geq p, \text{rank}(X) = q, \) and \( \Sigma > 0. \)

If the density of \( Y \) exists, then the density of \( E \) exists with the form of

\[
|\Sigma|^{-\frac{1}{2}n} g(\text{tr} \Sigma^{-1}(E'E)).
\]  

(3.2)

Hence the density of \( Y \) has the form of

\[
|\Sigma|^{-\frac{1}{2}n} g(\text{tr} \Sigma^{-1}(Y - XB)'(Y - XB)).
\]  

(3.3)
Section 3.1: Invariant tests and null distributions

Suppose we partition

\[
B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \quad X = \begin{pmatrix} X_1 & X_2 \end{pmatrix},
\]

such that \(B_1\) has \(q_1\) rows, \(B_2\) has \(q_2\) rows, \(X_1\) has \(q_1\) columns, and \(X_2\) has \(q_2\) columns. We wish to test \(H : B_1 = B_1^*\) against the general alternative \(K : B_1 \neq B_1^*\), where \(B_1^*\) is a given matrix.

To find the maximal invariants of the sufficient statistics, we need to find the sufficient statistics. Since

\[
Y - XB = (Y - X\hat{B}_n) + X(\hat{B}_n - B),
\]

we have

\[
(Y - XB)'(Y - XB)
= (Y - X\hat{B}_n)'(Y - X\hat{B}_n) + (\hat{B}_n - B)'X'X(\hat{B}_n - B)
= N\hat{\Sigma}_n + (\hat{B}_n - B)'A(\hat{B}_n - B),
\]

where \(N\hat{\Sigma}_n = (Y - X\hat{B}_n)'(Y - X\hat{B}_n)\), \(\hat{B}_n = A^{-1}C\), \(C = X'Y\), and

\[
A = X'X = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}.
\]

From (3.3), the density of \(Y\) has the form of

\[
|\Sigma|^{-\frac{1}{2}n}g(\text{tr}\Sigma^{-1}[N\hat{\Sigma}_n + (\hat{B}_n - B)'A(\hat{B}_n - B)]).
\]

By the Factorization theorem, \(\hat{\Sigma}_n, \hat{B}_n\) form a sufficient set of statistics for \(\Sigma, B\). Thus, there is an one-to-one correspondence between \((\hat{B}_1\Omega, \hat{B}_2\omega)\) and \((\hat{B}_1\Omega, \hat{B}_2\Omega)\) and \(\hat{\Sigma}_n, \hat{B}_1\Omega,\) and \(\hat{B}_2\omega\) form a sufficient set of statistics for \(\Sigma, B\). (Note that \(\hat{B}_2\omega = \hat{B}_2\Omega + (\hat{B}_1\Omega - B_1^*)A_{12}A_{22}^{-1}, \)

where \(\hat{B}_2\omega = A_{22}^{-1}[X_2'(Y - X_1B_1^*)]\).

We can reformulate the hypothesis as \(B_1 = 0\) (by replacing \(Y\) by \(Y - X_1B_1^*)\) and the
problem remains invariant under

\[(1) Y \mapsto Y + X_2 \Gamma, \tag{3.8}\]

\[(2) X_1 \mapsto X_1 C^*, \quad B_1 \mapsto C^{-1} B_1, \]

where $C^*$ is nonsingular,

\[(3) Y \mapsto YV, \]

where $V$ is nonsingular.

Under group (1),

\[
\hat{\Sigma}_0 \mapsto \hat{\Sigma}_0, \quad \hat{B}_{1\Omega} \mapsto \hat{B}_{1\Omega},
\]

\[
\hat{B}_{2\omega} \mapsto \hat{B}_{2\omega} + \Gamma. \tag{3.9}\]

So the only invariants of the sufficient statistics are $\hat{\Sigma}_0$ and $\hat{B}_{1\Omega}$ (by letting $\Gamma = -\hat{B}_{2\omega}$).

Under group (2),

\[
\hat{\Sigma}_0 \mapsto \hat{\Sigma}_0, \quad \hat{B}_{1\Omega} \mapsto C^*^{-1} \hat{B}_{1\Omega},
\]

\[
X_1^* \mapsto X_1^* C^*, \tag{3.10}\]

where $X_1^* = X_1 - X_2 A_{22}^{-1} A_{21}$. Let $C_1^* A_{11} C_1^* = I$ and $C_2^*$ be an orthogonal transformation such that

\[
C_2^*^{-1} C_1^*^{-1} \hat{B}_{1\Omega} = T, \tag{3.11}\]

where $t_{ii} = 0$, $i > v$, $t_{ii} \geq 0$, and $A_{11,2} = X_1^* X_1^* = A_{11} - A_{12} A_{22}^{-1} A_{21}$. Since $T$ is a function of $T'T = \hat{B}_{1\Omega} A_{11^2} \hat{B}_{1\Omega}$, the only invariants of the sufficient statistics are $N \hat{\Sigma}_0$ (denote as $G$) and $\hat{B}_{1\Omega} A_{11,2} \hat{B}_{1\Omega}$ (denote as $H$). Under group (3),

\[
G \mapsto V'GV, \quad H \mapsto V'HV, \tag{3.12}\]
where \( V \) is nonsingular. Hence the only invariants (i.e. maximal invariants) of the sufficient statistics are the roots of

\[
|H - \theta G| = 0. \tag{3.13}
\]

We have the following theorem.

**Theorem 3.1.**

The respective maximal invariants of the sufficient statistics for the sample space are the roots of (3.13).

Since the maximal invariants of the sufficient statistics are same as the ones in multivariate normal case, the invariant tests are same as the ones in multivariate normal case. Let \( \theta_1 \geq \theta_2 \geq \cdots \geq \theta_p \) be the roots of (3.13). Some invariant test criteria are listed as follows:

1. Lawley-Hotelling trace:

\[
\text{tr}HG^{-1} = \sum_{i=1}^{p} \theta_i. \tag{3.14}
\]

2. Bartlett-Nanda-Pillai trace:

\[
\text{tr}H(H + G)^{-1} = \sum_{i=1}^{p} \frac{\theta_i}{1 + \theta_i}. \tag{3.15}
\]

3. Roy's maximum root:

\[
\theta_1. \tag{3.16}
\]

4. Wilks' likelihood ratio criterion:

\[
U = \frac{|G|}{|G + H|} = \prod_{i=1}^{p} (1 + \theta_i)^{-1}. \tag{3.17}
\]

Under the null hypothesis, by Theorem 2.1, the joint distribution of maximal invariants of the sufficient statistics does not depend on \( \phi \). So the null distribution of an invariant test is the same for all \( \phi \)'s (i.e. it is same as the one in multivariate normal case.)
Chapter 3: Multivariate regression analysis

Anderson and Fang (1982b) derived that \( N \hat{\Sigma} \Omega \) is distributed as \( MG_{p,2}(\Sigma; \frac{n-p}{2}; \frac{\phi}{2}) \) (see page 4.) Next we want to find the null distribution of \( H \). Since

\[
\hat{B}_{1\Omega} = A^{(1)} X'E, \tag{3.18}
\]

and

\[
E \sim RU_{n\times p} D_{p\times p}, \tag{3.19}
\]

\[
H = \hat{B}_{1\Omega} A_{11,2} \hat{B}_{1\Omega}
= E' X A^{(1)} A_{11,2} A^{(1)} X'E
= E' F E,
\tag{3.20}
\]

where \( F = X A^{(1)} A_{11,2} A^{(1)} X' \), \( D'D = \Sigma > 0 \), and

\[
A^{-1} = \begin{pmatrix}
A^{(1)} \\
A^{(2)}
\end{pmatrix}.
\]

If \( R^2 \sim \chi^2_{n-p} \) (i.e., \( E \) has a multivariate normal distribution), then \( H = E' F E \sim W_p(\Sigma, q_1) \). From Cochran's Theorem for multivariate normal distributions, we have \( F^2 = F \) and \( \text{rank}(F) = q_1 \). Then by Cochran's Theorem for multivariate elliptically contoured distributions (Anderson and Fang (1982b)), \( H \sim MG_{p,2}(\Sigma; \frac{q_1}{2}; \frac{n-p}{2}; \phi) \).

3.2. Admissibility of invariant tests.

In multivariate normality, one way to approach admissibility is to apply either Stein's Theorem or Schwartz's Theorem (Section 8.10 of Anderson (1984)). In this section, we discuss the problem of admissibility and make an extension of Schwartz's Theorem to multivariate elliptically contoured distributions.

Theorem 3.2.

Assume that \((Y, m, \Omega^*, P^*)\) is a family of distributions and \((Y, m, \Omega, P)\) is a subfamily of \((Y, m, \Omega^*, P^*)\) such that for any bounded function \( f \), \( E_\omega f(Y) \equiv 0 \forall \omega \in \Omega \) implies \( f(Y) \equiv 0 \) a.e. \( m \). Suppose \( \Omega^*_0 \) and \( \Omega_0 \) are the nonempty proper subsets of \( \Omega^* \) and \( \Omega \), respectively,
such that \( \Omega_0 \subset \Omega_\circ \) and \( \Omega \setminus \Omega_0 \subset \Omega^* \setminus \Omega^*_0 \). If the test with acceptance region \( B \) is admissible for testing the hypothesis that \( \omega \in \Omega_0 \) against the alternative \( \omega \in \Omega \setminus \Omega_0 \), then the test with acceptance region \( B \) is admissible for testing the hypothesis that \( \omega \in \Omega_0^* \) against the alternative \( \omega \in \Omega^* \setminus \Omega^*_0 \).

**Proof.**

The critical function of the test with acceptance region \( B \) is \( \phi_B(y) = 0 \), \( y \in B \), and \( \phi_B(y) = 1 \), \( y \notin B \). Suppose \( \phi(y) \) is the critical function of a better test, that is

\[
\int \phi(y) \, dp_\omega(y) \leq \int \phi_B(y) \, dp_\omega(y), \quad \omega \in \Omega_0^*,
\]

\[
\int \phi(y) \, dp_\omega(y) \geq \int \phi_B(y) \, dp_\omega(y), \quad \omega \in \Omega^* \setminus \Omega^*_0,
\]

with strict inequality for some \( \omega \); we shall show that this assumption leads to a contradiction.

Now we restrict the above inequalities to \( \Omega \); we have

\[
\int (\phi(y) - \phi_B(y)) \, dp_\omega(y) \leq 0, \quad \omega \in \Omega_0^*,
\]

\[
\int (\phi(y) - \phi_B(y)) \, dp_\omega(y) \geq 0, \quad \omega \in \Omega \setminus \Omega_0.\]

Since the test with acceptance region \( B \) is admissible for testing the hypothesis that \( \omega \in \Omega_0 \) against the alternative \( \omega \in \Omega \setminus \Omega_0 \),

\[
E_\omega[\phi(Y) - \phi_B(Y)] \equiv 0 \quad \forall \omega \in \Omega.
\]

So \( \phi(Y) \equiv \phi_B(Y) \) a.e. \( m \). This leads to a contradiction. 

Now we go back to the problem of the multivariate regression model. Since \( p \geq q_1 \), there are \( p - q_1 \) roots of (3.13) identically 0. It seems reasonable that if a set of roots \( \theta = (\theta_1, \ldots, \theta_{q_1})' \) leads to acceptance, then a set of roots \( \theta^* = (\theta^*_1, \ldots, \theta^*_q)' \) such that \( \theta^*_i \leq \theta_i \), \( i = 1, \ldots, q_1 \), should also lead to acceptance. Such an acceptance set is called monotone. A
point \( \theta = (\theta_1, \ldots, \theta_{q_1})' \) is in the extended region \( A^* \) if the point \( \theta_\pi = (\theta_{i(1)}, \ldots, \theta_{i(q_1)})' \), where \( i(1), \ldots, i(q_1) \) is a permutation of \( 1, \ldots, q_1 \) such that \( \theta_{i(1)} \geq \cdots \geq \theta_{i(q_1)} \), is in \( A \). The following theorem is a generalization of Schwartz's Theorem.

**Theorem 3.3.**

Let \((\mathcal{Y}, m, \Omega^*, \mathcal{P}^*)\) be the family of multivariate elliptically contoured distributions and \( m \) be the Lebesgue measure. If the acceptance region \( A \) in the set \( 1 \geq \theta_1 \geq \cdots \geq \theta_{q_1} \geq 0 \) is montone and if the extended region \( A^* \) is closed and convex, then \( A \) is the acceptance region of an admissible test.

This theorem is proved using the following corollary from Theorem 3.2, and Theorem 3.4.1 of Ferguson (1967).

**Corollary 3.1.**

Let \((\mathcal{Y}, m, \Omega^*, \mathcal{P}^*)\) be the family of multivariate elliptically contoured distributions and \( m \) be the Lebesgue measure. If the acceptance region \( A \) in the set \( 1 \geq \theta_1 \geq \cdots \geq \theta_{q_1} \geq 0 \) is montone and if the extended region \( A^* \) is closed and convex, then \( A \) is the acceptance region of an admissible test among the class of tests based on the sufficient statistics \((\widehat{\Sigma}_\Omega, \widehat{B}_\Omega)\).

**Proof.**

Let \((\mathcal{Y}, m, \Omega, \mathcal{P})\) be the family of multivariate normal distributions. It is clear that \((\widehat{\Sigma}_\Omega, \widehat{B}_\Omega)\) is the complete sufficient statistics. Then this corollary follows from Theorem 3.2.

Next let us state Theorem 3.4.1 of Ferguson (1967).

**Theorem 3.4.**

Consider the game \((\Theta, \mathcal{A}, L)\) where the statistician observes a random vector \( X \) whose distribution depends on \( \theta \). If \( T \) is a sufficient statistic for \( \theta \), then the set \( D_0 \) of decision
Section 3.2: Admissibility of invariant tests

rules in \( D \), which are based on \( T \), forms an essentially complete class in the game \((\Theta, D, \hat{R})\).

Note:

\( \Theta \) is the parameter space; \( A \) is the space of actions; \( L \) is a loss function; \( D \) is the space of all behavioral decision rules; and \( \hat{R} \) is the risk function (see Ferguson (1967).)

Proof of Theorem 3.3.

For every fixed density \( g \) (see (3.2)), from Corollary 3.1, \( A \) is the acceptance region of an admissible test among the class of tests based on the sufficient statistic \((\hat{\Sigma}_n, \hat{B}_n)\). But \((\hat{\Sigma}_n, \hat{B}_n)\) is the sufficient statistic for \((\Sigma, B)\). Then, from Theorem 3.4, \( A \) is the acceptance region of an admissible test for this fixed \( g \). So \( A \) is the acceptance region of an admissible test.

From Theorem 3.3, we conclude that Wilks' likelihood ratio test, the Lawley-Hotelling trace test, the Bartlett-Nanda-Pillai trace test, and Roy's maximum root test are admissible.
Chapter 4

TESTING HYPOTHESIS FOR MEAN VECTOR

In this chapter an invariant test for equality between the mean vector and a specified vector is derived; other optimum properties derived in Chapter 3 are studied. In the multivariate normal case, the generalized $T^2$-test is the uniformly most powerful invariant test (UMPI), but in the multivariate elliptically contoured case, a weaker property called locally most powerful invariant test (LMPI) is derived in section 4.2.

4.1. Invariant test, null distribution, and admissibility.

Suppose $X_{n \times p} \sim LEC_{n \times p}(\mu, \Sigma, \phi)$, with $\Sigma > 0$, unknown, $n \geq p$, and the density of $X$ exists. We wish to test $H : \mu = \mu_0$ against the general alternative $K : \mu \neq \mu_0$, where $\mu_0$ is a known vector. We can reformulate the hypothesis as $\mu = 0$ (by replacing $X$ by $X - 1_n \mu_0'$) and the problem remains invariant under the following group:

$$X \mapsto XC,$$  \hspace{1cm} (4.1)

where $C$ is $p \times p$ nonsingular.

This problem is a special case of multivariate regression model, so we have the following corollaries.
Section 4.2: Locally most powerful invariant test

Corollary 4.1.

The respective maximal invariant of the sufficient statistics for the sample space is

\[ T^2 = n(\bar{x} - \mu_0)'S^{-1}(\bar{x} - \mu_0) \]

where \( \bar{x} = \frac{1}{n}X'1_n \) and \( S = \frac{1}{n-1}(X - 1_n\bar{x}')'(X - 1_n\bar{x})' \).

Corollary 4.1 implies that the \( T^2 \)-test whose acceptance region is \( T^2 \leq \lambda_1 \), a constant, is an invariant test.

Corollary 4.2.

For the null case, the distribution of \( T^2 \) does not depend on the distribution of the radius \( R \). That is, the null distribution is unique for any \( \phi \).

Corollary 4.3.

\( T^2 \)-test is admissible.

4.2. Locally most powerful invariant test.

In this section, we want to derive a nice result that the \( T^2 \)-test is a locally most powerful invariant test. First we need to find the non-null distribution of \( T^2 \).

If \( X_{n \times p} \sim LEC_{n \times p}(\mu, \Sigma, \phi) \), \( n \geq p > 1 \), and the density of \( X \) exists, then

\[ X \overset{d}{=} 1_n\mu' + RU_{n \times p}D_{p \times p}, \quad (4.2) \]

and

\[ \bar{x} = \frac{1}{n}X'1_n \]

\[ \overset{d}{=} \mu + \frac{1}{n}RD'U'1_n, \quad (4.3) \]

where \( D'D = \Sigma, R \sim F \) and independent of \( U \). So
\[
\frac{T^2}{n-1} = \frac{n}{n-1} (\bar{x} - \mu_0)' S^{-1} (\bar{x} - \mu_0) \\
\leq n[(\mu - \mu_0) + \frac{1}{n} R D' U' \mathbf{1}_n]' [R^2 D' U'(I - \frac{1}{n^2} \mathbf{1}_n \mathbf{1}_n^t) U D]^{-1} \\
\times [(\mu - \mu_0) + \frac{1}{n} R D' U' \mathbf{1}_n] \\
= \left[\frac{n^{\frac{1}{2}}}{R} (D^{-1})' (\mu - \mu_0) + \frac{1}{n^{\frac{1}{2}}} U' \mathbf{1}_n\right]' \left[U'(I - \frac{1}{n^2} \mathbf{1}_n \mathbf{1}_n^t) U\right]^{-1} \\
\times \left[\frac{n^{\frac{1}{2}}}{R} (D^{-1})' (\mu - \mu_0) + \frac{1}{n^{\frac{1}{2}}} U' \mathbf{1}_n\right]. \quad (4.4)
\]

The last equality of (4.4) is true because of \( \Pr\{R = 0\} = 0 \).

Let

\[
Y = \frac{n^{\frac{1}{2}}}{R} (D^{-1})' (\mu - \mu_0) + \frac{1}{n^{\frac{1}{2}}} U' \mathbf{1}_n, \quad (4.5)
\]

and \( Q \) be a \( p \times p \) orthogonal matrix with the first row \( \frac{1}{(Y'Y)^{\frac{1}{2}}} Y' \). Define

\[
V = QY
\]

\[
= \begin{pmatrix}
(Y'Y)^{\frac{1}{2}} \\
0 \\
\vdots \\
0
\end{pmatrix}, \quad (4.6)
\]
Section 4.2: Locally most powerful invariant test

and

\[ B = Q[U'(I - \frac{1}{n}1_n1_n')U]Q' \]

\[ = \begin{pmatrix} b_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}, \tag{4.7} \]

then

\[ \frac{T^2}{n-1} \overset{d}{=} V'B^{-1}V \]

\[ = Y'Yb^{11} \]

\[ = \frac{Y'Y}{b_{11.2}}, \tag{4.8} \]

where \( b^{11} \) is the \((1,1)\)th element of \( B^{-1} \) and \( b_{11.2} = b_{11} - B_{12}B_{22}^{-1}B_{21} \).

Let \( E \) be a \( n \times n \) orthogonal matrix with \( n \)th row \( \frac{1}{n^2}1_n \) and

\[ EU = \begin{pmatrix} u'_1 \\ \vdots \\ u'_n \end{pmatrix} \overset{d}{=} U, \tag{4.9} \]

then

\[ B = QU'UQ' - \frac{1}{n}QU'1_n1_n'UQ' \]

\[ = QU'(E'E)UQ' - \frac{1}{n}QU'1_n1_n'UQ' \]

\[ = Q\left(\sum_{i=1}^{n} u_iu'_i\right)Q' - Qu_nu'_nQ' \]

\[ = Q\left(\sum_{i=1}^{n-1} u_iu'_i\right)Q', \tag{4.10} \]
and

\[ Y = \frac{n^{\frac{1}{2}}}{R} (D^{-1})'(\mu - \mu_0) + u_n. \]  

(4.11)

From lemma 2 due to Cambanis, Huang and Simons (1981),

\[ U \triangleq \begin{pmatrix} (1 - R_1^2)^{\frac{1}{2}} U_1 \\ R_1 u^* \end{pmatrix}, \]

(4.12)

where \( R_1 (\geq 0), U_1 ((n - 1) \times p) \) and \( u^* (p \times 1) \) are independent, and \( R_1^2 \sim \text{Beta}\left(\frac{p}{2}, \frac{(n-1)p}{2}\right) \).

Thus

\[ B \triangleq (1 - R_1^2)QU_1U_1Q' \]

\[ \equiv B(R, R_1, U_1, u^*), \]  

(4.13)

and

\[ Y \triangleq \frac{n^{\frac{1}{2}}}{R} (D^{-1})'(\mu - \mu_0) + R_1 u^* \]

\[ \equiv Y(R, R_1, u^*). \]  

(4.14)

Consequently

\[
\Pr\{T^2 \leq z\} = \Pr \left\{ \frac{Y'(R, R_1, u^*)Y(R, R_1, u^*)}{b_{11.2}(R, R_1, U_1, u^*)} \leq \frac{z}{n - 1} \right\}
\]

\[
= \int_0^\infty \int_0^1 \Pr \left\{ \frac{Y'(r, v^{\frac{1}{2}}, u^*)Y(r, v^{\frac{1}{2}}, u^*)}{b_{11.2}(r, v^{\frac{1}{2}}, U_1, u^*)} \leq \frac{z}{n - 1} \right\} dv \, dF(r),
\]

(4.15)
where $\beta(\cdot ; a, b)$ is the density function of $\text{Beta}(a, b)$.

Conditioned on $u^* = u$ in $B(r, v^{\frac{1}{2}}, U_1, u^*)$, $Q$ is a constant orthogonal matrix. So

$$B(r, v^{\frac{1}{2}}, U_1, u) = (1 - v)Q U_1^T U_1 Q'$$

$$= (1 - v)U_1^T U_1,$$

(4.16)

does not depend on $u$. Hence $B(r, v^{\frac{1}{2}}, U_1, u^*)$ and $Y'(r, v^{\frac{1}{2}}, u^*)Y(r, v^{\frac{1}{2}}, u^*)$ are independent, as are $\delta_{11,2}(r, v^{\frac{1}{2}}, U_1, u^*)$ and $Y'(r, v^{\frac{1}{2}}, u^*)Y(r, v^{\frac{1}{2}}, u^*)$.

Let

$$U_1^T U_1 = E$$

$$= \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix},$$

(4.17)

and

$$U_1 = (u_1^* \quad U_2),$$

(4.18)

where $u_1^*$ is $(n - 1) \times 1$, $U_2$ is $(n - 1) \times (p - 1)$, then

$$(u_1^* \quad U_2) \overset{\Delta}{=} (R_2 v_1 \quad (1 - R_2^2)^{\frac{1}{2}} V_2),$$

(4.19)

where $R_2(\geq 0)$, $v_1$ and $V_2$ are independent, and $R_2^2 \sim \text{Beta}(\frac{n - 1}{2}, \frac{(n - 1)(p - 1)}{2})$. So

$$e_{11,2} = e_{11} - E_{12} E_{22}^{-1} E_{21}$$

$$= u_1^* u_1^* - u_1^* U_2 (U_2^T U_2)^{-1} U_2^T u_1^*$$

$$\overset{\Delta}{=} R_2^2 v_1 v_1 - R_2^2 v_1^T V_2 (V_2^T V_2)^{-1} V_2^T v_1$$

$$= R_2^2 v_1 (I - V_2 (V_2^T V_2)^{-1} V_2^T) v_1.$$

(4.20)
Since \( n \geq p > 1 \), \( V'_2 V_2 \) is nonsingular with probability one. \( A = I - V_2 (V'_2 V_2)^{-1} V'_2 \), conditioned on \( V_2 \), is a constant symmetric matrix, \( A = A^2 \) and \( \text{rank}(A) = (n-1) - (p-1) = n - p \). By Corollary 1 due to Anderson and Fang (1982a), \( v'_1 A v_1 \sim \text{Beta}(\frac{n-p}{2}, \frac{p-1}{2}) \) does not depend on \( V_2 \). So \( v'_1 (I - V_2 (V'_2 V_2)^{-1} V'_2) v_1 \sim \text{Beta}(\frac{n-p}{2}, \frac{p-1}{2}) \) and \( \epsilon_{11,2} \sim \text{Beta}(\frac{n-p}{2}, \frac{n(p-1)}{2}) \). Then

\[
\operatorname{Pr}\{T^2 \leq x\} = \int_0^\infty \int_0^1 \int_0^1 \operatorname{Pr}\left\{ Y'(r, v^1, u^1) Y(r, v^1, u^1) \leq \frac{z(1-v)w}{n-1} \right\} \times \beta(w; \frac{n-p}{2}, \frac{n(p-1)}{2}) \, dv \beta(v; \frac{p}{2}, \frac{(n-1)p}{2}) \, du \, dF(r). \tag{4.21}
\]

Let \( y = zw \), then

\[
\operatorname{Pr}\{T^2 \leq x\} = \int_0^\infty \int_0^1 \int_x^y \operatorname{Pr}\left\{ Y'(r, v^1, u^1) Y(r, v^1, u^1) \leq \frac{y(1-v)}{n-1} \right\} \times \frac{\Gamma\left(\frac{n(p-1)}{2}\right)}{\Gamma\left(\frac{n-p}{2}\right) \Gamma\left(\frac{n(p-1)}{2}\right)} \left[ z^{-\frac{n-p}{2}} y^{-\frac{n-p}{2} - 1} (1 - \frac{y}{z})^{-\frac{n(p-1)}{2} - 1} \right] \, dy \times \beta(v; \frac{p}{2}, \frac{(n-1)p}{2}) \, dv \, dF(r). \tag{4.22}
\]

The density of \( T^2 \) is

\[
f_r(x) = \int_0^\infty \int_0^1 \int_x^y \operatorname{Pr}\left\{ Y'(r, v^1, u^1) Y(r, v^1, u^1) \leq \frac{y(1-v)}{n-1} \right\} \times \frac{\Gamma\left(\frac{n(p-1)}{2}\right)}{\Gamma\left(\frac{n-p}{2}\right) \Gamma\left(\frac{n(p-1)}{2}\right)} \left[ z^{-\frac{n-p}{2}} y^{-\frac{n-p}{2} - 1} (1 - \frac{y}{z})^{-\frac{n(p-1)}{2} - 1} \right] \, dy \beta(v; \frac{p}{2}, \frac{(n-1)p}{2}) \, dv \, dF(r)
\]
Section 4.2: Locally most powerful invariant test

\[= \frac{1}{x} \int_0^\infty \int_0^1 \int_0^1 \Pr \left\{ Y'(r, u^1, u^*) Y(r, v^1, u^*) \leq \frac{x(1-v)w}{n-1} \right\} \]

\[\times \frac{n-p}{2} \left[ \beta(w; \frac{n-p}{2} + 1, \frac{n(p-1)}{2} - 1) - \beta(w; \frac{n-p}{2}, \frac{n(p-1)}{2}) \right] dw \]

\[\times \beta\left( v; \frac{n-1(p)}{2} \right) dv dF(r). \]

(4.23)

We have the following theorem

Theorem 4.1.

If \( X_{n \times p} \sim LEC_{n \times p}(\mu, \Sigma, \phi) \), \( n \geq p > 1 \), and the density of \( X \) exists, then the density of \( T^2 \) has the form (4.23).

Next we want to attack the problem of LMPI by using Theorem 4.1. First let us define LMPI.

Definition 4.1.

A level \( \alpha \) test \( \psi_0 \) is said to be locally most powerful if given any other level \( \alpha \) test \( \psi \), there exists \( \delta \) such that the power of \( \psi_0 \) is bigger than the power of \( \psi \) (i.e. \( P_{\psi_0}(\theta) \geq P_\psi(\theta) \) for all \( \theta \) with \( 0 < d(\theta) < \delta \), where \( d(\theta) \) is a measure of the distance of \( \theta \) from \( H_0 \).

Definition 4.2.

A level \( \alpha \) test \( \psi_0 \) is said to be locally most powerful invariant (LMPI) if the test is locally most powerful among the family of all invariant tests.

Theorem 4.2.

Assume \( X_{n \times p} \sim LEC_{n \times p}(\mu, \Sigma, \phi) \), \( n \geq p > 1 \), and that the density of \( X \) exists. For testing \( H_0 : \mu = \mu_0 \) vs \( H_1 : \mu \neq \mu_0 \), the \( T^2 \)-test is LMPI.
Chapter 4: Testing hypothesis for mean vector

Proof.

Since, from Theorem 4.1, the non-null density of $T^2$ has the form (4.23),

$$
\frac{\partial f_t(x)}{\partial r} = \frac{1}{x} \int_0^1 \int_0^1 \frac{\partial}{\partial r} \Pr \left\{ Y'(R, v \frac{1}{2}, u^*) Y(R, v \frac{1}{2}, u^*) \leq \frac{x(1-v)w}{n-1} \right\} \\
\times \frac{n-p}{2} \left[ \beta(w; \frac{n-p}{2} + 1, \frac{n(p-1)}{2} - 1) - \beta(w; \frac{n-p}{2}, \frac{n(p-1)}{2}) \right] \, dw \\
\times \beta(v; \frac{p}{2}, \frac{n-1}{2}) \, dv.
$$

(4.24)

But

$$
Y'(R, v \frac{1}{2}, u^*) Y(R, v \frac{1}{2}, u^*) = \frac{r^2}{R^2} + 2 \frac{(nu)^{\frac{1}{2}}}{R} (\mu - \mu_0)' D^{-1} u^* + v,
$$

(4.25)

where $r^2 = n(\mu - \mu_0)' \Sigma^{-1} (\mu - \mu_0)$. There exists a $p \times p$ orthogonal matrix $Q_1$ with first row $\frac{n \frac{1}{2} (\mu - \mu_0)' D^{-1}}{r}$ such that

$$
Q_1 u^* \overset{d}{=} \left( \begin{array}{c} BZ_1^{\frac{1}{2}} \\ Z_2 \end{array} \right)
$$

(4.26)

and

$$
Y'(R, v \frac{1}{2}, u^*) Y(R, v \frac{1}{2}, u^*) \overset{d}{=} \frac{r^2}{R^2} + 2 \frac{r}{R} (vZ_1)^{\frac{1}{2}} B + v,
$$

(4.27)
\[ \Pr \left\{ Y'(R, u^\frac{1}{3}, u^*) Y(R, u^\frac{1}{3}, u^*) \leq \frac{x(1-v)w}{n-1} \right\} \]

\[ = \Pr \left\{ \frac{r^2}{R^2} + 2 \frac{r}{R} (vz_1)^{\frac{1}{2}} B + v \leq \frac{x(1-v)w}{n-1} \right\} \]

\[ = \int_0^1 \frac{1}{2} \beta(z_1; \frac{1}{2}, \frac{p-1}{2}) \left[ \Pr \left\{ \frac{r^2}{R^2} + 2 \frac{r}{R} (vz_1)^{\frac{1}{2}} + v \leq \frac{x(1-v)w}{n-1} \right\} \right. \]
\[ \left. + \Pr \left\{ \frac{r^2}{R^2} - 2 \frac{r}{R} (vz_1)^{\frac{1}{2}} + v \leq \frac{x(1-v)w}{n-1} \right\} \right] \, dz_1, \tag{4.28} \]

where \( 0 \leq v, w \leq 1. \)

Since, under \( 0 \leq v, w, z_1 \leq 1 \) and \( \Pr\{R > 0\} = 1, \)

\[ \frac{r^2}{R^2} + 2 \frac{r}{R} (vz_1)^{\frac{1}{2}} + v \leq \frac{x(1-v)w}{n-1} \]

\[ \Leftrightarrow (\frac{r}{R} + (vz_1)^{\frac{1}{2}})^2 + v(1-z_1) \leq \frac{x(1-v)w}{n-1} \]

\[ \Leftrightarrow 0 \leq (\frac{r}{R} + (vz_1)^{\frac{1}{2}})^2 \leq \frac{x(1-v)w}{n-1} - v(1-z_1) \]

\[ \Leftrightarrow 0 \leq \left[ \frac{r}{R} + (vz_1)^{\frac{1}{2}} \right]^2 \leq \frac{x(1-v)w}{n-1} - v(1-z_1) \] \[ \text{and} \]

\[ \frac{x(1-v)w}{n-1} \geq v(1-z_1) \tag{4.29} \]

\[ \Leftrightarrow 0 \leq \frac{r}{R} \leq \left[ \frac{x(1-v)w}{n-1} - v(1-z_1) \right]^{\frac{1}{2}} - (vz_1)^{\frac{1}{2}} \text{ and} \]

\[ \frac{x(1-v)w}{n-1} \geq v(1-z_1), \]

and
\[
\frac{r^2}{R^2} - 2 \frac{r}{R} (vz_1)^{\frac{1}{2}} + v \leq \frac{\frac{1}{n-1}}{n-1} (1-v)w
\]

\[
\iff \left( \frac{r}{R} - (vz_1)^{\frac{1}{2}} \right)^2 + v(1-z_1) \leq \frac{1}{n-1} (1-v)w
\]

\[
\iff -\left[ \frac{\frac{1}{n-1}}{n-1} - v(1-z_1) \right]^{\frac{1}{2}} \leq \frac{r}{R} - (vz_1)^{\frac{1}{2}} \leq \left[ \frac{\frac{1}{n-1}}{n-1} - v(1-z_1) \right]^{\frac{1}{2}},
\]

\[
\frac{1}{n-1} \geq v(1-z_1)
\]

\[
\iff (vz_1)^{\frac{1}{2}} - \left[ \frac{\frac{1}{n-1}}{n-1} - v(1-z_1) \right]^{\frac{1}{2}} \leq \frac{r}{R} \leq (vz_1)^{\frac{1}{2}} + \left[ \frac{\frac{1}{n-1}}{n-1} - v(1-z_1) \right]^{\frac{1}{2}},
\]

\[
\frac{1}{n-1} \geq v(1-z_1),
\]

\[
\Pr \left\{ \mathcal{Y}'(R, u^{\frac{1}{2}}, u^*) \mathcal{Y}(R, v^{\frac{1}{2}}, u^*) \leq \frac{\frac{1}{n-1}}{n-1} (1-v)w \right\}
\]

\[
= \int_{0}^{1} \left[ \frac{1}{2} I_A \cap F \left( \Pr \left\{ R \geq \frac{r}{\left[ \frac{1}{n-1} - v(1-z_1) \right]^{\frac{1}{2}} - (vz_1)^{\frac{1}{2}} \right] } \right) 
\right.
\]

\[
+ \Pr \left\{ R \geq \frac{r}{\left[ \frac{1}{n-1} - v(1-z_1) \right]^{\frac{1}{2}} + (vz_1)^{\frac{1}{2}} } \right\} \right] + \frac{1}{2} I_A \cap F^c
\]

\[
\times \Pr \left\{ \frac{r}{\left[ \frac{1}{n-1} - v(1-z_1) \right]^{\frac{1}{2}} + (vz_1)^{\frac{1}{2}} } \leq R \leq \frac{r}{(vz_1)^{\frac{1}{2}} - \left[ \frac{1}{n-1} - v(1-z_1) \right]^{\frac{1}{2}} } \right\}]
\]

\[
\times \beta(z_1; \frac{1}{2}, \frac{p-1}{2}) dz_1,
\]

where \( A = \left\{ \frac{\frac{1}{n-1}}{n-1} > v(1-z_1) \right\}, \ F = \left\{ \left[ \frac{\frac{1}{n-1}}{n-1} - v(1-z_1) \right]^{\frac{1}{2}} > (vz_1)^{\frac{1}{2}} \right\} \).
and $0 \leq v, w \leq 1$.

We have, from (4.23),

$$\frac{\partial f_r(z)}{\partial r} \bigg|_{r=0}$$

$$= \frac{n-p}{2} \frac{1}{x} \int_0^1 \int_0^1 \int_0^1 \left[ \frac{1}{2} I_A \cap \mathbb{F} \left( \frac{-1}{\left[ \frac{x(1-v)w}{n-1} - v(1-z_1) \right]^{\frac{1}{2}}} \right) \right]$$

$$+ \frac{1}{2} I_A \cap \mathbb{F}^c \left( \frac{1}{(u z_1)^{\frac{1}{2}}} \left[ \frac{1}{\left[ \frac{z(1-v)w}{n-1} - v(1-z_1) \right]^{\frac{1}{2}}} \right] f(0) \right)$$

$$\times \beta(z_1; \frac{1}{2}, \frac{p-1}{2}) dz_1 \left[ \beta(w; \frac{n-p}{2}+1, \frac{n(p-1)}{2}-1) - \beta(w; \frac{n-p}{2}, \frac{n(p-1)}{2}) \right] dw$$

$$\times \beta(v; \frac{p}{2}, \frac{(n-1)p}{2}) dv$$

$$= \frac{n-p}{2} \frac{1}{x} f(0) \int_0^1 \int_0^1 \int_0^1 I_{\mathbb{F}} \left( \frac{-1}{\left[ \frac{x(1-v)w}{n-1} - v(1-z_1) \right]^{\frac{1}{2}}} \right)$$

$$\times \beta(z_1; \frac{1}{2}, \frac{p-1}{2}) dz_1 \left[ \beta(w; \frac{n-p}{2}+1, \frac{n(p-1)}{2}-1) - \beta(w; \frac{n-p}{2}, \frac{n(p-1)}{2}) \right] dw$$

$$\times \beta(v; \frac{p}{2}, \frac{(n-1)p}{2}) dv,$$

(4.32)
where \( f(\cdot) \) is the density function of \( R \).

Thus \( \frac{\partial f (x) |_{x=0}}{f(0)f_0(x)} \) does not depend on the density of \( R \). Since \( f(0) \) is constant and, under multivariate normality, the \( T^2 \)-test is LMP, the \( T^2 \)-test is LMP in multivariate elliptically contoured case.

Kariya (1981a, 1981b) has derived that \( T^2 \)-test is UMP under a subfamily of multivariate elliptically contoured distribution \( g \) (see (3.2)) is nonincreasing convex function.) For the more general case as the one in Theorem 4.2, the property of UMP is still an open problem.
Testing Various Other Hypothesis

In this chapter, invariant tests for testing various hypotheses are studied. There are five sections in this chapter. In each section, the invariant tests and their properties in different hypothesis testing are derived. The order of these problems are test for lack of correlations among the sets, test of equality of covariance matrices and mean vectors, test for known correlation coefficient, test for known partial correlation coefficient, and test for zero multiple correlation coefficient.

5.1. Criteria for testing lack of correlations among the sets.

Consider $X_{n \times p} \sim LEC_{n \times p}(\mu, \Sigma, \phi)$, with $\Sigma > 0$, and $n \geq p$. If the density of $X$ exists, then the density of $X$ has the form of

$$|\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(X - 1_n \mu')'(X - 1_n \mu)ight),$$  

where $1_n = (1, \ldots, 1)$.

Suppose we partition

$$\Sigma = \begin{pmatrix}
\Sigma_{11} & \cdots & \Sigma_{1q} \\
\vdots & & \vdots \\
\Sigma_{q1} & \cdots & \Sigma_{qq}
\end{pmatrix},$$

(5.2)

such that $\Sigma_{ij}, i = 1, \ldots, q, j = 1, \ldots, q$ is $p_i \times p_j$ matrix and $\sum_{i=1}^q p_i = p$. We wish to test $H_0 : \Sigma_{ij} = 0, i \neq j$ against the general alternative $H_1 : \Sigma_{ij} \neq 0$ for some $i \neq j$. First we have to find the sufficient statistics. Since
\[(X - 1_n \mu')'(X - 1_n \mu')
\quad = (X - 1_n \bar{x}')'(X - 1_n \bar{x}') + (1_n(\bar{x} - \mu)'(1_n(\bar{x} - \mu)')
\quad = A + n(\bar{x} - \mu)(\bar{x} - \mu)',
\quad \text{(5.3)}\]

the density of \(X\) has the form of
\[|\Sigma|^{-\frac{1}{2}} g(\text{tr} \Sigma^{-1}[A + n(\bar{x} - \mu)(\bar{x} - \mu)']),\]
\quad \text{(5.4)}

where \(\bar{x} = \frac{1}{n}X'1_n\) and \(A = (X - 1_n \bar{x}')(X - 1_n \bar{x}')\). By the Factorization theorem, \(A\) and \(\bar{x}\) form a sufficient set of statistics for \(\Sigma\) and \(\mu\). This problem remains invariant under
\[(1) X \mapsto X + 1_n d',
\quad (2) X \mapsto XC,\]
\quad \text{(5.5)}

where \(C\) is a matrix with nonsingular diagonal blocks of \(C_i\) with order \(p_i, i = 1, \ldots, q\), and off-diagonal blocks of 0's.

Under group (1),
\[A \mapsto A, \quad \bar{x} \mapsto \bar{x} + d.\]
\quad \text{(5.6)}

So the only invariants of the sufficient statistics are \(A\) (by letting \(d = -\bar{x}\).) Under group (2),
\[A \mapsto \begin{pmatrix} C_1' A_{11} C_1 & \cdots & C_1' A_{1q} C_q \\ \vdots & & \vdots \\ C_q' A_{q1} C_1 & \cdots & C_q' A_{qq} C_q \end{pmatrix},\]
\quad \text{(5.7)}

where \(A = \begin{pmatrix} A_{11} & \cdots & A_{1q} \\ \vdots & & \vdots \\ A_{q1} & \cdots & A_{qq} \end{pmatrix}\) and \(A_{ij}\) is \(p_i \times p_j\) matrix. Let \(A_\theta\) be a matrix with diagonal blocks of \(A_{ii}, i = 1, \ldots, q\), and off-diagonal blocks of 0's, \(B\) be a matrix with diagonal blocks
Section 5.1: Criteria for testing lack of correlations among the sets

of $B_i$, $i = 1, \ldots, q$, and off-diagonal blocks of 0's such that $B_i'A_iB_i = I$, $i = 1, \ldots, q$, then $B'A_0B = I$. The invariants of the sufficient statistics are

$$B'AB = \begin{pmatrix} I & B_1'A_1B_2 & \cdots & B_1'A_1B_q \\ B_2'A_2B_1 & I & \cdots & B_2'A_2B_q \\ \vdots & \vdots & \ddots & \vdots \\ B_q'A_qB_1 & B_q'A_qB_2 & \cdots & I \end{pmatrix}. \quad (5.8)$$

So $Q'B'(A - A_0)BQ$ is invariant with respect to group (2), where $Q$ is a matrix with orthogonal diagonal blocks and off-diagonal blocks of 0's. Hence we conclude that the invariant tests for the multivariate elliptically contoured distributions are same as the ones in multivariate normal case.

Some invariant test criteria are:

1. Likelihood ratio criterion:

$$|B'(A - A_0)B + I|. \quad (5.9)$$

2. Nagao's criterion:

$$\frac{1}{2} \text{tr}[B'(A - A_0)B]^2. \quad (5.10)$$

Since the likelihood ratio criterion is admissible in the multivariate normal case, from Theorem 3.2 and Theorem 3.4, we know that the likelihood ratio criterion is also admissible in the multivariate elliptically contoured case.

Next we want to find the null and non-null distributions for the above criteria. From Corollary 2.1, we have the following corollary.

**Corollary 5.1.**

*For both the null and non-null cases, the distribution of $B'(A - A_0)B$ does not depend on any particular underlying elliptically contoured distribution, i.e. the null and non-null distributions are same as the ones in the multivariate normal case.*
From the above corollary, we know that the distributions of likelihood ratio criterion
and Nagao's criterion are the same as the ones in multivariate normal case. So their
asymptotic distributions are the same as the ones in multivariate normal case.

5.2. Testing of equality of covariance matrices and mean
vectors.

Consider

\[
X = \begin{pmatrix}
  x'_{11} \\
  \vdots \\
  x'_{1n_i} \\
  \vdots \\
  x'_{q1} \\
  \vdots \\
  x'_{qn_q}
\end{pmatrix} \sim MEC_{nxp}(M; \Sigma_1, \ldots, \Sigma_1, \Sigma_2, \ldots, \Sigma_2, \ldots, \Sigma_q, \ldots, \Sigma_q; \phi)
\]

with \(n_i\) \(\Sigma_i\)'s, \(\Sigma_i > 0, n_i > p, i = 1, \ldots, q, \sum_{i=1}^{q} n_i = n\), \(x_{ij}; j = 1, \ldots, n_i, i = 1, \ldots, q\), are
\(p \times 1\) vectors, the rows of \(M\) are \(n_i (\mu_i')\)'s, \(i = 1, \ldots, q\), and the density of \(X\) exists. Then
the density of \(X\) has the following form

\[
\left( \prod_{i=1}^{q} |\Sigma_i|^{-\frac{1}{2}n_i} \right)^{\frac{1}{2}} \phi \left( \sum_{i=1}^{q} \text{tr} \Sigma_i^{-1}G_i \right),
\] (5.11)

where

\[
G_i = \sum_{j=1}^{n_i} (x_{ij} - \mu_i)(x_{ij} - \mu_i)',
\]

\[
= A_i + n_i(\bar{x}_i - \mu_i)(\bar{x}_i - \mu_i)',
\]

\[
A_i = \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)',
\]

and
Section 5.2: Testing of equality of covariance matrices and mean vectors

\[ x_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}. \]

By the Factorization theorem, \( A_i, x_i, i = 1, \ldots, q, \) form a sufficient set of statistics for \( \Sigma_i, \mu_i, i = 1, \ldots, q. \) Now we wish to test \( H_0 : \mu_1 = \cdots = \mu_q \) and \( \Sigma_1 = \cdots = \Sigma_q \) against the general alternative \( H_1 : \mu_i \neq \mu_j \) for some \( i \neq j \) or \( \Sigma_i \neq \Sigma_j \) for some \( i \neq j. \)

The problem remains invariant under

\begin{align*}
1) & x_{ij} \mapsto x_{ij} + d, \quad \forall i, j, \\
2) & D \mapsto DC^{-1}, \quad B \mapsto C^*B, \\
3) & x_{ij} \mapsto Cx_{ij}, \quad \forall i, j,
\end{align*}

(5.12)

where \( C^* \) and \( C \) are nonsingular, \( D = (\mu_1 - \mu_q, \ldots, \mu_{q-1} - \mu_q) \), and \( B \) is a \((q-1) \times n\) matrix with \((i, j)\)th element 1 when \( \sum_{k=1}^{i-1} n_k < j \leq \sum_{k=1}^{i} n_k \), 0 otherwise.

Under group (1),

\[ A_i \mapsto A_i, \quad \bar{x}_i \mapsto \bar{x}_i + d, \quad \forall i. \]

(5.13)

Hence the invariants of the sufficient statistics are \((A_1, \ldots, A_q)\) and \((\bar{x}_1 - \bar{x}_q, \ldots, \bar{x}_{q-1} - \bar{x}_q)\).

Under group (2), as in MANOVA, the invariants of the sufficient statistics are \((A_1, \ldots, A_q)\) and

\[ H = (\bar{x}_1 - \bar{x}_q, \ldots, \bar{x}_{q-1} - \bar{x}_q)E(\bar{x}_1 - \bar{x}_q, \ldots, \bar{x}_{q-1} - \bar{x}_q)', \]

where \( E = (e_{ij}) \) is a \((q-1) \times (q-1)\) matrix with the element \( e_{ij} = \delta_{ij}n_i - \frac{1}{n} n_i n_j. \) Since

\[ A = \sum_{i=1}^{q} \sum_{j=1}^{n_i} (x_{ij} - \bar{x})(x_{ij} - \bar{x})', \]

\[ = H + \sum_{i=1}^{q} A_i, \]

(5.14)
where \( \bar{x} = \frac{1}{n} \sum_{i=1}^{q} \sum_{j=1}^{n_i} x_{ij} \), the invariants of the sufficient statistics are \( A, A_1, \ldots, A_q \). Under group (3),

\[
A \mapsto CAC', \quad A_i \mapsto CA_iC', \quad \forall i,
\]

where \( C \) is nonsingular.

Let \(CAC' = I\), then \(QCA_iC'Q', i = 1, \ldots, q\), is invariant with respect to group (3), where \( Q \) is an \( p \times p \) orthogonal matrix. Therefore we conclude that the invariant tests in multivariate elliptically contoured distribution are same as the ones in multivariate normal case.

Some invariant test criteria are:

1. Likelihood ratio criterion:

\[
\prod_{i=1}^{q} |A_i|^{\frac{1}{2}n_i} \over |A|^{\frac{1}{2}n}.
\]

(5.16)

2. Bartlett modified likelihood ratio criterion:

\[
\prod_{i=1}^{q} |A_i|^{\frac{1}{2} (n_i - 1)} \over |A|^{\frac{1}{2} (n - q)}.
\]

(5.17)

For dealing with the null distribution we have the following corollary from Theorem 2.1.

**Corollary 5.2.**

Under \( \mu_1 = \cdots = \mu_q \), the joint distribution of \( CA_iC', i = 1, \ldots, q \), does not depend on any particular underlying elliptically contoured distribution, i.e. it is same as the one in the multivariate normal case.

From the above corollary, the null distribution of the invariant test criterion in the multivariate elliptically contoured distribution is the same as the one in the multivariate normal case.
5.3. Test of Hypothesis about correlation coefficient.

Assume $X_{n \times p} \sim LEC_{n \times p}(\mu, \Sigma, \phi)$, with $\Sigma > 0$ and $n \geq p$, and the density of $X$ exists, where $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, $\Sigma_{11} = \begin{pmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix}$ is a $2 \times 2$ matrix. We wish to test $H_0 : \rho = \rho_0$, where $|\rho_0| < 1$ is a given number.

Let $B = \begin{pmatrix} I_{2 \times 2} \\ 0 \end{pmatrix}$ be a $p \times 2$ matrix, then from Corollary 1 of Anderson and Fang (1982b), $XB \sim LEC_{n \times 2}(\mu B, \Sigma_{11}, \phi^*)$ with $\phi^* \in \Phi_{2n} \leftrightarrow R^* = R_{n, 1}^p R_{n, 1}^{-1} \sim Beta(n, \frac{1}{2} n(p - 2))$. So, without loss of generality, let $p = 2$. i.e. $X_{n \times 2} \sim LEC_{n \times 2}(\mu, \Sigma, \phi)$, with $\Sigma > 0$, and $n \geq 2$, and let the density of $X$ exist, where $\Sigma = \begin{pmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix}$. The density of $X$ has the form of

$$|\Sigma|^{-\frac{1}{2}} g(tr \Sigma^{-1} G), \quad (5.18)$$

where

$$G = (X - 1_n \mu')(X - 1_n \mu')$$
$$= A + n(\bar{x} - \mu)(\bar{x} - \mu)'$$

$$\bar{x} = \frac{1}{n} X'1_n,$$

and

$$A = (X - 1_n \bar{x}')(X - 1_n \bar{x}')$$
$$= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad (5.19)$$

By the Factorization theorem, $A$ and $\bar{x}$ form a sufficient set of statistics for $\Sigma$ and $\mu$. The problem remains invariant under
Chapter 5: Testing various other hypothesis

(1) \( X \mapsto X + 1_0d' \),
(2) \( X \mapsto XG \),

where \( C = \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix} \), and \( c_1, c_2 > 0 \).

Under group (1),

\[ A \mapsto A, \quad \bar{x} \mapsto \bar{x} + d, \]  

and the only invariants of the sufficient statistics are \( A \) (by letting \( d = -\bar{x} \)).

Under group (2),

\[ A \mapsto \begin{pmatrix} c_1^2a_{11} & c_1c_2a_{12} \\ c_1c_2a_{21} & c_2^2a_{22} \end{pmatrix}. \]

The only invariant (i.e. maximal invariant) of the sufficient statistics is \( r_{12} = \frac{a_{12}}{\sqrt{a_{11}a_{22}}} \) (by letting \( c_1 = a_{11}^{-\frac{1}{2}}, c_2 = a_{22}^{-\frac{1}{2}} \)). We conclude that the invariant tests in multivariate elliptically contoured distribution are same as the ones in multivariate normal case.

For finding the null and non-null distributions of \( r_{12} \), from Corollary 2.1, we have the following corollary.

**Corollary 5.3.**

*For both the null and non-null cases, the distribution of \( r_{12} \) does not depend on any particular underlying multivariate elliptically contoured distribution.*

From the above corollary, all the optimum properties of \( r_{12} \) for multivariate normal case is fulfilled for multivariate elliptically contoured case, such as : asymptotic normality, and "Fisher's z" test. From Corollary 2.2, \( r_{12} \) is UMPI against alternatives \( \rho > \rho_0 \).
5.4. Test of hypothesis about partial correlation coefficient.

Assume $X_{nxp} \sim \text{LECG}_{nxp}(\mu, \Sigma, \phi)$, with $\Sigma > 0$ and $n \geq p > q$, and the density of $X$ exists, where $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, and $\Sigma_{11}$ is a $q \times q$ matrix. We wish to test $H_0 : \rho_{12,q+1,\ldots,p} = \rho_0$, where $|\rho_0| < 1$ is a given number. Without loss of generality, assume that $q = 2$ (by the same technique as in section 5.3.)

Let $X = (X_1, X_2)$, where $X_1 = (x_1, x_2)$ is $n \times 2$, $X_2$ is $n \times (p-2)$, and $x_1, x_2$ are $n \times 1$ vectors. The density of $X$ has the form of

$$|\Sigma|^{-\frac{1}{2}} g(\text{tr} \Sigma^{-1} G), \quad (5.23)$$

where

$$G = (X - 1_n \mu')(X - 1_n \mu')$$

$$= A + n(\bar{x} - \mu)(\bar{x} - \mu)'$$

where

$$\bar{x} = \frac{1}{n} X' 1_n,$$

$$A = (X - 1_n \bar{x}')(X - 1_n \bar{x}')$$

$$= \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}, \quad (5.24)$$

and $\mathbf{A}_{11} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ is a $2 \times 2$ matrix.

By the Factorization theorem, $A$ and $\bar{x}$ form a sufficient set of statistics for $\Sigma$ and $\mu$.

The problem remains invariant under
(1) \( X \mapsto X + 1_n d' \),

(2) \( X \mapsto X \begin{pmatrix} I & 0 \\ 0 & C \end{pmatrix} \),

\[ (3) X \mapsto X \begin{pmatrix} I & 0 \\ B & I \end{pmatrix}, \tag{5.25} \]

(4) \( X_1 \mapsto X_1 \begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix} \),

where \( C \) is a \((p - 2) \times (p - 2)\) non-singular matrix, \( B \) is a \((p - 2) \times 2\) matrix, and \( b_1, b_2 > 0 \).

Under group (1),

\[ A \mapsto A, \quad \bar{x} \mapsto \bar{x} + d. \tag{5.26} \]

So the only invariants of the sufficient statistics are \( A \) (by letting \( d = -\bar{x} \)). Under group (2),

\[ A \mapsto \begin{pmatrix} A_{11} & A_{12}C \\ C'A_{21} & C'A_{22}C \end{pmatrix}. \tag{5.27} \]

Hence the only invariants of the sufficient statistics are \( A_{11}, A_{12}C \) (by letting \( C'A_{21}C = I \)).

Under group (3),

\[ A_{11} \mapsto A_{11} + B'C'A_{21} + A_{12}CB + B'B, \]

\[ A_{12}C \mapsto A_{12}C + B'. \tag{5.28} \]

The only invariants of the sufficient statistics are
Section 5.5: Test of hypothesis about the multiple correlation coefficient

\[ A_{11.2} = A_{11} - A_{12} A_{22}^{-1} A_{21} \]
\[ = \begin{pmatrix} a_{11.3,\ldots,p} & a_{12.3,\ldots,p} \\ a_{21.3,\ldots,p} & a_{22.3,\ldots,p} \end{pmatrix} \]

(by letting \( B' = -A_{12} \)). Under group (4),

\[ A_{11.2} \leftrightarrow \begin{pmatrix} b_1^2 a_{11.3,\ldots,p} & b_1 b_2 a_{12.3,\ldots,p} \\ b_1 b_2 a_{21.3,\ldots,p} & b_2^2 a_{22.3,\ldots,p} \end{pmatrix} \]  

(5.29)

Thus the only invariant (i.e., maximal invariant) of the sufficient statistics is \( r_{12.3,\ldots,p} = \frac{b_1 b_2}{\sqrt{a_{11.3,\ldots,p} a_{22.3,\ldots,p}}} \) (by letting \( b_i = a_{ii.3,\ldots,p}^{-\frac{1}{2}}, \) \( i = 1, 2 \)). We conclude that the invariant test in multivariate elliptically contoured distribution is same as the one in multivariate normal case.

Next we want to find the null and non-null distributions of \( r_{12.3,\ldots,p} \) for multivariate elliptically contoured distribution.

**Corollary 5.4.**

For both the null and non-null cases, the distribution of \( r_{12.3,\ldots,p} \) does not depend on the underlying multivariate elliptically contoured distribution.

**Proof.**

Clear from Corollary 2.1.

From the above corollary and Theorem 4.3.5 due to Anderson (1984), the distribution of a sample partial correlation coefficient, \( r_{12.3,\ldots,p} \), based on \( X_{nxp} \) with population partial correlation coefficient \( \rho_{12.3,\ldots,p} \) equal to a certain value, \( \rho \), is the same as the distribution of an ordinary correlation coefficient based on \( X_{[n-(p-1)]xp} \) with corresponding population correlation of \( \rho \). Also the asymptotic normality and "Fisher's z" test are fulfilled for the multivariate elliptically contoured case. Obviously, \( r_{12.3,\ldots,p} \) is UMPI against alternatives \( \rho_{12.3,\ldots,p} > \rho_0 \).
5.5. Test of hypothesis about the multiple correlation coefficient.

Assume $X_{n \times p} \sim LBC_{n \times p}(\mu, \Sigma, \phi)$, with $\Sigma > 0$ and $n \geq p$, and the density of $X$ exists, where $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, and $\Sigma_{11}$ is a $q \times q$ matrix. We wish to test $H_0 : \tilde{R}_{1, \varphi + 1, \ldots, \varphi} = 0$, against the general alternative $H_1 : \tilde{R}_{1, \varphi + 1, \ldots, \varphi} > 0$. So, without loss of generality, assume that $q = 1$ (by the same technique as the one in section 5.3.)

Let $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{1(2)} \\ \sigma_{1(2)} & \Sigma_{22} \end{pmatrix}$, then $\tilde{R}_{1, 2, \ldots, \varphi} = \frac{\sigma_{11}^2}{\sigma_{11}^2 \Sigma_{22}} \sigma_{1(2)}^2 = 0$ iff $\sigma_{1(1)} = 0$. The test is same as $H_0 : \sigma_{1(1)} = 0$ against $H_1 : \sigma_{1(1)} \neq 0$. This is a special case of testing for lack of correlations among the sets. Then, by the result in section 5.1, $B_2 a_{(1)} b_1$ are the invariants of the sufficient statistics, where $B_2 A_{22} B_2 = I$, $b_1^2 \sigma_{11} = 1$, and $A = \begin{pmatrix} a_{11} & a_{1(2)} \\ a_{1(2)} & A_{22} \end{pmatrix}$. Since the problem remains invariant under the group $\{Q \mid X \mapsto XQ\}$, where $Q = \begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix}$ and $Q_1$ is $(p - 1) \times (p - 1)$ orthogonal. Under this group,

$$B_2 a_{(1)} b_1 \mapsto Q_1 B_2 a_{(1)} b_1.$$  \hspace{1cm} (5.30)

The only invariant (i.e. maximal invariant) of the sufficient statistics is $b_1 (a_{(1)}' B_2 B_2 a_{(1)})^\frac{1}{2} = (a_{(1)}' A_{22}^{-1} a_{(1)})^\frac{1}{2} = R$ (by letting the first column of $Q_1$ is $B_2 a_{(1)}$.) Also, by Corollary 5.1, we claim that the null and non-null distributions of $R$ in multivariate elliptically contoured distribution are same as the ones in multivariate normal case. So $R$ is UMPI. Also, the invariant test is same as the one in multivariate normal case.
Chapter 6

MAXIMUM LIKELIHOOD ESTIMATES AND LIKELIHOOD RATIO TESTS

In this chapter maximum likelihood estimates (MLE) and likelihood ratio tests (LRT) for the multivariate elliptically contoured distribution are derived. We first give the result for a special case, then some examples, and finally the general theorem.

6.1. The result for a special case.

Anderson and Fang (1982c) attacked the problems of MLE and LRT and gave some conditions for finding MLE. In this section, more precise and weaker conditions are given. The following lemma contains the general idea for constructing the weaker conditions in finding MLE.

Lemma 6.1.

Assume that $g(\cdot)$ is a continuous function such that $g(x_1^2 + \cdots + x_N^2)$ is the density of $EC_N(0, 1, \phi)$ and $E(R^2) < \infty$ where $R \leftrightarrow \phi \in \Phi_N$. Then the function

$$h(x) \equiv x^{N/2} g(x), \quad x \geq 0, \quad (6.1)$$

has a maximum at some finite $x_0 > 0$.

Proof.

Since, from Cambanis, Huang and Simons (1981) (or Lemma 13.3.1 of Anderson
Chapter 6: Maximum likelihood estimates and likelihood ratio tests 45

(1984)), the density function of $R^2$ is

$$f(z) = \frac{\pi^{N/2}}{\Gamma(N/2)} z^{N/2 - 1} g(z),$$

$$h(z) = \frac{\Gamma(N/2)}{\pi^{N/2}} zf(z).$$

(6.2)

From $E(R^2) < \infty$, $\lim_{z \to \infty} h(z) = 0$. Furthermore $h(0) = 0$. (If not, there exists $\epsilon > 0$, $xf(z) > \epsilon \forall z > \delta$, so $1 \geq \int_0^\delta f(x) \, dx > \int_\delta^{\infty} z \, dx = \infty$ which is a contradiction.) Hence $h(z) \geq 0 \forall z \geq 0$, and $h(\cdot)$ is continuous. The assertion of the lemma follows.

The following theorem is the main theorem in this section and let us define the assumption for this theorem.

**Assumption A.**

$X \sim MEG_{nxp}(M; \Sigma^{(1)}, \ldots, \Sigma^{(q)}, \ldots, \Sigma^{(q)}, \ldots, \Sigma^{(q)}; \phi)$ with $n_i \Sigma^{(i)} \phi$, $n_i > p$, $i = 1, \ldots, q$, and $\sum_{i=1}^q n_i = n$. The rows of $M$ are $n_1 \mu^{(1)} \phi$, $\ldots$, $n_q \mu^{(q)} \phi$, successively, where $\Sigma^{(i)} = \text{diag}(\Sigma^{(i)}_{11}, \ldots, \Sigma^{(i)}_{kk})$, $1 \leq k \leq p$. The density function of $X$ is

$$\left[ \prod_{i=1}^q \prod_{j=1}^k \left| \Sigma^{(i)}_{ij} \right|^{-\frac{1}{2} n_i} \right] g \left( \sum_{i=1}^q \sum_{j=1}^k \text{tr} \Sigma^{(i)}_{ij}^{-1} G^{(i)}_{ij} \right),$$

(6.3)

where $G_i = \sum_{j=n_i-1+1}^n (x_{(j)} - \mu^{(i)})(x_{(j)} - \mu^{(i)})' = \begin{pmatrix} G_{11}^{(i)} & \cdots & G_{ik}^{(i)} \\ \vdots & \vdots & \vdots \\ G_{k1}^{(i)} & \cdots & G_{kk}^{(i)} \end{pmatrix}$ with $n_0 = 0$, $n_i = n_1 + \cdots + n_i$.

**Theorem 6.1.**

*Under the assumption A, if $f(\lambda) = \lambda^{-\frac{1}{2} np} g(\frac{\lambda}{p})$ attains its maximum at some finite*
Section 6.1: The result for a special case

\( \lambda_{\text{max}}(g) \), then the MLEs of \( \mu^{(i)} \) and \( \Sigma^{(i)} \) are

\[
\hat{\mu}^{(i)} = \bar{x}^{(i)}, \\
\hat{\Sigma}^{(i)} = \text{diag}(\hat{\Sigma}_{11}^{(i)}, \ldots, \hat{\Sigma}_{kk}^{(i)}),
\]

(6.4)
i = 1, \ldots, q, where \( \hat{\Sigma}_{jj}^{(i)} = \left( n_i \right) \lambda_{\text{max}}(g) W_{jj}^{(i)}, i = 1, \ldots, q, j = 1, \ldots, k. \) \( W_{jj}^{(i)} \) is obtained from \( G_{jj}^{(i)} \) by substituting \( \mu^{(i)} \) with \( \bar{x}^{(i)} = \frac{1}{n_i} \sum_{j=1}^{n_i} x^{(i)} \) and the corresponding maximized likelihood is

\[
\lambda_{\text{max}}(g)^{-\frac{1}{2}np} g \left( \frac{p}{\lambda_{\text{max}}(g)} \right) \prod_{i=1}^{q} \left( \frac{1}{n_i} \right)^{-\frac{1}{2}np} \prod_{j=1}^{k} |W_{jj}^{(i)}|^{-\frac{1}{2}n_i}. \]

(6.5)

Proof.

\[
\max_{\mu^{(i)}, \Sigma^{(i)}} L(\mu^{(1)}, \ldots, \mu^{(q)}; \Sigma^{(1)}, \ldots, \Sigma^{(q)}) \\
= \max_{\mu^{(i)}} \max_{\Sigma^{(i)}} L(\mu^{(1)}, \ldots, \mu^{(q)}; \Sigma^{(1)}, \ldots, \Sigma^{(q)}),
\]

(6.6)

where

\[
L \equiv L(\mu^{(1)}, \ldots, \mu^{(q)}; \Sigma^{(1)}, \ldots, \Sigma^{(q)}) \\
= \prod_{i=1}^{q} \prod_{j=1}^{k} \left| \Sigma_{jj}^{(i)} \right|^{-\frac{1}{2}n_i} g \left( \sum_{i=1}^{q} \sum_{j=1}^{k} \text{tr} \Sigma_{jj}^{(i)} G_{jj}^{(i)} \right).
\]

(6.7)

First we want to maximize \( L \) with respect to \( \Sigma^{(i)} \), \( i = 1, \ldots, q \). Since \( G_{jj}^{(i)} > 0, i = 1, \ldots, q, j = 1, \ldots, k \), with probability one, there exist nonsingular matrices \( C_{ij} \) such that \( C_{ij} C_{ij}^{-1} = G_{jj}^{(i)} \) and orthogonal matrices \( \Gamma_{ij} \) such that

\[
\Gamma_{ij} C_{ij}^{-1} \Sigma_{jj}^{(i)} C_{ij}^{-1} \Gamma_{ij} = \Lambda_{ij} \\
= \text{diag}(\lambda_{j1}^{(i)}, \ldots, \lambda_{jp_j}^{(i)}),
\]
Chapter 6: Maximum likelihood estimates and likelihood ratio tests

with positive \( \lambda_{j_0}^{(i)}, \alpha = 1, \ldots, p_j \). Therefore

\[
L = \left[ \prod_{i=1}^{q} \left( \prod_{j=1}^{k} \prod_{\alpha=1}^{p_j} \lambda_{j_0}^{(i)} \right)^{-\frac{1}{2}n_i} \right] g \left( \sum_{i=1}^{q} \sum_{j=1}^{k} \sum_{\alpha=1}^{p_j} 1/\lambda_{j_0}^{(i)} \right). \tag{6.8}
\]

It is a symmetric function of \( \lambda_{j_0}^{(i)} (j = 1, \ldots, k, \alpha = 1, \ldots, p_j) \) for \( i = 1, \ldots, q \). Hence \( \lambda_{j_0}^{(i)} = \lambda^{(i)}, j = 1, \ldots, k, \alpha = 1, \ldots, p_j, \) and

\[
L = \left[ \left( \prod_{i=1}^{q} \prod_{j=1}^{k} \left| G_{ij}^{(i)} \right|^{-\frac{1}{2}n_i} \right) \prod_{i=1}^{q} \lambda^{(i)-\frac{1}{2}n_ip} \right] g \left( p \sum_{i=1}^{q} \frac{1}{\lambda^{(i)}} \right). \tag{6.9}
\]

Let

\[
\frac{1}{\lambda} = \sum_{i=1}^{q} \frac{1}{\lambda^{(i)}}, \tag{6.10}
\]

\[
\beta_j = \frac{\lambda}{\lambda^{(j)}}, \quad j = 1, \ldots, q - 1.
\]

So maximising \( L \) with respect to \( \Sigma^{(i)}, i = 1, \ldots, q \), is the same as maximising \( L \) with respect to \( \lambda, \beta_j, j = 1, \ldots, q - 1 \).

Now

\[
L = \left( \prod_{i=1}^{q} \prod_{j=1}^{k} \left| G_{ij}^{(i)} \right|^{-\frac{1}{2}n_i} \right) \left( \prod_{i=1}^{q} \lambda^{(i)-\frac{1}{2}n_ip} \right) g \left( p \sum_{i=1}^{q} \frac{1}{\lambda^{(i)}} \right)
\]

\[
= \left( \prod_{i=1}^{q} \prod_{j=1}^{k} \left| G_{ij}^{(i)} \right|^{-\frac{1}{2}n_i} \right) \left[ \left( \prod_{i=1}^{q-1} \beta_i^{\frac{1}{2}n_ip} \right) \left( 1 - \sum_{i=1}^{q-1} \beta_i \right)^{\frac{1}{2}n_ip} \right] \lambda^{-\frac{1}{2}n_p} g \left( \frac{p}{\lambda} \right). \tag{6.11}
\]

So

\[
\max_{\lambda, \beta_1, \ldots, \beta_{q-1}} L
\]

\[
= \prod_{i=1}^{q} \prod_{j=1}^{k} \left| G_{ij}^{(i)} \right|^{-\frac{1}{2}n_i} \left[ \max_{\beta_1, \ldots, \beta_{q-1}} \left( \prod_{i=1}^{q-1} \beta_i^{\frac{1}{2}n_ip} \right) \left( 1 - \sum_{i=1}^{q-1} \beta_i \right)^{\frac{1}{2}n_ip} \right] \left[ \max_{\lambda} \lambda^{-\frac{1}{2}n_p} g \left( \frac{p}{\lambda} \right) \right]. \tag{6.12}
\]
Section 6.1: The result for a special case

$L$ attains its maximum at $\lambda = \lambda_{\text{max}}(g)$, $\beta_i = \frac{n_i}{n}$, $i = 1, \ldots, q - 1$ and

$$L^* = \max_{\Sigma^{(i)}} \frac{p}{\lambda_{\text{max}}(g)} \prod_{i=1}^{q} \left( \left( \frac{n}{n_i} \right)^{-\frac{1}{2}n_i} \prod_{j=1}^{k} |G_{jj}^{(i)}|^{-\frac{1}{2}n_i} \right).$$

(6.13)

Since $G_i = W_i + n_i(\bar{x}^{(i)} - \mu^{(i)})^T(\bar{x}^{(i)} - \mu^{(i)})^T$, maximizing $L^*$ with respect to $\mu^{(i)}$, $i = 1, \ldots, q$, shows that $L$ attains its maximum at $\mu^{(i)} = \bar{x}^{(i)}$. This completes the proof.

Combining the above lemma and theorem, we have the following corollary.

Corollary 6.1.

Under the assumption $A$, if $g(\cdot)$ is continuous and $E(R^2) < \infty$ where $R \leftrightarrow \phi \in \Phi_{np}$, then the MLEs of $\mu^{(i)}$ and $\Sigma^{(i)}$ are

$$\hat{\mu}^{(i)} = \bar{x}^{(i)},$$

$$\hat{\Sigma}^{(i)} = \text{diag}(\hat{\Sigma}_{11}^{(i)}, \ldots, \hat{\Sigma}_{kk}^{(i)}),$$

(6.14)

$i = 1, \ldots, q$, where $\hat{\Sigma}_{jj}^{(i)} = \frac{1}{n_i} \lambda_{\text{max}}(g) W_{jj}^{(i)}$, $i = 1, \ldots, q$, $j = 1, \ldots, k$, and the corresponding maximum of the likelihood function is

$$\lambda_{\text{max}}(g)^{-\frac{1}{2}np} \left( \frac{p}{\lambda_{\text{max}}(g)} \right) \prod_{i=1}^{q} \left( \left( \frac{n}{n_i} \right)^{-\frac{1}{2}n_i} \prod_{j=1}^{k} |W_{jj}^{(i)}|^{-\frac{1}{2}n_i} \right).$$

(6.15)

Under some conditions, Theorem 6.1 and Corollary 6.1 show that the maximum of the likelihood function for the multivariate elliptically contoured distribution is proportional to that of the multivariate normal distribution. So the following likelihood ratio criteria in the multivariate elliptically contoured distribution are the same as those in the multivariate normal case:

1. The criterion for testing for lack of correlation among the sets ($q = 1$):

$$r_1 = \frac{|W|}{\prod_{j=1}^{k} |W_{jj}|}.$$  

(6.16)
(2) The criterion for testing the hypothesis that a mean vector is equal to a given vector \( \mu_0 \) \((q = k = 1)\):

\[
\tau_2 = \frac{|W|}{|W_0|} = \frac{1}{1 + T^2/(n - 1)}.
\]

(6.17)

where \( W_0 = \sum_{j=1}^{n} (x_{ij} - \mu_0)(x_{ij} - \mu_0)' \) and \( T^2 \) is the Hotelling \( T^2 \)-statistic (see Chapter 4.)

(3) The criterion for testing the hypothesis of equality of covariance matrices \((k = 1)\):

\[
\tau_3 = \prod_{i=1}^{q} \left[ \left( \frac{|W_i|}{\sum_{j=1}^{q} W_j} \right)^{\frac{1}{2} n_i} \left( \frac{n}{n_i} \right)^{\frac{1}{2} p n_i} \right].
\]

(6.18)

(4) The criterion for testing equality of several means \((k = 1)\):

\[
\tau_4 = \frac{|W|}{\sum_{j=1}^{q} W_j}.
\]

(6.19)

(5) The criterion for testing equality of several means and covariance matrices \((k = 1)\):

\[
\tau_5 = \prod_{i=1}^{q} \left( \frac{|W_i|}{|W|} \right)^{\frac{1}{2} n_i} \left( \frac{n}{n_i} \right)^{\frac{1}{2} p n_i}.
\]

(6.20)

(6) The criterion for testing the hypothesis that a covariance matrices is proportional to a given matrix \( \Sigma_0 \) \((q = k = 1)\):

\[
\tau_6 = \frac{|\Sigma_0^{-1} W|}{(tr \Sigma_0^{-1} W/p)^p}.
\]

(6.21)
Section 6.1 gives six examples for finding LRT. In this section, three examples that do not fulfill the assumptions in Theorem 6.1 are given. From the results of these three examples, we get the idea of general theorem in section 6.3. The following three examples satisfy the assumption that $f(\cdot)$ attains its maximum at $\lambda_{\text{max}}(g)$.

**Example 1.** Tests of hypotheses concerning subvectors of $\mu$.

If $X_{n \times p} \sim LEC_{n \times p} \left( \left( \begin{array}{c} \mu_1 \\ \mu_2 \end{array} \right), \Sigma, \phi \right)$ with $\Sigma > 0$ and $n \geq p$. We wish to test $H_0 : \mu_1 = 0$. The likelihood function has the following form:

$$L(\mu_1, \mu_2, \Sigma) = |\Sigma|^{-\frac{1}{2}n} g(\text{tr} \Sigma^{-1} G),$$

(6.22)

where $G = \left( X - I_n \left( \begin{array}{c} \mu_1 \\ \mu_2 \end{array} \right) \right)^{\prime \prime} \left( X - I_n \left( \begin{array}{c} \mu_1 \\ \mu_2 \end{array} \right) \right)$.

Let $\Omega = \{ (\mu, \Sigma) \mid \Sigma > 0 \}$ and $\omega = \{ (\mu, \Sigma) \mid \mu_1 = 0, \Sigma > 0 \}$. It is clear that

$$\max_{\Omega} L(\mu_1, \mu_2, \Sigma) = \lambda_{\text{max}}(g)^{-\frac{1}{2}n} g \left( \frac{p}{\lambda_{\text{max}}(g)} \right) |W|^{-\frac{1}{2}n},$$

(6.23)

where $W$ is obtained from $G$ by substituting $\mu_i$ with $x_i, i = 1, 2$. Since

$$\max_{\omega} L(\mu_1, \mu_2, \Sigma) = \max_{\mu_2} \max_{\Sigma} L(0, \mu_2, \Sigma),$$

(6.24)

using a similar argument as the proof of Theorem 6.1, we get

$$\max_{\Sigma} L(0, \mu_2, \Sigma) = \lambda_{\text{max}}(g)^{-\frac{1}{2}n} g \left( \frac{p}{\lambda_{\text{max}}(g)} \right) |H|^{-\frac{1}{2}n},$$

(6.25)

where
Chapter 6: Maximum likelihood estimates and likelihood ratio tests

\[ H = \left( X - I_n \begin{pmatrix} 0 \\ \mu_2 \end{pmatrix} \right)' \left( X - I_n \begin{pmatrix} 0 \\ \mu_2 \end{pmatrix} \right) \]

\[ = W + n \left( \bar{x}_1 \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 - \mu_2 \end{pmatrix} \right)' \]

Thus

\[ \left( \bar{x}_1', \left( \bar{x}_2 - \mu_2 \right)' \right) = \left( \bar{x}_1', \left( \bar{x}_2 - \mu_2 \right)' - \bar{x}_1' W_{11}^{-1} W_{12} \right) \begin{pmatrix} I & W_{11}^{-1} W_{12} \\ 0 & I \end{pmatrix}, \quad (6.26) \]

\[ \begin{pmatrix} I & W_{11}^{-1} W_{12} \\ 0 & I \end{pmatrix}^{-1} = \begin{pmatrix} I & -W_{11}^{-1} W_{12} \\ 0 & I \end{pmatrix}, \quad (6.27) \]

and

\[ \begin{pmatrix} I & 0 \\ -W_{21} W_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix} \begin{pmatrix} I & -W_{11}^{-1} W_{12} \\ 0 & I \end{pmatrix} = \begin{pmatrix} W_{11} & 0 \\ 0 & W_{22-1} \end{pmatrix}, \quad (6.28) \]

imply that

\[ |H| = \left| W + n \left( \bar{x}_1 \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 - \mu_2 \end{pmatrix} \right)' \right| \]

\[ = |W| \left[ 1 + n(\bar{x}_1', (\bar{x}_2 - \mu_2)') W^{-1} \left( \bar{x}_1 \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 - \mu_2 \end{pmatrix} \right) \right] \]

\[ = |W| \left[ 1 + n(\bar{x}_1', (\bar{x}_2 - \mu_2)' - \bar{x}_1' W_{11}^{-1} W_{12}) \begin{pmatrix} W_{11}^{-1} & 0 \\ 0 & W_{22-1} \end{pmatrix} \right] \]
\[ \frac{\mathbf{x}_1}{(\mathbf{x}_2 - \mu_2) - W_2W_{11}^{-1}\mathbf{x}_1} \] 

\[ = |W| \left[ 1 + n\{\mathbf{x}_1'W_{11}^{-1}\mathbf{x}_1 + [(\bar{x}_2 - \mu_2) - W_2W_{11}^{-1}\mathbf{x}_1]'W_{22}^{-1} \right] \left( [(\bar{x}_2 - \mu_2) - W_2W_{11}^{-1}\mathbf{x}_1] \right), \]

where \( W = \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix} \) and \( W_{22} = W_{22} - W_2W_{11}^{-1}W_{12} \). The maximum likelihood estimate \( \hat{\mu}_2 \) of \( \mu_2 \) is

\[ \hat{\mu}_2 = \bar{x}_2 - W_2W_{11}^{-1}\mathbf{x}_1. \]  

(6.30)

So

\[ \max_\omega L(\mu_1, \mu_2, \Sigma) = \lambda_{\max}(g)^{-\frac{1}{2}} g \left( \frac{p}{\lambda_{\max}(g)} \right) |W|^{-\frac{1}{2}} (1 + n\mathbf{x}_1'W_{11}^{-1}\mathbf{x}_1)^{-\frac{1}{2}}, \]  

(6.31)

and the likelihood ratio criterion is

\[ \frac{\max_\omega L(\mu_1, \mu_2, \Sigma)}{\max_\Omega L(\mu_1, \mu_2, \Sigma)} = \frac{|W|^{\frac{1}{2}}}{|W|^{\frac{1}{2}} (1 + n\mathbf{x}_1'W_{11}^{-1}\mathbf{x}_1)^{\frac{1}{2}}} \]  

(6.32)

\[ = \left( \frac{1}{1 + n\bar{x}_1'W_{11}^{-1}\bar{x}_1} \right)^{\frac{1}{2}}. \]

Example 2.

If \( X_{nxp} \sim LEC_{nxp}(\mu, \Sigma, \phi) \) with \( \Sigma > 0 \) and characteristic roots \( \lambda_i, i = 1, \ldots, p \), with \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0 \). We wish to test \( H_0 : \lambda_{i+1} = \cdots = \lambda_{i+k}, i + k \leq p \). Since \( \Sigma > 0 \), there exists an orthogonal matrix \( \Gamma \) such that \( \Sigma = \Gamma \Lambda \Gamma' \) where \( \Lambda \) is an diagonal matrix with the (\ell, \ell)th elements \( \lambda_\ell, \ell = 1, \ldots, p \) and the likelihood function has the form

\[ L(\mu, \Gamma, \Lambda) = \prod_{j=1}^{p} \lambda_j^{-\frac{1}{2}} g(\text{tr} \Lambda^{-1} H), \]  

(6.33)
where $H = \Gamma' \Gamma$.

Let $\Omega = \{(\mu, \Sigma) \mid \Sigma > 0\}$ and $\omega = \{(\mu, \Sigma) \mid \lambda_{i+1} = \cdots = \lambda_{i+k}\}$. It is clear that

$$\max_{\Omega} L(\mu, \Sigma) = \lambda_{\max}(g)^{-\frac{1}{2} np} \left( \frac{p}{\lambda_{\max}(g)} \right) |W|^{-\frac{1}{2} n}. \tag{6.34}$$

Since

$$\max_{\omega} L(\mu, \Sigma) = \max_{\mu, \Gamma \in \Omega_{\mu}, \Lambda \in \mathcal{M}_{(p,1,k)}} L(\mu, \Gamma, \Lambda), \tag{6.35}$$

where $\Omega_{\mu}$ is the class of $p$-dimensional orthogonal matrix, and $\mathcal{M}_{(p,1,k)} = \{\Lambda \mid 0 < \lambda_1 < \cdots < \lambda_{i+k+1} < \lambda_{i+k} = \cdots = \lambda_{i+1} < \lambda_i < \cdots < \lambda_1\}$, by Theorem 6.1,

$$\max_{\Lambda \in \mathcal{M}_{(p,1,k)}} L(\mu, \Gamma, \Lambda) = \lambda_{\max}(g)^{-\frac{1}{2} np} g \left( \frac{p}{\lambda_{\max}(g)} \right) \left[ \left( \frac{1}{k} \sum_{j=1}^{k} h_{i+j} \right) ^k \prod_{i+i+j+1 \leq i \leq \infty} h_{ij} \right]^{-\frac{1}{2} n}$$

$$= \lambda_{\max}(g)^{-\frac{1}{2} np} g \left( \frac{p}{\lambda_{\max}(g)} \right) \left[ \left( \frac{1}{k} \sum_{j=1}^{k} \gamma_{i+j} G_{\gamma_{i+j}} \right) ^k \prod_{i+i+j+1 \leq i \leq \infty} \gamma_{ij} G_{\gamma_{ij}} \right]^{-\frac{1}{2} n}, \tag{6.36}$$

where $H = (h_{ij}) = (\gamma_{ij} G_{\gamma_{ij}})$, $\Gamma = (\gamma_1, \cdots, \gamma_p)$, and $h_{pp} \leq h_{p-1} \leq \cdots \leq h_{11}$. Therefore

$$G = W + n(\bar{x} - \mu)(\bar{x} - \mu)' \tag{6.37}$$
\[ h_{jj} = \gamma'_j G \gamma_j \]
\[ = \gamma'_j (W + n(\bar{x} - \mu)(\bar{x} - \mu)^t) \gamma_j \]
\[ = \gamma'_j W \gamma_j + n \gamma'_j \gamma_j (\bar{x} - \mu)(\bar{x} - \mu)^t \gamma_j \]
\[ \geq \gamma'_j W \gamma_j, \quad (6.38) \]
and
\[ \max_{\mu} \lambda_{\max}(g)^{-\frac{1}{2}} p g \left( \frac{p}{\lambda_{\max}(g)} \right)^{-\frac{1}{2}} \left[ \left( \frac{1}{k} \sum_{j=1}^{k} \gamma'_{i+j} G \gamma_{i+j} \right) \prod_{i+j\leq p} \gamma'_{i} \gamma_{j} \right] \left( \frac{1}{k} \sum_{j=1}^{k} \gamma'_{i+j} W \gamma_{i+j} \right)^{-\frac{1}{2}} \cdot \quad (6.39) \]

imply that
\[ \lambda_{\max}(g)^{-\frac{1}{2}} p g \left( \frac{p}{\lambda_{\max}(g)} \right)^{-\frac{1}{2}} \left[ \left( \frac{1}{k} \sum_{j=1}^{k} \gamma'_{i+j} W \gamma_{i+j} \right) \prod_{i+j\leq p} \gamma'_{i} \gamma_{j} \right] \left( \frac{1}{k} \sum_{j=1}^{k} \gamma'_{i+j} W \gamma_{i+j} \right)^{-\frac{1}{2}} \cdot \quad (6.39) \]

Let \( F = \Gamma' W \Gamma = (\gamma'_i G \gamma_j) = (f_{ij}) \), \( f_{pp} \leq \cdots \leq f_{11} \), and \( 0 \leq r_p \leq \cdots \leq r_1 \) be the characteristic roots of \( W \). Then, by Theorem 9.B.1 due to Marshall and Olkin (1979), \( f \prec r \) (\( r \) majorizes \( f \)) on \( D_p = \{ x \in \mathbb{R}^p : x_1 \geq \cdots \geq x_p \} \), where \( f = (f_{11}, \cdots, f_{pp}) \) and \( r = (r_1, \cdots, r_p) \). Define
\[ \psi(x) = \left( \frac{1}{k} \sum_{j=1}^{k} x_{i+j} \right)^{\frac{1}{k}} \prod_{i+j\leq p} x_j, \quad (6.40) \]
where \( x \in \mathbb{R}^p \). Then \( \psi \) is continuous on \( D_p \) and continuously differentiable on the interior of \( D_p \). Since
\[ \psi_\ell(x) = \frac{\partial \psi(x)}{\partial x_\ell} \]

\[
\begin{align*}
\left\{ \begin{array}{l}
\left( \frac{1}{k} \sum_{j=1}^{k} x_{i+j} \right)^k \\
\prod_{i+k+1 \leq \ell \leq p \atop 1 \leq j \leq \ell} x_j \\
\left( \frac{1}{k} \sum_{j=1}^{k} x_{i+j} \right)^{k-1} \\
\prod_{i+k+1 \leq \ell \leq p \atop 1 \leq j \leq \ell} x_j \\
i + 1 \leq \ell \leq i + k
\end{array} \right.
\end{align*}
\]

is decreasing in \( \ell = 1, \ldots, p \), by Theorem 3.A.3 due to Marshall and Olkin (1979),

\[
\max_{f_{i\ell} \leq \cdots \leq f_{i1}} \lambda_{\max}(g)^{-\frac{1}{2}np} \left( \frac{p}{\lambda_{\max}(g)} \right)^{\frac{1}{2}n} \left[ \left( \frac{1}{k} \sum_{j=1}^{k} f_{i+j} \right)^k \prod_{i+k+1 \leq \ell \leq p} f_{ij} \right]^{-\frac{1}{2}n}
\]

\[
= \lambda_{\max}(g)^{-\frac{1}{2}np} \left( \frac{p}{\lambda_{\max}(g)} \right)^{\frac{1}{2}n} \left[ \left( \frac{1}{k} \sum_{j=1}^{k} r_{i+j} \right)^k \prod_{i+k+1 \leq \ell \leq p} r_j \right]^{-\frac{1}{2}n}
\]

Thus the likelihood ratio criterion is

\[
\frac{\max_\omega L(\mu, \Sigma)}{\max_\Omega L(\mu, \Sigma)} = \left( \frac{\prod_{j=1}^{p} r_j}{\left( \frac{1}{k} \sum_{j=1}^{k} r_{i+j} \right)^k \prod_{i+k+1 \leq \ell \leq p} r_j} \right)^{\frac{1}{2}n}
\]

\[
= \left( \frac{\prod_{j=1}^{k} r_{i+j}}{\left( \frac{1}{k} \sum_{j=1}^{k} r_{i+j} \right)^k} \right)^{\frac{1}{2}n}
\]
Example 5.  

**Circularly symmetric model**

If $X_{n \times p} \sim LEC_{n \times p}(\mu, \Sigma, \phi)$ with $\Sigma > 0$ and $n \geq pk$. Consider the following four hypotheses:

$$H_1: \Sigma = \text{diag}(\Sigma_0, \cdots, \Sigma_0),$$

$$H_2: \Sigma = \begin{pmatrix}
\Sigma_0 & \Sigma_1 & \cdots & \Sigma_1 & \Sigma_0 \\
\Sigma_1 & \Sigma_0 & \cdots & \Sigma_1 & \Sigma_0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\Sigma_1 & \Sigma_1 & \cdots & \Sigma_0 & \Sigma_1 \\
\Sigma_1 & \Sigma_1 & \cdots & \Sigma_1 & \Sigma_0 
\end{pmatrix},$$

$$H_3: \Sigma = \begin{pmatrix}
\Sigma_0 & \Sigma_1 & \cdots & \Sigma_{p-2} & \Sigma_{p-1} \\
\Sigma_{p-1} & \Sigma_0 & \cdots & \Sigma_{p-3} & \Sigma_{p-2} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\Sigma_2 & \Sigma_3 & \cdots & \Sigma_0 & \Sigma_1 \\
\Sigma_1 & \Sigma_2 & \cdots & \Sigma_{p-1} & \Sigma_0 
\end{pmatrix},$$

$$H_4: \Sigma > 0,$$

where $\Sigma_i, i = 0, 1, \ldots, p - 1$ are $k \times k$ symmetric and $\Sigma$ is $pk \times pk$ positive definite.

If $\Sigma$ is circularly symmetric, then

$$\Sigma = (W_0 \otimes \Sigma_0) + (W_1 \otimes \Sigma_1) + \cdots + (W_{p-1} \otimes \Sigma_{p-1}),$$

(6.44)

where $W_0 = I$, $W_j = \begin{pmatrix} 0 & I_{p-j} \\ I_j & 0 \end{pmatrix}$, $j = 1, \ldots, p - 1$, and $W_{p-1} = W_1^I$. From the theorem in Olkin (1972), there exists $\Gamma \in O_p$, with elements $r_{jk}$ given by

$$r_{jk} = p^{-\frac{1}{2}} \left\{ \sin \left( 2\pi \frac{(j-1)(k-1)}{p} \right) + \cos \left( 2\pi \frac{(j-1)(k-1)}{p} \right) \right\},$$

(6.45)

such that $(\Gamma \otimes I)\Sigma(\Gamma' \otimes I) = \text{diag}(\Psi_1, \cdots, \Psi_p) \equiv D_\Psi$ where $\Psi_j, j = 1, \ldots, p$, are $k \times k$ positive
definite. The parameter space for each of the hypotheses $H_1 - H_4$ becomes

\[ \omega_1 = \{ \mathbf{D}_\psi \mid \psi_1 = \cdots = \psi_p > 0, \text{given } \psi_j = \psi_{p-j+2}, \ j = 2, \cdots, p \}, \]

\[ \omega_2 = \{ \mathbf{D}_\psi \mid \psi_1 > 0, \psi_2 = \cdots = \psi_p > 0, \text{given } \psi_j = \psi_{p-j+2}, \ j = 2, \cdots, p \}, \]

\[ \omega_3 = \{ \mathbf{D}_\psi \mid \psi_1 > 0, \psi_j = \psi_{p-j+2} > 0, \ j = 2, \cdots, p \}, \]

\[ \omega_4 = \{ \Sigma \mid \Sigma > 0 \}. \]

Let $\mathbf{V} = (\Gamma \otimes I)\mathbf{S}(\Gamma' \otimes I) = (\mathbf{V}_{ij})$, $i = 1, \ldots, p$, $j = 1, \ldots, p$.

When $p = 2m$,

\[ (\mathbf{V}_1, \cdots, \mathbf{V}_{m+1}) = (\mathbf{V}_{11}, \mathbf{V}_{22} + \mathbf{V}_{pp}, \cdots, \mathbf{V}_{mm} + \mathbf{V}_{m+2} \mathbf{m} + 2, \mathbf{V}_{m+1} \mathbf{m} + 1), \]  

(6.46)

while $p = 2m + 1$,

\[ (\mathbf{V}_1, \cdots, \mathbf{V}_{m+1}) = (\mathbf{V}_{11}, \mathbf{V}_{22} + \mathbf{V}_{pp}, \cdots, \mathbf{V}_{m+1} \mathbf{m} + 1 + \mathbf{V}_{m+2} \mathbf{m} + 2), \]

(6.47)

where $\mathbf{V}_j = \mathbf{V}_{p-j+2}$, $j = 2, \ldots, p$, and $\mathbf{S} = (\mathbf{X} - \mathbf{1}_n \mathbf{x}')'(\mathbf{X} - \mathbf{1}_n \mathbf{x}')$.

Under $\omega_1$, $\omega_2$, $\omega_3$, the likelihood function has the form

\[ L(\mu, \psi_1, \cdots, \psi_p) = \prod_{i=1}^{p} |\psi_i|^{-\frac{1}{2}n} g(\operatorname{tr} \mathbf{D}_\psi^{-1}(\Gamma \otimes I)\mathbf{S}(\Gamma' \otimes I)), \]

(6.48)

using a similar argument as the proof of Theorem 6.1,

\[ \max_{\omega_1} L(\mu, \psi_1, \cdots, \psi_p) = \lambda_{\max}(g)^{-\frac{1}{2}npk} g\left(\frac{pk}{\lambda_{\max}(g)}\right) \frac{1}{p} \sum_{i=1}^{p} \mathbf{V}_{ii} \right|^{-\frac{1}{2}np}, \]

(6.49)
Section 6.3: General theorem

\[
\max_{\omega_3} L(\mu, \Psi_1, \ldots, \Psi_p) = \lambda_{\max}(g)^{-\frac{1}{2}npk} g \left( \frac{p k}{\lambda_{\max}(g)} \right) \left| V_{11} \right|^{-\frac{1}{2}n} \left| \frac{1}{p-1} \sum_{i=2}^{p} V_{ii} \right|^{-\frac{1}{2}n(p-1)},
\]

(6.50)

\[
\max_{\omega_3} L(\mu, \Psi_1, \ldots, \Psi_p) = \lambda_{\max}(g)^{-\frac{1}{2}npk} g \left( \frac{p k}{\lambda_{\max}(g)} \right) 2^{2k(p-m-1)} \prod_{i=1}^{p} \left| V_{ii} \right|^{-\frac{1}{2}n},
\]

(6.51)

\[
\max_{\omega_3} L(\mu, \Psi_1, \ldots, \Psi_p) = \lambda_{\max}(g)^{-\frac{1}{2}npk} g \left( \frac{p k}{\lambda_{\max}(g)} \right) \left| V \right|^{-\frac{1}{2}n}
\]

(6.52)

and the likelihood ratio criterion is same as the one in the multivariate elliptically contoured case.

6.3. General theorem.

From three examples in section 6.2, we can derive the general theorem for finding MLE. First, let’s generate the assumption for this general theorem.

Assumption B.

Let \( x_{N \times 1} \) have the density \( |\Sigma|^{-\frac{1}{2}} g((x - \mu) \Sigma^{-1}(x - \mu)) \) for \( \Sigma_{N \times N} \) positive definite and \( \mu_{N \times 1} \). Let \( \omega_0 \equiv \omega_m \times \omega_c \), where \( \omega_m \subset \mathbb{R}^N \) and \( \omega_c \) is a subset of positive definite matrices such that if \( \Sigma \in \omega_c \), then \( \alpha \Sigma \in \omega_c \) for every \( \alpha > 0 \). Suppose that the MLE’s under normality, \( \hat{\mu} \in \omega_m \) and \( \hat{\Sigma} \in \omega_c \), exist, are unique. \( \hat{\Sigma} \) is positive definite (with probability 1), and let the maximum of likelihood function be \( \hat{L} \).

Theorem 6.2.

Under the assumption B, if \( f(\lambda) \equiv \lambda^{-\frac{1}{2}N} g(\frac{1}{\lambda}) \) attains its maximum at some finite \( \lambda_{\max}(g) \), then \( \hat{\mu} = \hat{\mu} \) and \( \hat{\Sigma} = N \lambda_{\max}(g) \hat{\Sigma} \) are maximum likelihood estimates for \( g(\cdot) \) and
the corresponding maximum of likelihood function is

\[
\frac{\lambda_{\text{max}}(g)^{-\frac{1}{2}N} g \left( \frac{1}{\lambda_{\text{max}}(g)} \right)}{N^{\frac{1}{2}N} (2\pi)^{-\frac{1}{2}N}} e^{\frac{1}{2}N L}.
\]

(6.53)

Proof.

By definition, the likelihood function has the form

\[
L \equiv L(\mu, \Sigma) = |\Sigma|^{-\frac{1}{2}} g(\text{tr} \Sigma^{-1} G),
\]

(6.54)

where \( G = (x - \mu)(x - \mu)' \). Since \( \Sigma > 0 \), there exists an orthogonal matrix \( \Gamma \) such that \( \Sigma = \Gamma \Lambda \Gamma' \), where \( \Lambda \) is a diagonal matrix with diagonal elements \( \lambda_i, i = 1, \ldots, N \), and \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N > 0 \) are the characteristic roots of \( \Sigma \). Therefore

\[
L = \left[ \prod_{i=1}^{N} \lambda_i^{-\frac{1}{2}} \right] g \left( \sum_{i=1}^{N} \lambda_i^{-1} h_{ii} \right),
\]

(6.55)

where \( H = \Gamma G \Gamma' = (h_{ij}) \).

Let \( \frac{1}{\lambda} = \sum_{i=1}^{N} \lambda_i^{-1} h_{ii}, \beta_j = \lambda_j^{-1} h_{jj}, j = 1, \ldots, N - 1 \), then

\[
L = \left[ \prod_{i=1}^{N-1} \left( \frac{\beta_i}{h_{ii}} \right)^{-\frac{1}{2}} \right] \left( \frac{1 - \sum_{i=1}^{N-1} \beta_i}{h_{NN}} \right)^{-\frac{1}{2}} \lambda^{-\frac{1}{2}N} g \left( \frac{1}{\lambda} \right).
\]

(6.56)

Define \( \omega_\lambda = \{ (\Gamma, \beta_1, \ldots, \beta_{N-1}) : \Sigma \in \omega_\epsilon \} \). Let \( \Sigma^* = \alpha \Sigma, \alpha > 0 \), then \( \lambda_i^* = \alpha \lambda_i, i = 1, \ldots, N \) and \( \lambda^* = \alpha \lambda \). Therefore the range of \( \lambda \) is \( (0, \infty) \) and \( \omega_\lambda^* \) does not depend on
Section 6.3: General theorem

\[ \lambda \text{ (defined as } \omega^*) \text{. So} \]

\[
\max_{\mu \in \omega_m} \max_{\Sigma \in \omega_e} L(\mu, \Sigma) = \max_{(\Gamma, \beta_1, \ldots, \beta_{N-1}) \in \omega^*} \left[ \prod_{i=1}^{N-1} \left( \frac{\beta_i}{\hat{h}_{ii}} \right)^{-\frac{1}{2}} \right] \left( \frac{1 - \sum_{i=1}^{N-1} \beta_i}{h_{NN}} \right)^{-\frac{1}{2}} \max_{\lambda \in (0, \infty)} \lambda^{-\frac{1}{2}N} g \left( \frac{1}{\lambda} \right) \tag{6.57} \]

\[
= \lambda_{\max}(g)^{-\frac{1}{2}N} g \left( \frac{1}{\lambda_{\max}(g)} \right) \max_{\mu \in \omega_m} \left[ \prod_{i=1}^{N-1} \left( \frac{\beta_i}{\hat{h}_{ii}} \right)^{-\frac{1}{2}} \right] \left( \frac{1 - \sum_{i=1}^{N-1} \beta_i}{h_{NN}} \right)^{-\frac{1}{2}} \tag{6.58} \]

Since
\[
\max_{(\Gamma, \beta_1, \ldots, \beta_{N-1}) \in \omega^*} \left[ \prod_{i=1}^{N-1} \left( \frac{\beta_i}{\hat{h}_{ii}} \right)^{-\frac{1}{2}} \right] \left( \frac{1 - \sum_{i=1}^{N-1} \beta_i}{h_{NN}} \right)^{-\frac{1}{2}} \]
does not depend on \( g \), and under normality, \( g(x) = (2\pi)^{-\frac{1}{2}N} e^{-\frac{1}{2}x^T x} \) and \( \lambda_{\max}(g) = \frac{1}{N} \), it is clearly that \( \hat{\mu} = \tilde{\mu}, \hat{\Sigma} = N\lambda_{\max}(g)\tilde{\Sigma} \), and the corresponding maximum of likelihood function is

\[
\lambda_{\max}(g)^{-\frac{1}{2}N} g \left( \frac{1}{\lambda_{\max}(g)} \right) e^{\frac{1}{2}N} L. \tag{6.58} \]

Combining the above theorem and Lemma 6.1, we have the following corollary.

Corollary 6.2.

Under the same assumptions as Theorem 6.2, if \( g(\cdot) \) is continuous and \( E(R^2) < \infty \) where \( R \leftrightarrow \phi \in \Phi_N \), then \( \hat{\mu} = \tilde{\mu}, \hat{\Sigma} = N\lambda_{\max}(g)\tilde{\Sigma} \), and the corresponding maximum of likelihood function is (6.58)

Corollary 6.3.

Let \( \mu \in \Omega_m \subset \mathbb{R}^N \) and \( \Sigma \in \Omega_e \), satisfying the conditions of Theorem 6.2, and let the null hypothesis \( H \) be \( \mu \in \omega_m \subset \Omega_m \) and \( \Sigma \in \omega_e \subset \Omega_e \). Then the likelihood ratio criterion is independent of \( \phi \).
Proof.

The LRC is

\[
\left( \frac{\hat{L}_{\omega}}{\hat{L}_{\Omega}} \right) = \left( \frac{L_{\omega}}{L_{\Omega}} \right), \tag{6.59}
\]

from Theorem 6.2, which does not depend on \( R \) (i.e., \( \phi \)).

Remarks:

1. For every random matrix \( X \) distributed as multivariate elliptically contoured distribution, we can write it in a vector form which has an elliptically contoured distribution and a density function same as the one in Theorem 6.2. Hence Theorem 6.2 is more general than Theorem 6.1, and the three examples in section 2 also follow from Theorem 6.2.

2. Since, from (1.6),

\[
\text{Cov}(x) = E(R^2D'u^{(N)}u^{(N)'}D)
= E(R^2)D'Cov(u^{(N)})D, \tag{6.60}
\]

\( E(R^2) < \infty \) iff each element of \( \text{Cov}(x) \) is finite. So the assumptions in Corollary 6.2 — \( R \) has finite second moment and \( g(\cdot) \) is continuous — are quite reasonable assumptions.
Chapter 7

NONCENTRAL DISTRIBUTIONS OF QUADRATIC FORMS

In usual multivariate analysis, we know the distribution of $x^{(1)'}x^{(1)}$, where $x = \begin{pmatrix} x^{(1)} \\ x^{(2)} \end{pmatrix}$

$\sim N_n \begin{pmatrix} (\mu_1) \\ (\mu_2) \end{pmatrix}, I$. Anderson and Fang (1982a) derive the distribution of $x^{(1)'}x^{(1)}$, where

$x = \begin{pmatrix} x^{(1)} \\ x^{(2)} \end{pmatrix} \sim EC_n \begin{pmatrix} (\mu_1) \\ (\mu_2) \end{pmatrix}, I; \phi$, under the condition that $\mu_1 = 0$. In this chapter, we derive the distribution of $x^{(1)'}x^{(1)}$ when $\mu_1 \neq 0$.

If $x = \begin{pmatrix} x^{(1)} \\ x^{(2)} \end{pmatrix} \sim EC_n \begin{pmatrix} (\mu_1) \\ (\mu_2) \end{pmatrix}, I; \phi$, where $x^{(1)}$ is $k$-dimensional vector, then

$x \sim \mu + RU^{(n)}$, where $R (\geq 0)$ is independent of $U^{(n)} = \begin{pmatrix} U_1^{(n)} \\ U_2^{(n)} \end{pmatrix}$, and $x^{(1)} \sim \mu_1 + RU_1^{(n)}$.

From Lemma 2 of Cambanis, Huang and Simons (1981), $U_1^{(n)} \sim R_k n U^{(k)}$, where $R_k n (\geq 0)$ and $U^{(k)}$ are independent and $R_k^2 \sim \text{Beta}(\frac{k}{2}, \frac{n-k}{2})$. Thus $x^{(1)} \sim \mu_1 + RR_k n U^{(k)}$ and the distribution of $x^{(1)'}x^{(1)}$ is

\[
\text{Prob}\{x^{(1)'}x^{(1)} \leq x\} = \text{Prob}\{(RR_k n U^{(k)} + \mu_1)'(RR_k n U^{(k)} + \mu_1) \leq x\} \\
= \text{Prob}\{R^2 R_k^2 + 2RR_k n U^{(k)}'\mu_1 + \|\mu_1\|^2 \leq x\} \\
= \text{Prob}\{V + 2V^{\frac{1}{2}}U^{(k)}'\mu_1 + \|\mu_1\|^2 \leq x\}, \tag{7.1}
\]

where $V = R^2 R_k^2$. By making an orthogonal transformation, we get $U^{(k)}'\mu_1 \sim u_{1}||\mu_1||$, where $U^{(k)} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$. Since we know the distribution of $V$ and $u_{1}$ (see Anderson and Fang
\[ (1982a) , \]

\[ \text{Prob}\{x^{(1)}'x^{(1)} \leq z\} = \text{Prob}\{V + 2V^{\frac{1}{2}}u_1\|\mu_1\| + \|\mu_1\|^2 \leq z\} \]

\[ = \int_0^\infty \text{Prob}\left\{ u_1 \leq \frac{z - \|\mu_1\|^2 - v}{2v^{\frac{1}{2}}\|\mu_1\|} \right\} g(v) \, dv \]

\[ = \int_0^\infty \left[ \int_{-\infty}^z \frac{\Gamma\left(\frac{k}{2}\right)}{\Gamma\left(\frac{k-1}{2}\right)\pi^\frac{k-1}{2}} (1 - u^2)^\frac{k-1}{2} I_{|u| \leq 1} \, du \right] g(v) \, dv , \]

where \( b = \frac{z - \|\mu_1\|^2 - v}{2v^{\frac{1}{2}}\|\mu_1\|} \),

\[ g(v) = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-k}{2}\right)\Gamma\left(\frac{n-k}{2}\right)} v^{\frac{k}{2}-1} \int_0^\infty r^{-(n-2)}(r^2 - v)^{\frac{n-k}{2}-1} \, dF(r) , \]

and \( R \sim F \).

For \( x \geq 0 \) and \( v \geq 0 \),

\[ \frac{x - \|\mu_1\|^2 - v}{2v^{\frac{1}{2}}\|\mu_1\|} \leq 1 \]

\[ \iff x \leq (v^{\frac{1}{2}} + \|\mu_1\|)^2 \]

\[ \iff x^{\frac{1}{2}} - \|\mu_1\| \leq v^{\frac{1}{2}} \]

\[ \iff (x^{\frac{1}{2}} - \|\mu_1\|)^2 \leq v \quad \text{or} \quad z \leq \|\mu_1\|^2 , \]
and

$$\frac{x - \|\mu_1\|^2 - v}{2v^{\frac{1}{2}}\|\mu_1\|} \geq -1$$

$$\Leftrightarrow x \geq (v^{\frac{1}{2}} - \|\mu_1\|)^2$$

$$\Leftrightarrow x^{\frac{1}{2}} + \|\mu_1\| \geq v^{\frac{1}{2}} \quad \text{and} \quad v^{\frac{1}{2}} \geq \|\mu_1\| - x^{\frac{1}{2}}$$

$$\Leftrightarrow (x^{\frac{1}{2}} + \|\mu_1\|)^2 \geq v \quad \text{and} \quad (v \geq (\|\mu_1\| - x^{\frac{1}{2}})^2 \quad \text{or} \quad x \geq \|\mu_1\|^2).$$

Let $A = \{x - \|\mu_1\|^2 - v}{2v^{\frac{1}{2}}\|\mu_1\|} \leq 1\}, B = \{x - \|\mu_1\|^2 - v}{2v^{\frac{1}{2}}\|\mu_1\|} \geq -1\}, A_1 = \{(x^{\frac{1}{2}} - \|\mu_1\|)^2 \leq v\}$,

$B_1 = \{x \leq \|\mu_1\|^2\}$, and $C_1 = \{(x^{\frac{1}{2}} + \|\mu_1\|)^2 \geq v\}$. Then

$$A \cap B = A_1 \cap C_1.$$

We have, from (7.2),

$$\text{Prob}\{x^{(1)'x^{(1)} \leq z}\} = \int_0^\infty \left[\int_{-\infty}^6 \frac{\Gamma(\frac{b}{2})}{\Gamma(\frac{b+1}{2})\pi^{\frac{1}{2}}} (1 - u^2)^{\frac{k-1}{2}} - 1 (I_A \cap B + I_{A^*} d\nu) g(\nu) d\nu \right]$$

$$= \int_{(x^{\frac{1}{2}} + \|\mu_1\|)^2}^{(x^{\frac{1}{2}} - \|\mu_1\|)^2} \int_{-1}^6 \frac{\Gamma(\frac{b}{2})}{\Gamma(\frac{b+1}{2})\pi^{\frac{1}{2}}} (1 - u^2)^{\frac{k-1}{2}} - 1 du g(\nu) d\nu$$

$$+ I_{(x \geq \|\mu_1\|)^2} \int_0^{(x^{\frac{1}{2}} - \|\mu_1\|)^2} g(\nu) d\nu. \quad (7.5)$$
So, from Leibniz’s rule for differentiating an integral, the density of $x^{(1)}_1 x^{(1)}_2$ is

\[
\int_{(s^{1/2} - ||\mu_1||)^2}^{(s^{1/2} + ||\mu_1||)^2} \left(2v^{1/2} ||\mu_1||\right)^{-1} \frac{\Gamma(\frac{s}{2})}{\Gamma(\frac{h-1}{2} + \frac{s}{2})} (1 - k^2)^{s/2 - 1} g(v) \, dv
\]

\[
= \int_{(s^{1/2} - ||\mu_1||)^2}^{(s^{1/2} + ||\mu_1||)^2} \frac{\Gamma(\frac{s}{2})}{\Gamma(\frac{h-1}{2} + \frac{s}{2})} (2v^{1/2} ||\mu_1||)^{-1} (2v^{1/2} ||\mu_1||)^{-1} (v - (s^{1/2} - ||\mu_1||)^2) \times \left[(s^{1/2} + ||\mu_1||)^2 - v\right]^{s/2 - 1} g(v) \, dv
\]

\[
= \int_{(s^{1/2} - ||\mu_1||)^2}^{(s^{1/2} + ||\mu_1||)^2} \frac{1}{\Gamma(\frac{h-1}{2} + \frac{s}{2}) (2||\mu_1||)^{k-2} 2^{k/2}} \times \left[(v - (s^{1/2} - ||\mu_1||)^2) \left[(s^{1/2} + ||\mu_1||)^2 - v\right]^{s/2 - 1} e^{-\frac{1}{2}v} h(v) \, dv,
\]

where

\[
h(v) = \frac{\Gamma(\frac{h}{2}) 2^{k/2}}{v^{k-1} e^{-\frac{1}{2}v^2}} g(v).
\]

Let $d = z + ||\mu_1||^2$, $c = 2x^{1/2} ||\mu_1||$ and $w = \frac{v - d}{c}$. The density of $x^{(1)}_1 x^{(1)}_2$ is

\[
e^{-c^{k-2}e^{-\frac{1}{2}d^2}} \frac{\Gamma(\frac{h}{2} - \frac{1}{2}) (2||\mu_1||)^{k-2} 2^{k/2}}{\Gamma(\frac{h-1}{2} + \frac{s}{2})} \int_{-1}^{1} (1 - w^2)^{s/2} e^{-\frac{1}{2}c w h(cw + d)} \, dw
\]

\[
= \frac{2x^{k-1} e^{-\frac{1}{2}(s+||\mu_1||)^2} c^{k-2} e^{-\frac{1}{2}c w h(cw + d)} I^*}{\Gamma\left(\frac{h-1}{2}\right) \Gamma\left(\frac{k}{2}\right) 2^{k/2}} I^*,
\]

where $I^* = \int_{-1}^{1} (1 - w^2)^{s/2} e^{-\frac{1}{2}c w h(cw + d)} \, dw$. And

\[
I^* = \int_{0}^{1} (1 - w^2)^{s/2} e^{-\frac{1}{2}c w h(cw + d)} \, dw + \int_{0}^{1} (1 - w^2)^{s/2} e^{\frac{1}{2}c w h(cw + d)} \, dw
\]

\[
= \int_{0}^{1} (1 - w^2)^{s/2} \left[e^{-\frac{1}{2}c w h(cw + d)} + e^{\frac{1}{2}c w h(cw + d)}\right] \, dw.
\]
Now we make a transformation \( \omega = s^{\frac{1}{2}} \),

\[
I^* = \int_0^1 (1-s)^{\frac{k-x}{2}} \left[ e^{-\frac{1}{2}cs^{\frac{1}{2}}} h(cs^{\frac{1}{2}} + d) + e^{\frac{1}{2}cs^{\frac{1}{2}}} h(-cs^{\frac{1}{2}} + d) \right] \frac{1}{2} s^{-\frac{1}{2}} ds
\]

\[
= \frac{1}{2} \int_0^1 (1-s)^{\frac{k-x}{2}} \left[ \sum_{j=0}^{\infty} \frac{(-\frac{1}{2}cs^{\frac{1}{2}})^j}{j!} h(cs^{\frac{1}{2}} + d) + \sum_{j=0}^{\infty} \frac{(\frac{1}{2}cs^{\frac{1}{2}})^j}{j!} h(-cs^{\frac{1}{2}} + d) \right] s^{-\frac{1}{2}} ds
\]

\[
= \frac{1}{2} \int_0^1 (1-s)^{\frac{k-x}{2}} \left\{ \left[ h(cs^{\frac{1}{2}} + d) + h(-cs^{\frac{1}{2}} + d) \right] \sum_{j=0}^{\infty} \frac{(\frac{1}{2}cs^{\frac{1}{2}})^{2j}}{(2j)!} 
+ \left[ h(-cs^{\frac{1}{2}} + d) - h(cs^{\frac{1}{2}} + d) \right] \sum_{j=0}^{\infty} \frac{(\frac{1}{2}cs^{\frac{1}{2}})^{2j+1}}{(2j+1)!} \right\} s^{-\frac{1}{2}} ds
\]

(7.9)

\[
= \frac{1}{2} \Gamma(\frac{1}{2}) \int_0^1 (1-s)^{\frac{k-x}{2}} \left\{ \left[ h(cs^{\frac{1}{2}} + d) + h(-cs^{\frac{1}{2}} + d) \right] \left[ \sum_{j=0}^{\infty} \frac{(\frac{1}{2}cs^{\frac{1}{2}})^{2j}}{2^{2j} j! \Gamma(j + \frac{1}{2})} \right] s^{-\frac{1}{2}} 
+ \left[ h(-cs^{\frac{1}{2}} + d) - h(cs^{\frac{1}{2}} + d) \right] \frac{e}{2} \sum_{j=0}^{\infty} \frac{(\frac{1}{2}cs^{\frac{1}{2}})^{2j}}{2^{2j+1} j! \Gamma(j + \frac{3}{2})} \right\} ds
\]

\[
= \frac{\Gamma(\frac{1}{2})cs^{\frac{1}{2}}}{4} \int_0^1 (1-s)^{\frac{k-x}{2}} s^{-\frac{1}{2}} \left\{ \left[ h(cs^{\frac{1}{2}} + d) + h(-cs^{\frac{1}{2}} + d) \right] I_{-\frac{1}{2}} \left( \frac{c}{2} s^{\frac{1}{2}} \right) 
+ \left[ h(-cs^{\frac{1}{2}} + d) - h(cs^{\frac{1}{2}} + d) \right] I_{\frac{1}{2}} \left( \frac{c}{2} s^{\frac{1}{2}} \right) \right\} ds
\]

\[
= \frac{\Gamma(\frac{1}{2})cs^{\frac{1}{2}}}{2^{\frac{3}{2}} \mu_1^{\frac{1}{2}}} \int_0^1 (1-s)^{\frac{k-x}{2}} s^{-\frac{1}{2}} \left\{ \left[ h(cs^{\frac{1}{2}} + d) + h(-cs^{\frac{1}{2}} + d) \right] I_{-\frac{1}{2}} \left( x^{\frac{1}{2}} \mu_1 \sigma^{\frac{1}{2}} \right) 
+ \left[ h(-cs^{\frac{1}{2}} + d) - h(cs^{\frac{1}{2}} + d) \right] I_{\frac{1}{2}} \left( x^{\frac{1}{2}} \mu_1 \sigma^{\frac{1}{2}} \right) \right\} ds,
\]
where $I_m$ is the modified Bessel function of the first kind and order $m$.

Hence the density of $x^{(1)'}x^{(1)}$ is

$$
\frac{z^{\frac{2m-1}{2}} e^{-\frac{1}{2}(x+||\mu_1||^2)||\mu_1||^\frac{1}{2}}}{\Gamma\left(\frac{k-1}{2}\right)2^{\frac{k+2}{2}}} \int_0^1 (1-s)^{\frac{k-1}{2}} s^{-\frac{1}{2}} \left\{ h(cs^{\frac{1}{2}} + d) + h(-cs^{\frac{1}{2}} + d) \right\} ds \times I_{-\frac{1}{2}} \left( z^\frac{1}{2} ||\mu_1|| s^{\frac{1}{2}} \right) + \left[ h(-cs^{\frac{1}{2}} + d) - h(cs^{\frac{1}{2}} + d) \right] I_{\frac{1}{2}} \left( z^\frac{1}{2} ||\mu_1|| s^{\frac{1}{2}} \right)
$$

(7.10)

where

$$
h(v) = \frac{\Gamma\left(\frac{k}{2}\right)2^{\frac{k}{2}}}{\sqrt{\pi}^{k-1} e^{-\frac{1}{2}v}} g(v),
$$

$d = x + ||\mu_1||^2$, $e = 2z^\frac{1}{2} ||\mu_1||$, and $I_m$ is the modified Bessel function of the first kind and order $m$.

Two examples are given in the following:

Example 1. **Multivariate normal distribution.**

If $x = \begin{pmatrix} x^{(1)} \\ x^{(2)} \end{pmatrix}$ ~ $N_n \left( \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, I \right)$ and $x^{(1)}$ is $k$-dimensional vector. When $\mu_1 = 0$, the density of $x^{(1)'}x^{(1)}$ is

$$
g(v) = \frac{1}{\Gamma\left(\frac{k}{2}\right)2^{\frac{k}{2}}} v^{\frac{k-1}{2}} e^{-\frac{1}{2}v}.
$$

If $\mu_1 \neq 0$, then, from (7.10) ($h(v) = 1$), the density of $x^{(1)'}x^{(1)}$ is

$$
2z^{\frac{2m-1}{4}} e^{-\frac{1}{2}(x+||\mu_1||^2)||\mu_1||^\frac{1}{2}} \int_0^1 (1-s)^{\frac{k-3}{2}} s^{-\frac{1}{2}} I_{-\frac{1}{2}} \left( z^\frac{1}{2} ||\mu_1|| s^{\frac{1}{2}} \right) ds 
$$

(7.11)

$$
= \frac{1}{2} \left( \frac{z}{||\mu_1||^2} \right)^\frac{1}{2}(k-2) I_{\frac{1}{2}}(k-2) \left( z^\frac{1}{2} ||\mu_1|| \right) \exp \left[ -\frac{1}{2}(||\mu_1||^2 + z) \right],
$$

where $I_m$ is the modified Bessel function of the first kind and order $m$. 
Example 2. Multivariate \( t \)-distribution.

From Anderson and Fang (1982a), if \( y_n = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \) has \( n \)-dimensional multivariate \( t \)-distribution with \( \ell \) degree of freedom, where \( y_1 \) is \( k \)-dimensional vector and \( \mu_1 = 0 \), then \( \frac{1}{2} y_1'y_1 \) has \( F \)-distribution with degrees of freedom \( k \) and \( \ell \) and the density of \( y_1'y_1 \) is

\[
g(v) = \frac{\Gamma\left(\frac{k+\ell}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{\ell}{2}\right)} \ell^{-\frac{k}{2}} v^{-\frac{k-2}{2}} \left(1 + \frac{v}{\ell}\right)^{-\frac{k+\ell}{2}}.
\]

If \( \mu_1 \neq 0 \), then, from (7.7), the density of \( y_1'y_1 \) is

\[
\frac{e^{k-2}}{\Gamma\left(\frac{k-1}{2}\right)\Gamma\left(\frac{\ell}{2}\right)\Gamma\left(\frac{1}{2}\right)} \int_{-1}^{1} (1 - w^2)^{-\frac{k-2}{2}} \frac{\Gamma\left(\frac{k+\ell}{2}\right)}{\Gamma\left(\frac{\ell}{2}\right)} \left(1 + \frac{c w + d + \ell}{\ell}\right)^{-\frac{k+\ell}{2}} dw
\]

\[
= \frac{x^{\frac{k}{2}-1} \Gamma\left(\frac{k+\ell}{2}\right) \ell^{\frac{1}{2}}}{\Gamma\left(\frac{k-1}{2}\right)\Gamma\left(\frac{\ell}{2}\right)\Gamma\left(\frac{1}{2}\right)} \int_{0}^{1} (1 - s)^{-\frac{k-2}{2}} \left[(c s^{-\frac{1}{2}} + d + \ell)^{-\frac{k+\ell}{2}} + (-c s^{-\frac{1}{2}} + d + \ell)^{-\frac{k+\ell}{2}}\right] \frac{1}{2} s^{-\frac{1}{2}} ds
\]

\[
= \frac{x^{\frac{k}{2}-1} \Gamma\left(\frac{k+\ell}{2}\right) \ell^{\frac{1}{2}}}{\Gamma\left(\frac{k-1}{2}\right)\Gamma\left(\frac{\ell}{2}\right)\Gamma\left(\frac{1}{2}\right)} \sum_{j=0}^{\infty} \frac{\Gamma\left(\frac{k+\ell}{2} + 2j\right)}{\Gamma\left(\frac{k+\ell}{2}\right)(2j)!} (d + \ell)^{-\frac{k+\ell}{2}-2j} \int_{0}^{1} (1 - s)^{-\frac{k-2}{2}} (c s^{-\frac{1}{2}})^{2j} s^{-\frac{1}{2}} ds
\]

\[
= \frac{x^{\frac{k}{2}-1} \Gamma\left(\frac{k+\ell}{2}\right) \ell^{\frac{1}{2}}}{\Gamma\left(\frac{k-1}{2}\right)\Gamma\left(\frac{\ell}{2}\right)\Gamma\left(\frac{1}{2}\right)} \sum_{j=0}^{\infty} \frac{\Gamma\left(\frac{k+\ell}{2} + 2j\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{k+\ell}{2}\right)2^{2j} j!\Gamma\left(j + \frac{1}{2}\right)} (d + \ell)^{-\frac{k+\ell}{2}-2j} e^{2j} \frac{\Gamma\left(j + \frac{1}{2}\right)\Gamma\left(\frac{k-1}{2}\right)}{\Gamma\left(\frac{k}{2} + j\right)}
\]

\[
= \frac{x^{\frac{k}{2}-1} \ell^{\frac{1}{2}}}{\Gamma\left(\frac{1}{2}\right)} \sum_{j=0}^{\infty} \frac{\Gamma\left(\frac{k+\ell}{2} + 2j\right) \mu_1^{2j}}{j! (x + \mu_1^2 + \ell)^{\frac{k+\ell}{2}+2j} \Gamma\left(\frac{k}{2} + j\right)} (z + \mu_1^2 + \ell)^{-\frac{k+\ell}{2}+2j} \Gamma\left(\frac{k}{2} + j\right)
\]

(7.12)
References


References


TECHNICAL REPORTS
U.S. ARMY RESEARCH OFFICE - CONTRACT DAAG29-82-K-0156


"Invariant Tests and Likelihood Ratio Tests for Multivariate Elliptically Contoured Distributions"

T. W. Anderson and Huang Hsu

Department of Statistics - Sequoia Hall
Stanford University
Stanford, CA 94305

U. S. Army Research Office
Post Office Box 12211
Research Triangle Park, NC 27709

Approved for public release; distribution unlimited.

Invariant test, likelihood ratio test, multivariate elliptically contoured distribution, admissibility locally most powerful invariant test, maximal invariant, maximum likelihood estimation.
20. Abstract

The usual assumption in multivariate hypothesis testing is that the sample consists of \( n \) independent, identically distributed Gaussian \( p \)-vectors. In this dissertation this assumption is weakened by considering a class of distributions for which the vector observation are not necessarily Gaussian or independent. This class consists of the elliptically symmetric laws with densities of the form \( f(\mathbf{X}_{n \times p}) = f(\text{tr}(\mathbf{X}-\mathbf{M})'(\mathbf{X}-\mathbf{M})\Sigma^{-1}) \).

The following hypothesis testing problems are considered: testing for equality between the mean vector and a specified vector, lack of correlations among different sets, equality of covariance matrices and mean vectors, equality between the correlation coefficient and a specified number, and MANOVA. For each of the above hypotheses, invariant tests and their properties are developed. These include the uniformly most powerful test, the locally most powerful test, admissibility, and null and non-null distributions. Further, under the assumptions that \( g(\cdot) \) is continuous, each element of the covariance matrix of \( \mathbf{X} \) is finite, and the null hypothesis is scalar-invariant, it is shown that the usual normal-theory likelihood ratio tests are exactly robust for the null case under this wider class (i.e. the likelihood ratio tests, sampling from this general class, are the same as the usual normal-theory likelihood ratio tests and their null distributions are the same.)