Elliptically Contoured Models in Statistics

Mathematics and Its Applications

Volume 240

Elliptically Contoured Models in Statistics

by

A. K. Gupta

Department of Mathematics and Statistics,
Bowling Green State University,
Bowling Green, Ohio, U.S.A.

and

T. Varga

AB-AEGON General Insurance Company,
Budapest, Hungary

SPRINGER SCIENCE+BUSINESS MEDIA, B.V.

Library of Congress Cataloging-in-Publication Data

Gupta, A. K. (Arjun K.), 1938-
Elliptically contoured models in statistics / by A.K. Gupta and T. Varga.
p. cm. -- (Mathematics and its applications ; v. 240)
Includes bibliographical references and indexes.
ISBN 978-94-010-4719-7 ISBN 978-94-011-1646-6 (eBook)
DOI 10.1007/978-94-011-1646-6
1. Distribution (Probability theory) 2. Multivariate analysis. I. Varga, T. (Tamas), 1960- . II. Title. III. Series: Mathematics and its applications (Kluwer Academic Publishers) ; v. 240.
QA273.6.G87 1993
519.2'4--dc20 92-44825

ISBN 978-94-010-4719-7

Printed on acid-free paper

Dedicated to the memory of my mother and father. AKG

TABLE OF CONTENTS

PREFACE

In multivariate statistical analysis, elliptical distributions have recently provided an alternative to the normal model. Most of the work, however, is spread out in journals throughout the world and is not easily accessible to the investigators. Fang, Kotz, and Ng presented a systematic study of multivariate elliptical distributions, however, they did not discuss the matrix variate case. Recently Fang and Zhang have summarized the results of generalized multivariate analysis which include vector as well as the matrix variate distributions. On the other hand, Fang and Anderson collected research papers on matrix variate elliptical distributions, many of them published for the first time in English. They published very rich material on the topic, but the results are given in paper form which does not provide a unified treatment of the theory. Therefore, it seemed appropriate to collect the most important results on the theory of matrix variate elliptically contoured distributions available in the literature and organize them in a unified manner that can serve as an introduction to the subject.

The book will be useful for researchers, teachers, and graduate students in statistics and related fields whose interests involve multivariate statistical analysis. Parts of this book were presented by Arjun K.Gupta as a one semester course at Bowling Green State University. Some new results have also been included which generalize the results in Fang and Zhang. Knowledge of matrix algebra and statistics at the level of Anderson is assumed. However, Chapter 1 summarizes some results of matrix algebra. This chapter also contains a brief review of the literature and a list of mathematical symbols used in the book.

Chapter 2 gives the basic properties of the matrix variate elliptically contoured distributions, such as the probability density function and expected values. It also presents one of the most important tools of the theory of elliptical distributions, the stochastic representation.

The probability density function and expected values are investigated in detail in Chapter 3.

Chapter 4 focuses on elliptically contoured distributions that can be represented as mixtures of normal distributions.

The distributions of functions of random matrices with elliptically contoured distributions are discussed in Chapter 5. Special attention is given to quadratic forms.

Characterization results are given in Chapter 6.

The last three chapters are devoted to statistical inference. Chapter 7 focuses on estimation results, whereas Chapter 8 is concerned with hypothesis testing problems. Inference for linear models is studied in Chapter 9. Finally, an up to date bibliography has been provided, along with author and subject indexes.

We would like to thank the Department of Mathematics and Statistics, Bowling Green State University, for supporting our endeavour and for providing the necessary facilities to accomplish the task. The first author is thankful to the Biostatistics Department, University of Michigan, for providing him the opportunity to organize the material in its final form. Thanks are also due to Professors A. M. Kshirsagar, D. K. Nagar, M. Siotani, and J. Tang for many helpful discussions. He would also like to acknowledge his wife, Meera, and his children, Alka, Mita, and Nisha for their support throughout the writing of the book.

Particular gratitude is rendered to Cynthia Patterson for her efficient technical word processing. We also express our appreciation to the publisher and especially to Dr. D. J. Larner for his sustained interest in the project.

September 1992

A. K. Gupta
T. Varga

CHAPTER 1

PRELIMINARIES

1.1. INTRODUCTION AND LITERATURE REVIEW

Matrix variate distributions have been studied by statisticians for a long time. The first results on this topic were published by Hsu and Wishart. These distributions proved to be useful in statistical inference. For example, the Wishart distribution is essential when studying the sample covariance matrix in the multivariate normal theory. Random matrices can also be used to describe repeated measurements on multivariate variables. In this case, the assumption of the independence of the observations, a commonly used condition in statistical analysis, is often not feasible. When analyzing data sets like these, the matrix variate elliptically contoured distributions can be used to describe the dependence structure of the data. This is a rich class of distributions containing the matrix variate normal, contaminated normal, Cauchy and Student's t-distributions. The fact that the distributions in this class possess certain properties, similar to those of the normal distribution, makes them especially useful. For example, many testing procedures developed for the normal theory to test various hypotheses can be used for this class of distributions, too.

Matrix variate elliptically contoured distributions represent an extension of the concept of elliptical distributions from the vector to the matrix case. Important distribution results on vector variate elliptical distributions were derived by Kelker (1970), Chu (1973), Dawid (1977) and Cambanis, Huang, and Simons (1981). Quadratic forms in elliptical distributions were studied by Cacoullos and Koutras (1984), Fang and Wu (1984), Anderson and Fang (1987), and Smith (1989). Problems related to moments were considered by Berkane and Bentler (1986a). Characterization results were given by Kingman (1972), Khatri and Mukerjee (1987), and Berkane and Bentler (1986b). Kariya (1981), Kuritsyn (1986), Anderson, Fang, and Hsu (1986), Jajuga (1987), Cellier, Fourdrinier, and Robert (1989) and Grübel and Rocke (1989) focused on inference problems. Asymptotic results were obtained by Browne (1984), Hayakawa (1987), Khatri (1988) and Mitchell (1989). Special aspects of elliptical distributions were discussed by Khatri (1980), Sampson (1983), Mitchell and Krzanowski (1985), Cacoullos and Koutras (1985), Khattree and Peddada (1987), and Cléroux and Ducharme (1989). Krishnaiah and Lin (1986) introduced the concept of complex elliptical distributions. Some of the early results in elliptical distributions were summarized in Muirhead (1982) and Johnson (1987). More extensive reviews of papers on this topic were provided by Chmielewski (1981), and Bentler and Berkane (1985). The most recent summary of distribution results was given by Fang, Kotz, and Ng (1990).

Some of the papers mentioned above also contain results on matrix variate elliptically contoured distributions; for example, Anderson and Fang (1987), and Anderson, Fang, and Hsu (1986). Other papers, like Chmielewski (1980), Richards (1984), Khatri (1987), and Sutradhar and Ali (1989), are also concerned with matrix variate elliptical distributions. Fang and Anderson (1990) is a collection of papers on matrix variate elliptical distributions. Many of these papers were originally published in Chinese journals and this is their first publication in English. Recently Fang and Zhang (1990) have provided an excellent account of spherical and related distributions.

The purpose of the present book is to provide a unified treatment of the theory of matrix variate elliptically contoured distributions, to present the most important results on the topic published in various papers and books, and to give their proofs.

1.2. NOTATION

We denote matrices by capital letters, vectors by small bold letters and scalars by small letters. We use the same notation for a random variable and its values. Also the following notations will be used in the sequel.

$\mathbb{R}^p$	: the p-dimensional real space.
$\mathcal{B}(\mathbb{R}^p)$	: the Borel sets in $\mathbb{R}^p$.
S_p	: the unit sphere in $\mathbb{R}^p$.
$\mathbb{R}^+$	: the set of positive real numbers.
$\mathbb{R}_0^+$	: the set of nonnegative real numbers.
$\chi_A(x)$	: the indicator function of A, that is $\chi_A(x) = 1$ if $x \in A$ and $\chi_A(x) = 0$ if $x \notin A$.
$\chi(x \geq t)$	: the same as $\chi_{[t,\infty)}(x)$ (t is a real number).
$A \in \mathbb{R}^{p \times n}$	: A is a $p \times n$ dimensional real matrix.
a_{ij}	: the $(i,j)^{th}$ element of the matrix A.
A'	: transpose of A.
$rk(A)$	: rank of A.
$A > O$	: the square matrix A is positive definite (see also Section 1.3).
$A \geq O$	: the square matrix A is positive semidefinite (see also Section 1.3).
$\|A\|$	: determinant of the square matrix A.
$tr(A)$	: trace of the square matrix A.
$etr(A)$	: $\exp\{tr(A)\}$ if A is a square matrix.
$\lVert A \rVert$	: norm of A defined by $\lVert A \rVert = (tr(A'A))^{\frac{1}{2}}$.
A^-	: generalized inverse of A, that is $AA^-A = A$ (see also Section 1.3).
$A^{\frac{1}{2}}$	: let the spectral decomposition of $A \geq 0$ be GDG', and define $A^{\frac{1}{2}} = GD^{\frac{1}{2}}G'$ (see also Section 1.3).
$O(p)$	: the set of $p \times p$ dimensional orthogonal matrices.
I_p	: the $p \times p$ dimensional identity matrix.
$\mathbf{e}_p$	: the p-dimensional vector whose elements are 1's; that is, $\mathbf{e}_p = (1,1,\dots,1)'$.
$A \otimes B$	: Kronecker product of the matrices A and B (see also Section 1.3).
$A \geq B$	: $A - B$ is positive semidefinite.
$A > B$	: $A - B$ is positive definite.

vec(A)	:	the vector $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ where $\mathbf{a}_i$ denotes the i^{th} column of the $p \times n$ matrix A, $i = 1,2,\dots,n$.
$J(X \to f(X))$	:	the Jacobian of the matrix transformation f.
$X \sim \mathcal{D}$	:	the random matrix X is distributed according to the distribution $\mathcal{D}$.
$X \approx Y$	:	the random matrices X and Y are identically distributed.
Cov(X)	:	covariance matrix of the random matrix X; that is Cov(X) = Cov(vec(X')).
$\phi_X(T)$	:	the characteristic function of the random matrix X at T; that is $\mathcal{E}(\mathrm{etr}(iT'X))$, $X,T \in \mathbb{R}^{p\times n}$.

For a review of Jacobians, see Press (1972) and Siotani, Hayakawa and Fujikoshi (1985). We also use the following notations for some well known probability distributions.

UNIVARIATE DISTRIBUTIONS:

$N(\mu,\sigma^2)$: normal distribution; its probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\},$$

where $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$, and $x \in \mathbb{R}$.

$B(a,b)$: beta distribution; its probability density function is

$$f(x) = \frac{1}{\beta(a,b)} x^{a-1}(1-x)^{b-1},$$

where $a > 0$, $b > 0$, $\beta(a,b) = \dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ and $0 < x < 1$.

t_n : Student's t-distribution; its probability density function is
$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \frac{1}{\left(1+\frac{x^2}{n}\right)^{\frac{n+1}{2}}},$$
where $n > 0$, and $x \in \mathbb{R}$.

χ_n^2 : chi-square distribution; its probability density function is
$$f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} \exp\left\{-\frac{x}{2}\right\},$$
where $n > 0$, and $x \geq 0$.

χ_n : chi distribution; its probability density function is
$$f(x) = \frac{1}{2^{\frac{n}{2}-1}\Gamma\left(\frac{n}{2}\right)} x^{n-1} \exp\left\{-\frac{x^2}{2}\right\},$$
where $n > 0$, and $x \geq 0$.

$F_{n,m}$: F distribution; its probability density function is
$$f(x) = \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)} \left(\frac{n}{m}\right)^{\frac{n}{2}} \frac{x^{\frac{n}{2}-1}}{\left(1+\frac{n}{m}x\right)^{\frac{n+m}{2}}},$$
where $n,m = 1,2,\ldots,$ and $x > 0$.

$U_{p,m,n}$: U distribution, which is the same as the distribution of $\prod_{i=1}^{p} v_i$; where v_i's are independent and
$$v_i \sim B\left(\frac{n+1-i}{2}, \frac{m}{2}\right).$$

For the U distribution, see Anderson (1984), pp. 301-304.

MULTIVARIATE DISTRIBUTIONS:

$N(\mu,\Sigma)$: multivariate normal distribution; its characteristic function is

$$\phi_{\mathbf{x}}(\mathbf{t}) = \exp\left\{i\mathbf{t}'\mu + \frac{1}{2}\mathbf{t}'\Sigma\mathbf{t}'\right\},$$

where $\mathbf{x},\mathbf{t},\mu \in \mathbb{R}^p$, $\Sigma \in \mathbb{R}^{p\times p}$, and $\Sigma \geq O$.

$D(m_1,\ldots,m_p;m_{p+1})$: Dirichlet distribution; its probability density function is

$$f(\mathbf{x}) = \frac{\Gamma\left(\sum_{i=1}^{p+1} m_i\right)}{\prod_{i=1}^{p+1} \Gamma(m_i)} \cdot \prod_{i=1}^{p} x_i^{m_i-1}\left(1 - \sum_{i=1}^{p} x_i\right)^{m_{p+1}-1}$$

where $\mathbf{x} = (x_1,x_2,\ldots,x_p)' \in \mathbb{R}^p$, $0 < \sum_{i=1}^{p} x_i < 1$, and $m_i > 0$, $i = 1,2,\ldots,p$.

MATRIX VARIATE DISTRIBUTIONS:

$N_{p,n}(M,\Sigma \otimes \Phi)$: matrix variate normal distribution; its characteristic function is

$$\phi_X(T) = \mathrm{etr}\left(iT'M + \frac{1}{2}T'\Sigma T\Phi\right),$$

where $M,X,T \in \mathbb{R}^{p\times n}$, $\Sigma \in \mathbb{R}^{p\times p}$, $\Sigma \geq O$, $\Phi \in \mathbb{R}^{n\times n}$, and $\Phi \geq O$.

$W_p(\Sigma,n)$: Wishart distribution; its probability density function is

$$f(X) = \frac{|X|^{\frac{n-p-1}{2}}\,\mathrm{etr}\left(-\frac{1}{2}\Sigma^{-1}X\right)}{2^{\frac{np}{2}}\,|\Sigma|^{\frac{n}{2}}\,\Gamma_p\left(\frac{n}{2}\right)},$$

where $X \in \mathbb{R}^{p\times p}$, $X > O$, $\Sigma \in \mathbb{R}^{p\times p}$, $\Sigma > O$, p,n are integers, $n \geq p$, and

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

$B_p^I(a,b)$: matrix variate beta distribution of type I; its probability density function is

$$f(X) = \frac{|X|^{a-\frac{p+1}{2}} |I_p - X|^{b-\frac{p+1}{2}}}{\beta_p(a,b)},$$

where $a > \frac{p-1}{2}$, $b > \frac{p-1}{2}$,

$$\beta_p(a,b) = \frac{\Gamma_p(a)\Gamma_p(b)}{\Gamma_p(a+b)}, \quad X \in \mathbb{R}^{p\times p} \text{ and } O < X < I_p.$$

$B_p^{II}(a,b)$: matrix variate beta distribution of type II; its probability density function is

$$f(X) = \frac{|X|^{a-\frac{p+1}{2}} |I_p + X|^{-(a+b)}}{\beta_p(a,b)}$$

where $a > \frac{p-1}{2}$, $b > \frac{p-1}{2}$, $X \in \mathbb{R}^{p\times p}$ and $X > O$.

$T_{p,n}(m,M,\Sigma,\Phi)$: matrix variate T distribution; its probability density function is

$$f(X) = \frac{\pi^{-\frac{np}{2}} \Gamma_p\left(\frac{n+m+p-1}{2}\right)}{\Gamma_p\left(\frac{m+p-1}{2}\right) |\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}}$$

$$\cdot\, |I_p + \Phi^{-1}(X-M)\Sigma^{-1}(X-M)'|^{-\frac{n+m+p-1}{2}}$$

where $m > 0$, $X,T,M \in \mathbb{R}^{p\times n}$, $\Sigma \in \mathbb{R}^{p\times p}$, $\Sigma > O$, $\Phi \in \mathbb{R}^{n\times n}$, and $\Phi > O$.

For further discussion of $B_p^I(a,b)$, $B_p^{II}(a,b)$, see Olkin and Rubin (1964) and Javier and Gupta (1985b), and for results on $T_{p,n}(m,M,\Sigma,\Phi)$, see Dickey (1967).

1.3. SOME RESULTS FROM MATRIX ALGEBRA

In this section, we give some results from matrix algebra which are used in the subsequent chapters. Except for the results on the generalized inverse, we do not prove the theorems since they can be found in any book of linear algebra (e.g. Magnus and Neudecker, 1988). Other books, like Anderson (1984) and Muirhead (1982), discuss these results in the appendices of their books.

DEFINITION 1.3.1 Let A be a $p \times p$ matrix. Then, A is called

i) symmetric if $A' = A$.

ii) idempotent if $A^2 = A$.

iii) nonsingular if $|A| \neq 0$.

iv) orthogonal if $AA' = A'A = I_p$.

v) positive semidefinite and this is denoted by $A \geq O$ if A is symmetric and for every p-dimensional vector $\mathbf{v}$, $\mathbf{v}'A\mathbf{v} \geq 0$.

vi) positive definite and this is denoted by $A > O$ if A is symmetric and for every p-dimensional nonzero vector $\mathbf{v}$, $\mathbf{v}'A\mathbf{v} > 0$.

vii) permutation matrix if in each row and each column of A exactly one element is 1 and all the others are 0.

viii) signed permutation matrix if in each row and each column of A exactly one element is 1 or -1 and all the others are 0.

THEOREM 1.3.1. *Let A be $p \times p$, and B be $q \times p$ matrices. Then, we have the following results.*

i) If $A > O$, then $A^{-1} > O$.

ii) If $A \geq O$, then $BAB' \geq O$.

iii) If $q \leq p$, $A > O$ and $\mathrm{rk}(B) = q$, then $BAB' > O$.

DEFINITION 1.3.2. Let A be a $p \times p$ matrix. Then, the roots (with multiplicity) of the equation

$$|A - \lambda I_p| = 0$$

are called the characteristic roots of A.

THEOREM 1.3.2. *Let* A *be a* $p \times p$ *matrix and* $\lambda_1, \lambda_2, \ldots, \lambda_p$ *its characteristic roots. Then,*

i) $|A| = \prod_{i=1}^{p} \lambda_i.$

ii) $\mathrm{tr}(A) = \sum_{i=1}^{p} \lambda_i.$

iii) $\mathrm{rk}(A)$ = *the number of nonzero characteristic roots.*

iv) A *is nonsingular if and only if the characteristic roots are nonzero.*

v) *Further, if we assume that* A *is symmetric, then the characteristic roots of* A *are real.*

vi) A *is positive semidefinite if and only if the characteristic roots of* A *are nonnegative.*

vii) A *is positive definite if and only if the characteristic roots of* A *are positive.*

The next theorem gives results on the rank of matrices.

THEOREM 1.3.3. *i) Let* A *be a* $p \times q$ *matrix. Then,* $\mathrm{rk}(A) \le \min(p,q)$ *and* $\mathrm{rk}(A) = \mathrm{rk}(A') = \mathrm{rk}(AA') = \mathrm{rk}(A'A)$. *If* $p = q$, *then* $\mathrm{rk}(A) = p$ *if and only if* A *is nonsingular.*

ii) *Let* A *and* B *be* $p \times q$ *matrices. Then,* $\mathrm{rk}(A + B) \le \mathrm{rk}(A) + \mathrm{rk}(B)$.

iii) *Let* A *be a* $p \times q$, B *a* $q \times r$ *matrix. Then,* $\mathrm{rk}(AB) \le \min(\mathrm{rk}(A),\mathrm{rk}(B))$. *If* $p = q$ *and* A *is nonsingular then* $\mathrm{rk}(AB) = \mathrm{rk}(B)$.

DEFINITION 1.3.3. Let A be a $p \times q$ matrix. If $\mathrm{rk}(A) = \min(p,q)$, then A is called a full rank matrix.

In the following theorem we list some of the properties of the trace function.

THEOREM 1.3.4. *i) Let* A *be a* $p \times p$ *matrix Then,* $\mathrm{tr}(A) = \mathrm{tr}(A')$, *and* $\mathrm{tr}(cA) = c\,\mathrm{tr}(A)$ *where* c *is a scalar.*

ii) *Let* A *and* B *be* $p \times q$ *matrices. Then,* $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$.

iii) *Let* A *be a* $p \times q$, and B *a* $q \times p$ *matrix. Then,* $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

THEOREM 1.3.5. *Let the* $p \times p$ *matrix* A *be defined by*

$$a_{ij} = \begin{cases} x, & \text{if } i = j \\ y, & \text{if } i \neq j \end{cases} .$$

Then, $|A| = (x - y)^{p-1}(x + (p - 1)y)$.

Now we give some matrix factorization theorems.

THEOREM 1.3.6. (Singular value decomposition of a matrix) *Let* A *be a* $p \times q$ *matrix with* $p \geq q$. *Then, there exist a* $p \times p$ *orthogonal matrix* G, *a* $q \times q$ *orthogonal matrix* H *and a* $q \times q$ *positive semidefinite diagonal matrix* D *such that*

$$A = G\begin{pmatrix} D \\ O \end{pmatrix}H,$$

where O *denotes the* $(p - q) \times q$ *zero matrix. Moreover,* $\mathrm{rk}(D) = \mathrm{rk}(A)$.

THEOREM 1.3.7. (Spectral decomposition of a symmetric matrix) *Let* A *be a* $p \times p$ *symmetric matrix. Then, there exist a* $p \times p$ *orthogonal matrix* G *and a* $p \times p$ *diagonal matrix* D *such that*

$$A = GDG'. \tag{1.1}$$

Moreover, if A *is of the form (1.1) then the diagonal elements of* D *are the characteristic roots of* A.

DEFINITION 1.3.4. Let A be a $p \times p$ positive semidefinite matrix with spectral decomposition $A = GDG'$. Let $D^{\frac{1}{2}}$ be the diagonal matrix whose elements are the square roots of the elements of D. Then we define $A^{\frac{1}{2}}$ as $GD^{\frac{1}{2}}G'$.

THEOREM 1.3.8. *Let* A *and* B *be* $p \times p$ *matrices. Assume* A *is positive definite and* B *is positive semidefinite. Then, there exist a* $p \times p$ *nonsingular matrix* C *and a* $p \times p$ *diagonal matrix* D *such that*

$$A = CC' \text{ and } B = CDC'. \tag{1.2}$$

Moreover, if A *and* B *are of the form (1.2), then the diagonal elements of* D *are the roots of the equation* $|B - \lambda A| = 0$.

THEOREM 1.3.9. *Let* A *be a* $p \times q$ *matrix with* $\text{rk}(A) = q$. *Then there exist a* $p \times p$ *orthogonal matrix* G *and a* $q \times q$ *positive definite matrix* B *such that*

$$A = G \begin{bmatrix} I_q \\ O \end{bmatrix} B,$$

where O *denotes the* $(p - q) \times q$ *zero matrix.*

THEOREM 1.3.10. (The rank factorization of a square matrix) *Let* A *be a* $p \times p$ *matrix with* $\text{rk}(A) = q$. *Then,* $AA' = BB'$ *if and only if there exists a* $p \times q$ *matrix of rank* q *such that* $A = BB'$.

THEOREM 1.3.11. (Vinograd's Theorem) *Assume* A *is a* $p \times q$, *and* B *is a* $p \times r$ *matrix, where* $q \leq r$. *Then,* $AA' = BB'$ *if and only if there exists a* $q \times r$ *matrix* H *with* $HH' = I_q$ *such that* $B = AH$.

THEOREM 1.3.12. *Let* A *be a* $p \times p$, *symmetric idempotent matrix of rank* q. *Then, there exists a* $p \times p$ *orthogonal matrix* G *such that*

$$A = G\begin{pmatrix} I_q & O \\ O & O \end{pmatrix}G',$$

where the O's *denote zero matrices of appropriate dimensions.*

THEOREM 1.3.13. *Let* $A_1, A_2, \ldots, A_n$ *be* $p \times p$, *symmetric, idempotent matrices. Then, there exists a* $p \times p$ *orthogonal matrix such that* $G'A_iG$ *is diagonal for every* $1 \le i \le n$ *if and only if* $A_iA_j = A_jA_i$ *for every* $1 \le i,j \le n$.

THEOREM 1.3.14. *Let* $A_1, A_2, \ldots, A_n$ *be* $p \times p$, *symmetric idempotent matrices. Then, there exists a* $p \times p$ *orthogonal matrix* G *such that*

$$G'A_1G = \begin{pmatrix} I_{r_1} & O \\ O & O \end{pmatrix}, \quad G'A_2G = \begin{pmatrix} O & O & O \\ O & I_{r_2} & O \\ O & O & O \end{pmatrix},$$

$$\ldots, G'A_nG = \begin{pmatrix} O & O & O \\ O & I_{r_n} & O \\ O & O & O \end{pmatrix},$$

where $r_i = \mathrm{rk}(A_i)$, $i = 1,\ldots,n$, *if and only if* $A_iA_j = O$ *for every* $i \ne j$.

Next, we give some results for the Kronecker product, also called direct product of matrices.

DEFINITION 1.3.5. *Let* $A = (a_{ij})$ *be a* $p \times q$, B *an* $r \times s$ *matrix. Then the Kronecker product of* A *and* B, *denoted by* $A \otimes B$, *is the* $(pr) \times (qs)$ *matrix defined by*

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \ldots & a_{1q}B \\ a_{21}B & a_{22}B & \ldots & a_{2q}B \\ \vdots & & & \\ a_{p1}B & a_{p2}B & \ldots & a_{pq}B \end{pmatrix}.$$

THEOREM 1.3.15. *i) If* c *and* d *are scalars, then* $(cA) \otimes (dB) = cd(A \otimes B)$.

ii) *If* A *and* B *are of equal dimension, then* $(A + B) \otimes C = (A \otimes C) + (B \otimes C)$, *and* $C \otimes (A + B) = (C \otimes A) + (C \otimes B)$.

iii) $(A \otimes B) \otimes C = A \otimes (B \otimes C)$.

iv) $(A \otimes B)' = A' \otimes B'$.

v) *If* A *and* B *are square matrices then* $\text{tr}(A \otimes B) = \text{tr}(A)\,\text{tr}(B)$.

vi) *If* A *is* $p \times q$, B *is* $r \times s$, C *is* $q \times n$, *and* D *is* $s \times v$, *then* $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$.

vii) *If* A *and* B *are nonsingular matrices, then* $A \otimes B$ *is also nonsingular and* $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.

viii) *If* A *and* B *are orthogonal matrices, then* $A \otimes B$ *is also orthogonal.*

ix) *If* A *and* B *are positive semidefinite matrices, then* $A \otimes B$ *is also positive semidefinite.*

x) *If* A *and* B *are positive definite matrices, then* $A \otimes B$ *is also positive definite.*

xi) *If* A *is* $p \times p$, *and* B *is* $q \times q$ *matrix, then* $|A \otimes B| = |A|^q\,|B|^p$.

xii) *If* A *is* $p \times p$, B *is* $q \times q$, $\lambda_1, \lambda_2, \ldots, \lambda_p$ *are the characteristic roots of* A, *and* $\mu_1, \mu_2, \ldots, \mu_q$ *are the characteristic roots of* B, *then* $\lambda_i \mu_j$, $i = 1, \ldots, p$, $j = 1, \ldots, q$ *are the characteristic roots of* $A \otimes B$.

THEOREM 1.3.16. *Let* A_1 *and* A_2 *be* $p \times q$, B_1 *and* B_2 *be* $r \times s$ *nonzero matrices. Then,* $A_1 \otimes B_1 = A_2 \otimes B_2$, *if and only if there exists a nonzero real number* c *such that* $A_2 = cA_1$ *and* $B_2 = \frac{1}{c} B_1$.

DEFINITION 1.3.6. Let X be a $p \times n$ matrix and denote the columns of X by $x_1, x_2, \ldots, x_n$. *Then* $\text{vec}(X) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$.

THEOREM 1.3.17. *i) Let* X *be* $p \times n$, A $q \times p$, *and* B $n \times m$ *matrices. Then,* $\text{vec}((AXB)') = (A \otimes B') \otimes \text{vec}(X')$.

ii) *Let* X *and* Y *be* $p \times n$, A $p \times p$ *and* B $n \times n$ *matrices. Then*

$\text{tr}(X'AYB) = (\text{vec}(X'))' (A \otimes B') \text{vec}(Y')$.

iii) Let X *and* Y *be* $p \times n$ *dimensional matrices. Then* $\text{tr}(X'Y) = (\text{vec}(X'))' \text{vec}(Y')$.

For a more extensive study of the Kronecker product, see Graybill (1969).

Now we give some results on the generalized inverse of a matrix. Since this concept is not so widely used in statistical publications, we prove the theorems in this part of the chapter. For more on generalized inverse, see Rao and Mitra (1971).

DEFINITION 1.3.7. Let A be a $p \times q$ matrix. If there exists a $q \times p$ matrix B such that $ABA = A$, then B is called a generalized inverse of A and is denoted by A^-.

It follows from the definition that A^- is not necessarily unique. For example, for any real number a, $(1,a)$ is a generalized inverse of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. However, if A is a nonsingular square matrix, then A^- is unique as the following theorem shows.

THEOREM 1.3.18. *Let* A *be a* $p \times p$ *nonsingular matrix. Then,* A^{-1} *is the one and only one generalized inverse of* A.

PROOF: The matrix A^{-1} is a generalized inverse of A since $AA^{-1}A = I_pA = A$. On the other hand, from $AA^-A = A$ we get $A^{-1}AA^-AA^{-1} = A^{-1}AA^{-1}$. Hence, $A^- = A^{-1}$. ■

Next, we show that every matrix has a generalized inverse.

THEOREM 1.3.19. *Let* A *be a* $p \times q$ *matrix. Then* A *has a generalized inverse.*

PROOF: First, we prove that the theorem is true for diagonal matrices. Let D be $n \times n$ diagonal and define the $n \times n$ diagonal matrix B by

$$b_{ii} = \begin{cases} \dfrac{1}{d_{ii}} & \text{if } d_{ii} \neq 0 \\ 0 & \text{if } d_{ii} = 0. \end{cases}$$

Then, $DBD = D$. Hence B, is a generalized inverse of D.

Next, assume that A is a $p \times q$ matrix with $p \geq q$. Using Theorem 1.3.6, we can find a $p \times p$ orthogonal matrix G, a $q \times q$ orthogonal matrix H, and a positive semidefinite diagonal matrix D such that $A = G\begin{pmatrix} D \\ O \end{pmatrix}H$. We already know that D has a generalized inverse. Define

$$B = H'(D^-,O)G'.$$

Then, we obtain

$$ABA = G\begin{pmatrix} D \\ O \end{pmatrix}HH'(D^-,O)G'\,G\begin{pmatrix} D \\ O \end{pmatrix}H = G\begin{pmatrix} D \\ O \end{pmatrix}H = A.$$

So B is a generalized inverse of A. If A is $p \times q$ dimensional with $p < q$, then A' has a generalized inverse B. So $A'BA' = A'$. Therefore, $AB'A = A$. Hence, B' is a generalized inverse of A. ■

We know that if A is a nonsingular square matrix, then $(A^{-1})' = (A')^{-1}$. The generalized inverse has the same property as the next theorem shows.

THEOREM 1.3.20. *Let* A *be a* $p \times q$ *matrix. Then,* $A^{-\prime} = A^{\prime -}$*; that is* B *is a generalized inverse of* A*, if and only if* B' *is a generalized inverse of* A'.

PROOF: First, assume B is a generalized inverse of A. Then, $ABA = A$. Hence, $A'B'A' = A'$. So B' is a generalized inverse of A'.

On the other hand, assume B' is a generalized inverse of A'. Then, $A'B'A' = A'$. Hence, $ABA = A$ and therefore B is a generalized inverse of A. ■

For nonsingular square matrices of equal dimension, we have

$(AB)^{-1} = B^{-1}A^{-1}$. However, for the generalized inverse, $(AB)^- = B^-A^-$ does not always hold. For example, consider $A = (1,0)$ and $B = \begin{pmatrix}1\\0\end{pmatrix}$. Then, for any real numbers a and b, $\begin{pmatrix}1\\a\end{pmatrix}$ is a generalized inverse of A and (1,b) is a generalized inverse of B. However, $B^-A^- = 1 + ab$ is a generalized inverse of $AB = 1$ only if $a = 0$ or $b = 0$. In special cases, however, $(AB)^- = B^-A^-$ is true as the next theorem shows.

THEOREM 1.3.21. *Let* A *be a* $q \times q$, B *a* $p \times p$, *and* C *a* $q \times q$ *matrix. Assume* B *and* C *are nonsingular. Then,* $(BAC)^- = C^{-1}A^-B^{-1}$.

PROOF: The matrix F is a generalized inverse of $(BAC)^-$ iff BACFBAC = CAB. This is equivalent to ACFBA = A; that is $CFB = A^-$, or $F = C^{-1}A^-B^{-1}$. ■

THEOREM 1.3.22. *Let* A *be a* $p \times q$ *matrix with* $\mathrm{rk}(A) = q$. *Then,* $A^-A = I_q$.

PROOF: It follows from Theorem 1.3.6, that there exists a $p \times p$ orthogonal matrix G, a $q \times q$ orthogonal matrix H, and a $q \times q$ positive definite diagonal matrix D such that

$$A = G\begin{pmatrix}D\\O\end{pmatrix}H.$$

Then, $AA^-A = A$ can be written as

$$G\begin{pmatrix}D\\O\end{pmatrix}HA^-A = G\begin{pmatrix}D\\O\end{pmatrix}H.$$

Premultiplying the last equation by G', we obtain

$$\begin{pmatrix}D\\O\end{pmatrix}HA^-A = \begin{pmatrix}D\\O\end{pmatrix}H.$$

Hence,

$$\begin{pmatrix} DHA^-A \\ O \end{pmatrix} = \begin{pmatrix} DH \\ O \end{pmatrix}.$$

and consequently $DHA^-A = DH$. Now, D and H are $q \times q$ nonsingular matrices, so we get $A^-A = I_q$. ■

THEOREM 1.3.23. *Let* A *be a* $p \times p$ *matrix of rank* q *and let* $A = BB'$ *be a rank factorization of* A *as it is defined in Theorem 1.3.10. Then,* $B^-B = I_q$ *and* $B^-AB^{-'} = I_q$. *Moreover,* $B^{-'}B^-$ *is a generalized inverse of* A.

PROOF: Since B is $p \times q$ dimensional and $rk(B) = q$, from Theorem 1.3.22 we get $B^-B = I_q$ and

$$B^-AB^{-'} = B^-BB'B^{-'} = B^-B(B^-B)' = I_qI_q = I_q.$$

We also have

$$AB^{-'}B^-A = BB'B^{-'}B^-BB' = B(B^-B)'(B^-B)B' = BI_qI_qB' = BB' = I_q. ■$$

1.4. A FUNCTIONAL EQUATION

We close this chapter with a result from the theory of functional equations that will prove to be useful in the derivation of many theorems about elliptically contoured distributions. The theorem gives the solution of a variant of Hamel's equation (or Cauchy's equation).

THEOREM 1.4.1. *Let* f *be a real function defined on* $\mathbb{R}_0^+$ *the set of nonnegative numbers. Assume that* f *is bounded in each finite interval and satisfies the equation*

$$f(x + y) = f(x)\,f(y) \quad \textit{for all } x,y \in \mathbb{R}_0^+.$$

Then either $f(x) = 0$ *for all* x *or* $f(x) = e^{ax}$ *where* $a \in \mathbb{R}$.

PROOF. See Feller (1957), p. 413. ■

COROLLARY 1.4.1.1. *Let* f *be a bounded, not identically zero function defined on* $\mathbb{R}_0^+$. *If* f *satisfies the equation*

$$f(x + y) = f(x)\, f(y) \quad \textit{for all} \quad x,y \in \mathbb{R}_0^+$$

then $f(x) = e^{-kx}$ *where* $k \geq 0$.

CHAPTER 2

BASIC PROPERTIES

2.1. DEFINITION

In the literature, several definitions of elliptically contoured distributions can be found, e.g. see Anderson and Fang (1982b), Fang and Chen (1984), and Sutradhar and Ali (1989). We will use the following definition given in Gupta and Varga (1991c).

DEFINITION 2.1.1. Let X be a random matrix of dimensions $p \times n$. Then, X is said to have a matrix variate elliptically contoured (m.e.c.) distribution if its characteristic function has the form $\phi_X(T) = \text{etr}(iT'M)$ $\psi(\text{tr}(T'\Sigma T\Phi))$ with T: $p \times n$, M: $p \times n$, Σ: $p \times p$, Φ: $n \times n$, $\Sigma \geq O$, $\Phi \geq O$ and ψ: $[0,\infty) \to \mathbb{R}$.

This distribution will be denoted by $E_{p,n}(M,\Sigma \otimes \Phi,\psi)$.

REMARK 2.1.1. If in Definition 2.1.1 $n = 1$, we say that X has a vector variate elliptically contoured distribution. It is also called multivariate elliptical distribution. Then the characteristic function of X takes on the form $\phi_{\mathbf{x}}(\mathbf{t}) = \exp(i\mathbf{t}'\mathbf{m})\, \psi(\mathbf{t}'\Sigma\mathbf{t})$, where $\mathbf{t}$ and $\mathbf{m}$ are p-dimensional vectors. This definition was given by many authors, e.g. Kelker (1970), Cambanis, Huang and Simons (1981) and Anderson and Fang (1987). In this case, in the notation $E_{p,n}(M,\Sigma \otimes \Phi,\psi)$, the index n can be dropped; that is, $E_p(\mathbf{m},\Sigma,\psi)$ will denote the distribution $E_{p,1}(\mathbf{m},\Sigma,\psi)$.

REMARK 2.1.2. It follows from Definition 2.1.1 that $|\psi(t)| \leq 1$ for $t \in \mathbb{R}_0^+$.

The following theorem shows the relationship between matrix variate and vector variate elliptically contoured distributions.

THEOREM 2.1.1. *Let* X *be a* $p \times n$ *random matrix and* $\mathbf{x} = \mathrm{vec}(X')$. *Then,* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ *if and only if* $\mathbf{x} \sim E_{pn}(\mathrm{vec}(M'), \Sigma \otimes \Phi, \psi)$.

PROOF: Note that $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ iff

$$\phi_X(T) = \mathrm{etr}(iT'M)\ \psi(\mathrm{tr}(T'\Sigma T\Phi)). \tag{2.1}$$

On the other hand, $\mathbf{x} \sim E_{pn}(\mathrm{vec}(M'), \Sigma \otimes \Phi, \psi)$ iff $\phi_{\mathbf{x}}(\mathbf{t}) = \exp(i\mathbf{t}'\ \mathrm{vec}(M'))\ \psi(\mathbf{t}'(\Sigma \otimes \Phi)\mathbf{t})$. Let $\mathbf{t} = \mathrm{vec}(T')$. Then

$$\phi_{\mathbf{x}}(\mathbf{t}) = \exp(i(\mathrm{vec}(T'))'\mathrm{vec}(M'))\psi((\mathrm{vec}(T'))'(\Sigma \otimes \Phi)\ \mathrm{vec}(T')). \tag{2.2}$$

Now, using Theorem 1.3.17, we can write

$$(\mathrm{vec}(T'))'\ (\mathrm{vec}(M')) = \mathrm{tr}(T'M) \tag{2.3}$$

and

$$(\mathrm{vec}(T'))'\ (\Sigma \otimes \Phi)\ \mathrm{vec}(T') = \mathrm{tr}(T'\Sigma T\Phi). \tag{2.4}$$

From (2.1), (2.2), (2.3), and (2.4) it follows that $\phi_X(T) = \phi_{\mathbf{x}}(\mathrm{vec}(T'))$. This completes the proof. ■

The next theorem shows that linear functions of a random matrix with m.e.c. distribution have elliptically contoured distributions also.

THEOREM 2.1.2. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$. *Assume* C: $q \times m$, A: $q \times p$, *and* B: $n \times m$ *are constant matrices. Then,* $AXB + C \sim E_{q,m}(AMB + C, (A\Sigma A') \otimes (B'\Phi B), \psi)$.

PROOF: The characteristic function of $Y = AXB + C$ can be written as

$$\phi_Y(T) = \mathcal{E}(\mathrm{etr}(iT'Y))$$

$$= \mathcal{E}(\mathrm{etr}(iT'(AXB + C)))$$

$$= \mathcal{E}(\mathrm{etr}(iT'AXB))\ \mathrm{etr}(iT'C)$$
$$= \mathcal{E}(\mathrm{etr}(iBT'AX))\ \mathrm{etr}(iT'C)$$

$$= \phi_X(A'TB')\ \mathrm{etr}(iT'C)$$

$$= \mathrm{etr}(iBT'AM)\ \psi(\mathrm{tr}(BT'A\Sigma A'TB'\Phi))\ \mathrm{etr}(iT'C)$$

$$= \mathrm{etr}(iT'(AMB + C))\ \psi(\mathrm{tr}(T'(A\Sigma A) + (B'\Phi B))).$$

This is the characteristic function of $E_{q,m}(AMB + C,(A\Sigma A') \otimes (B'\Phi B),\psi)$. ■

COROLLARY 2.1.2.1. *Let* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$, *and let* $\Sigma = AA'$ *and* $\Phi = BB'$ *be rank factorizations of* Σ *and* Φ. *That is,* A *is* $p \times p_1$ *and* B *is* $n \times n_1$ *matrix, where* $p_1 = \mathrm{rk}(\Sigma)$, $n_1 = \mathrm{rk}(\Phi)$. *Then,*

$$A^-(X - M)B'^- \sim E_{p_1,n_1}(O,I_{p_1} \otimes I_{n_1},\psi).$$

Conversely, if $Y \sim E_{p_1,n_1}(O,I_{p_1} \otimes I_{n_1},\psi)$, *then*

$$AYB' + M \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$$

with $\Sigma = AA'$ *and* $\Phi = BB'$.

PROOF: Let $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$, $\Sigma = AA'$, and $\Phi = BB'$ be the rank factorizations of Σ and Φ. Then, it follows from Theorem 2.1.2 that

$$A^-(X - M)B'^- \sim E_{p_1,n_1}(O,(A^-\Sigma A) \otimes (B^-\Phi B'^-),\psi).$$

Using Theorem 1.3.23, we get $A^-\Sigma A^{-'} = I_{p_1}$ and $B^-\Phi B'^- = I_{n_1}$, which completes the proof of the first part of the theorem. The second part follows directly from Theorem 2.1.2. ■

If $\mathbf{x} \sim E_p(\mathbf{o},I_p,\psi)$, then it follows from Theorem 2.1.2 that

$Gx \sim E_p(\mathbf{o}, I_p, \psi)$ for every $G \in O(p)$. This gives rise to the following definition.

DEFINITION 2.1.2. The distribution $E_p(\mathbf{o}, I_p, \psi)$ is called spherical distribution.

A consequence of the definition of the m.e.c. distribution is that if X has m.e.c. distribution, then X' also has m.e.c. distribution. This is shown in the following theorem.

THEOREM 2.1.3. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$. *Then,* $X' \sim E_{n,p}(M', \Phi \otimes \Sigma, \psi)$.

PROOF: We have

$$\begin{aligned}
\phi_{X'}(T) &= \mathcal{E}(\mathrm{etr}(iT'X')) \\
&= \mathcal{E}(\mathrm{etr}(iXT)) \\
&= \mathcal{E}(\mathrm{etr}(iTX)) \\
&= \mathrm{etr}(iTM)\ \psi(\mathrm{tr}(T\Sigma T'\Phi)) \\
&= \mathrm{etr}(iT'M')\ \psi(\mathrm{tr}(T'\Phi T\Sigma)).
\end{aligned}$$

This is the characteristic function of $E_{n,p}(M', \Phi \otimes \Sigma, \psi)$. ■

The question arises whether the parameters in the definition of a m.e.c. distribution are uniquely defined. The answer is they are not. To see this assume that a, b, and c are positive constants such that $c = ab$, $\Sigma_2 = a\Sigma_1$, $\Phi_2 = b\Phi_1$, $\psi_2(z) = \psi_1\left(\frac{1}{c}z\right)$. Then, $E_{p,n}(M, \Sigma_1 \otimes \Phi_1, \psi_1)$ and $E_{p,n}(M, \Sigma_2 \otimes \Phi_2, \psi_2)$ define the same m.e.c. distribution. However, this is the only way that two formulae define the same m.e.c. distribution as shown in the following theorem.

THEOREM 2.1.4. *Let* $X \sim E_{p,n}(M_1,\Sigma_1 \otimes \Phi_1,\psi_1)$ *and at the same time* $X \sim E_{p,n}(M_2,\Sigma_2 \otimes \Phi_2,\psi_2)$. *If* X *is nondegenerate, then there exist positive constants* $a, b,$ *and* c *such that* $c = ab$ *and* $M_2 = M_1$, $\Sigma_2 = a\Sigma_1$, $\Phi_2 = b\Phi_1$, *and* $\psi_2(z) = \psi_1\left(\frac{1}{c}z\right)$.

PROOF: The proof follows the lines of Cambanis, Huang and Simons (1981). First of all, note that the distribution of X is symmetric about M_1 as well as about M_2. Therefore, $M_1 = M_2$ must hold. Let $M = M_1$. Let us introduce the following notations.

$$\Sigma_\ell = {}_\ell\sigma_{ij}, \quad i,j = 1,\dots,p; \quad \ell = 1,2,$$

$$\Phi_\ell = {}_\ell\phi_{ij}, \quad i,j = 1,\dots,n; \quad \ell = 1,2.$$

Let $\mathbf{k}(a)$ denote the p-dimensional vector whose a^{th} element is 1 and all the others are 0 and $\mathbf{l}(b)$ denote the n-dimensional vector whose b^{th} element is 1 and all the others are 0. Since X is nondegenerate, it must have an element $x_{i_0j_0}$ which is nondegenerate. Since $x_{i_0j_0} = \mathbf{k}'(i_0)\, X\, \mathbf{l}(j_0)$, from Theorem 2.1.2 we get

$$x_{i_0j_0} \sim E_1(m_{i_0j_0}, {}_1\sigma_{i_0i_0}\, {}_1\phi_{j_0j_0}, \psi_1) \quad \text{and}$$

$$x_{i_0j_0} \sim E_1(m_{i_0,j_0}, {}_2\sigma_{i_0i_0}\, {}_2\phi_{j_0j_0}, \psi_2).$$

Therefore, the characteristic function of $x_{i_0j_0} - m_{i_0j_0}$ is

$$\begin{aligned}\phi(t) &= \psi_1(t^2\, {}_1\sigma_{i_0i_0}\, {}_1\phi_{j_0j_0}) \\ &= \psi_2(t^2\, {}_2\sigma_{i_0i_0}\, {}_2\phi_{j_0j_0}) \qquad (2.5)\end{aligned}$$

with $t \in \mathbb{R}$.

Since, ${}_\ell\sigma_{i_0i_0}$ and ${}_\ell\phi_{j_0j_0}$ ($\ell = 1,2$) are diagonal elements of positive semidefinite matrices, they cannot be negative, and since $x_{i_0j_0}$ is nondegenerate, they cannot be zero either. So, we can define

$$c = \frac{{}_2\sigma_{i_0 i_0}\, {}_2\phi_{j_0 j_0}}{{}_1\sigma_{i_0 i_0}\, {}_1\phi_{j_0 j_0}}.$$

Then, $c > 0$ and $\psi_2(z) = \psi_1\left(\frac{1}{c}z\right)$ for $z \in [0,\infty)$.

We claim that $\Sigma_2 \otimes \Phi_2 = c(\Sigma_1 \otimes \Phi_1)$. Suppose this is not the case. Then, there exists $\mathbf{t} \in \mathbb{R}^{pn}$ such that $\mathbf{t}'(\Sigma_2 \otimes \Phi_2)\mathbf{t} \neq c\mathbf{t}'(\Sigma_1 \otimes \Phi_1)\mathbf{t}$. From Theorem 1.3.17, it follows that there exists $T_0 \in \mathbb{R}^{p\times n}$ such that $\mathrm{tr}(T_0'\Sigma_2 T_0 \Phi_2) \neq c\,\mathrm{tr}(T_0'\Sigma_1 T_0 \Phi_1)$.

Define $T = uT_0$, $u \in \mathbb{R}$. Then, the characteristic function of $X - M$ at uT_0 is

$$\psi_1(u\,\mathrm{tr}(T_0'\Sigma_1 T_0 \Phi_1)) = \psi_2(u\,c\,\mathrm{tr}(T_0' T_0 \Phi_1)).$$

On the other hand, the characteristic function of $X - M$ at uT_0 can be expressed as $\psi_2(u\,\mathrm{tr}(T_0'\Sigma_2 T_0 \Phi_2))$. So

$$\psi_2(u\,c\,\mathrm{tr}(T_0'\Sigma_1 T_0 \Phi_1)) = \psi_2(u\,\mathrm{tr}(T_0'\Sigma_2 T_0 \Phi_2)). \tag{2.6}$$

If $\mathrm{tr}(T_0'\Sigma_1 T_0 \Phi_1) = 0$ or $\mathrm{tr}(T_0'\Sigma_2 T_0 \Phi_2) = 0$, then from (2.6) we get that $\psi(u) = 0$ for every $u \in \mathbb{R}$. However, this is impossible since X is nondegenerate.

If $\mathrm{tr}(T_0'\Sigma_1 T_0 \Phi_1) \neq 0$ and $\mathrm{tr}(T_0'\Sigma_2 T_0 \Phi_2) \neq 0$, then define $d = c\frac{\mathrm{tr}(T_0'\Sigma_1 T_0 \Phi_1)}{\mathrm{tr}(T_0'\Sigma_2 T_0 \Phi_2)}$. Then, $d \neq 0$, $d \neq 1$, and from (2.6) we get $\psi_2(u) = \psi_2(du)$. By induction, we get

$$\psi_2(u) = \psi_2(d^n u) \text{ and } \psi_2(u) = \psi_2\left(\left(\frac{1}{d}\right)^n u\right), \quad n = 1,2,\ldots.$$

Now either $d^n \to 0$ or $\left(\frac{1}{d}\right)^n \to 0$, and from the continuity of the characteristic function and the fact that $\psi_2(0) = 1$ it follows that $\psi_2(u) = 0$ for every $u \in \mathbb{R}$. However, this is impossible. So, we must have

$\Sigma_2 \otimes \Phi_2 = c(\Sigma_1 \otimes \Phi_1)$. From Theorem 1.3.16 it follows that there exist $a > 0$ and $b > 0$ such that $\Sigma_2 = a\Sigma_1$, $\Phi_2 = b\Phi_1$, and $ab = c$. This completes the proof. ■

An important subclass of the class of the m.e.c. distributions is the class of matrix variate normal distributions.

DEFINITION 2.1.3. The $p \times n$ random matrix X is said to have a matrix variate normal distribution if its characteristic function has the form $\phi_X(T) = \operatorname{etr}(iT'M) \operatorname{etr}\left(-\frac{1}{2}T'\Sigma T\Phi\right)$, with T: $p \times n$, M: $p \times n$, Σ: $p \times p$, Φ: $n \times n$, $\Sigma \geq O$, $\Phi \geq O$. This distribution is denoted by $N_{p,n}(M,\Sigma \otimes \Phi)$.

The next theorem shows that the matrix variate normal distribution can be used to represent samples taken from multivariate normal distributions.

THEOREM 2.1.5. *Let* $X \sim N_{p,n}(\mu e_n', \Sigma \otimes I_n)$, *where* $\mu \in \mathbb{R}^p$. *Let* $x_1, x_2, \ldots, x_n$ *be the columns of* X. *Then,* $x_1, x_2, \ldots, x_n$ *are independent identically distributed random vectors with common distribution* $N_p(\mu,\Sigma)$.

PROOF: Let $T = (t_1, t_2, \ldots, t_n)$ be $p \times n$ matrix. Then

$$\phi_X(T) = \operatorname{etr}\left(i\begin{pmatrix} t_1' \\ t_2' \\ \vdots \\ t_n' \end{pmatrix}(\mu,\mu,\ldots,\mu)\right)\operatorname{etr}\left(-\frac{1}{2}\begin{pmatrix} t_1' \\ t_2' \\ \vdots \\ t_n' \end{pmatrix}\Sigma(t_1,t_2,\ldots,t_n)\right)$$

$$= \exp\left(i\sum_{j=1}^{n} t_j'\mu\right)\exp\left(-\frac{1}{2}\sum_{j=1}^{n} t_j'\Sigma t_j\right)$$

$$= \prod_{j=1}^{n} \exp\left(i\, t_j'\mu - \frac{1}{2}t_j'\Sigma\, t_j\right);$$

which shows that $x_1,x_2,\dots,x_n$ are independent, each with distribution $N_p(\mu,\Sigma)$. ■

2.2. PROBABILITY DENSITY FUNCTION

If $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ defines an absolutely continuous elliptically contoured distribution, Σ and Φ must be positive definite. Assume this is not the case. For example, $\Sigma \geq O$ but Σ is not positive definite. Then, from Theorem 1.3.7, it follows that $\Sigma = GDG'$ where $G \in O(n)$, and D is diagonal and $d_{11} = 0$. Let $Y = G'(X - M)$. Then, $Y \sim E_{p,n}(O,D \otimes \Phi,\psi)$, and the distribution of Y is also absolutely continuous. On the other hand, $y_{11} \sim E_1(0,0,\psi)$ so y_{11} is degenerate. But the marginal of an absolutely continuous distribution cannot be degenerate. Hence, we get a contradiction. So, $\Sigma > O$ and $\Phi > O$ must hold when the m.e.c. distribution is absolutely continuous.

The probability density function (p.d.f.) of a m.e.c. distribution is of a special form as the following theorem shows.

THEOREM 2.2.1. *Let X be a $p \times n$ dimensional random matrix whose distribution is absolutely continuous. Then, $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ if and only if the p.d.f. of X has the form*

$$f(X) = |\Sigma|^{-\frac{n}{2}} |\Phi|^{-\frac{p}{2}} h(\mathrm{tr}((X - M)' \Sigma^{-1}(X - M)\Phi^{-1})), \tag{2.7}$$

where h and ψ determine each other for specified p and n.

PROOF: I. First, we prove that if $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ and $E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ is absolutely continuous, then the p.d.f. of X has the form (2.7).

Step 1. Assume that $M = O$ and $\Sigma \otimes \Phi = I_{pn}$. Then, $X \sim E_{p,n}(O,I_p \otimes I_n,\psi)$. We want to show that the p.d.f. of X depends on X only through $\mathrm{tr}(X'X)$. Let $\mathbf{x} = \mathrm{vec}(X')$. From Theorem 2.1.1 we know that $\mathbf{x} \sim E_{pn}(\mathbf{o},I_{pn},\psi)$. Let $H \in O(pn)$, then, in view of Theorem 2.1.2,

$$Hx \sim E_{pn}(o, HH', \psi) = E_{pn}(o, I_{pn}, \psi).$$

Thus, the distribution of x is invariant under orthogonal transformation. Therefore, using Theorem 1.3.11, we conclude that the p.d.f. of x depends on x only through $x'x$. Let us denote the p.d.f. of x by $f_1(x)$. We have $f_1(x) = h(x'x)$. Clearly, h only depends on p, n, and ψ. It follows from Theorem 1.3.17, that $x'x = tr(X'X)$. Thus, denoting the p.d.f. of X by $f(X)$, we get $f(X) = h(tr(X'X))$.

Step 2. Now, let $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$. From Corollary 2.1.2.1, it follows that $Y = \Sigma^{-\frac{1}{2}}(X - M)\Phi^{-\frac{1}{2}} \sim E_{p,n}(O, I_p \otimes I_n, \psi)$. Therefore, if $g(Y)$ is the p.d.f. of Y, then $g(Y) = h(tr(Y'Y))$. The Jacobian of the transformation $Y \to X$ is $|\Sigma^{-\frac{1}{2}}|^n \, |\Phi^{-\frac{1}{2}}|^p$ (see Press, 1972, p. 45). So the p.d.f. of X is

$$f(X) = h(tr(\Phi^{-\frac{1}{2}}(X - M)' \, \Sigma^{-\frac{1}{2}} \Sigma^{-\frac{1}{2}}(X - M)\Phi^{-\frac{1}{2}}))|\Sigma|^{-\frac{n}{2}} |\Phi|^{-\frac{p}{2}}$$

$$= |\Sigma|^{-\frac{n}{2}} |\Phi|^{-\frac{p}{2}} h(tr((X - M)' \, \Sigma^{-1}(X - M)\Phi^{-1})).$$

II. Next, we show that if a random matrix X has the p.d.f. of the form (2.7), then its distribution is elliptically contoured. That is, assume that the $p \times n$ random matrix X has the p.d.f.

$$f(X) = |\Sigma|^{-\frac{n}{2}} |\Phi|^{-\frac{p}{2}} h(tr((X - M)' \, \Sigma^{-1}(X - M)\Phi^{-1})),$$

then we want to show that $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$. Let $Y = \Sigma^{-\frac{1}{2}}(X - M)\Phi^{-\frac{1}{2}}$. Then, the p.d.f. of Y is $g(Y) = h(tr(Y'Y))$. Let $y = vec(Y')$. Then, the p.d.f. of y is $g_1(y) = h(y'y)$. The characteristic function of y is

$$\phi_y(t) = \int_{\mathbb{R}^{pn}} \exp(it'y) \, h(y'y) dy,$$

where $t \in \mathbb{R}^{pn}$.

Next, we prove that if t_1 and t_2 are vectors of dimension pn such that $t_1't_1 = t_2't_2$, then $\phi_y(t_1) = \phi_y(t_2)$. Using Theorem 1.3.11, we see that there exists $H \in O(pn)$, such that $t_1'H = t_2'$. Therefore,

$$\begin{aligned}\phi_y(t_2) &= \int_{\mathbb{R}^{pn}} \exp(it_2'y)\, h(y'y)dy \\ &= \int_{\mathbb{R}^{pn}} \exp(it_1'Hy)\, h(y'y)dy.\end{aligned}$$

Let $z = Hy$. The Jacobian of the transformation $y \to z$ is $|H'|^{pn} = 1$. So

$$\begin{aligned}\int_{\mathbb{R}^{pn}} \exp(it_1'Hy)\, h(y'y)dy &= \int_{\mathbb{R}^{pn}} \exp(it_1'z)\, h(z'HH'z)dz \\ &= \int_{\mathbb{R}^{pn}} \exp(it_1'z)\, h(z'z)dz \\ &= \phi_y(t_1).\end{aligned}$$

This means that $\phi_y(t_1) = \phi_y(t_2)$. Therefore, $\phi_y(t)$ is a function of $t_1't_1$, which implies that $\phi_Y(T)$ is a function of $tr(T'T)$. Therefore, there exists a function ψ such that $\phi_Y(T) = \psi(tr(T'T))$. That is, $y \sim E_{p,n}(O, I_p \otimes I_n, \psi)$. Using Corollary 2.1.2.1, we get $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$. ■

Next we prove a lemma which will be useful for further study of m.e.c. distributions.

LEMMA 2.2.1. *Let* f *be a function* $f\colon A \times \mathbb{R}^p \to \mathbb{R}^q$, *where* A *can be any set. Assume there exists a function* $g\colon A \times \mathbb{R} \to \mathbb{R}^q$ *such that* $f(a,x) = g(a,x'x)$ *for any* $a \in A$ *and* $x \in \mathbb{R}^p$. *Then, we have*

$$\int_{x\in\mathbb{R}^p} f(a,x)dx = \frac{2\pi^{\frac{p}{2}}}{\Gamma\left(\frac{p}{2}\right)} \int_0^\infty g(a,r^2)dr$$

for any $a \in A$.

PROOF: Let $x = (x_1,x_2,\ldots,x_p)'$ and introduce the polar coordinates

$$x_1 = r \sin\theta_1 \sin\theta_2 \ldots \sin\theta_{p-2} \sin\theta_{p-1}$$

$$x_2 = r \sin\theta_1 \sin\theta_2 \ldots \sin\theta_{p-2} \cos\theta_{p-1}$$

$$x_3 = r \sin\theta_1 \sin\theta_2 \ldots \cos\theta_{p-2}$$

$$\vdots$$

$$x_{p-1} = r \sin\theta_1 \cos\theta_2$$

$$x_p = r \cos\theta_1,$$

where $r > 0$, $0 < \theta_i < \pi$, $i = 1,2,\ldots,p-2$, and $0 < \theta_{p-1} < 2\pi$. Then, the Jacobian of the transformation $(x_1,x_2,\ldots,x_p) \to (r,\theta_1,\theta_2,\ldots,\theta_{p-1})$ is $r^{p-1} \sin^{p-2}\theta_1 \sin^{p-3}\theta_2 \ldots \sin\theta_{p-2}$.

We also have $x'x = r^2$. Thus,

$$\int_{x\in\mathbb{R}^p} f(a,x)dx = \int_{x\in\mathbb{R}^p} g(a,x'x)dx$$

$$= \int_0^\infty \int_0^\pi \int_0^\pi \cdots \int_0^{2\pi} g(a,r^2)r^{p-1} \sin^{p-2}\theta_1 \sin^{p-3}\theta_2 \ldots \sin\theta_{p-2}d\theta_{p-1}\ldots d\theta_2 d\theta_1 dr$$

$$= \frac{2\pi^{\frac{p}{2}}}{\Gamma\left(\frac{p}{2}\right)} \int_0^\infty r^{p-1}\, g(a,r^2)dr. \blacksquare$$

The next theorem is due to Fang, Kotz, and Ng(1990).

THEOREM 2.2.2. *Let* $g: \mathbb{R}_0^+ \to \mathbb{R}_0^+$ *be a measurable function. Then, there exists a constant* c *such that*

$$c\, g(\mathrm{tr}(X'X)),\ X \in \mathbb{R}^{p\times n}$$

is the p.d.f. of the $p \times n$ *random matrix* X *if and only if*

$$0 < \int_0^\infty r^{np-1}\, g(r^2)dr < \infty.$$

Moreover, the relationship between g *and* c *is given by*

$$c = \frac{\Gamma\left(\frac{pn}{2}\right)}{2\pi^{\frac{pn}{2}} \int_0^\infty r^{np-1}\, g(r^2)dr}.$$

PROOF: By definition, $c\, g(\mathrm{tr}(X'X))$, $X \in \mathbb{R}^{p\times n}$, is the p.d.f. of a $p \times n$ random matrix X iff $c\, g(\mathbf{y}'\mathbf{y})$; $\mathbf{y} \in \mathbb{R}^{pn}$ is the p.d.f. of a pn-dimensional random vector y. On the other hand, $c\, g(\mathbf{y}'\mathbf{y})$, $\mathbf{y} \in \mathbb{R}^{pn}$ is the p.d.f. of a pn-dimensional random vector **y** iff

$$\int_{\mathbb{R}^{pn}} c\, g(\mathbf{y}'\mathbf{y})d\mathbf{y} = 1.$$

From Lemma 2.2.1, we get

$$\int_{\mathbb{R}^{pn}} c\, g(y'y)dy = c \frac{2\pi^{\frac{pn}{2}}}{\Gamma\left(\frac{pn}{2}\right)} \int_0^\infty r^{pn-1} g(r^2)dr.$$

Hence, we must have

$$0 \leq \int_0^\infty r^{pn-1} g(r^2)dr < \infty$$

and $c = \dfrac{\Gamma\left(\frac{pn}{2}\right)}{2\pi^{\frac{pn}{2}} \int_0^\infty r^{pn-1} g(r^2)dr}$. ■

2.3. MARGINAL DISTRIBUTIONS

Using Theorem 2.1.2, we can derive the marginal distributions of a m.e.c. distribution.

THEOREM 2.3.1. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$, *and partition* X, M, *and* Σ *as*

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad M = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}, \quad \text{and } \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

where X_1 *is* $q \times n$, M_1 *is* $q \times n$, *and* Σ_{11} *is* $q \times q$, $1 \leq q < p$. *Then,* $X_1 \sim E_{q,n}(M_1, \Sigma_{11} \otimes \Phi, \psi)$.

PROOF: Let $A = (I_q, O)$ be of dimensions $q \times p$. Then, $AX = X_1$, and from Theorem 2.1.2, we obtain $X_1 \sim E_{q,n}\left((I_q,O)\, M_1, \left((I_q,O)\, \Sigma \begin{pmatrix} I_q \\ O \end{pmatrix}\right) \otimes \Phi, \psi\right)$ i.e. $X_1 \sim E_{q,n}(M_1, \Sigma_{11} \otimes \Phi, \psi)$. ■

If we partition X vertically, we obtain the following result.

THEOREM 2.3.2. *Let* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$, *and partition* $X, M,$ *and* Φ *as*

$$X = (X_1,X_2),\ M = (M_1,M_2) \text{ and } \Phi = \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix},$$

where X_1 *is* $p \times m$, M_1 *is* $p \times m$, *and* Φ_{11} *is* $m \times m$, $1 \le m < n$. *Then,*

$$X_1 \sim E_{p,n}(M_1,\Sigma \otimes \Phi_{11},\psi). \tag{2.8}$$

PROOF: From Theorem 2.1.3, it follows that

$$X' = \begin{pmatrix} X_1' \\ X_2' \end{pmatrix} \sim E_{n,p}\left(\begin{pmatrix} M_1' \\ M_2' \end{pmatrix}, \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix} \otimes \Sigma,\psi\right).$$

Then (2.8) follows directly from Theorem 2.3.1. ■

THEOREM 2.3.3. *Let* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$, *then* $x_{ij} \sim E_1(m_{ij},\sigma_{ij}\phi_{jj},\psi)$.
PROOF: The result follows from Theorem 2.3.1, and Theorem 2.3.2. ■

REMARK 2.3.1. It follows from Theorem 2.3.1, and Theorem 2.3.2 that if $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ and Y is a $q \times m$ submatrix of X, then Y also has m.e.c. distribution; $Y \sim E_{q,m}(M^*,\Sigma^* \otimes \Phi^*,\psi)$.

2.4. EXPECTED VALUE AND COVARIANCE

In this section, the first two moments of a m.e.c. distribution will be derived. In Chapter 3, moments of higher orders will also be obtained.

THEOREM 2.4.1. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$.

a) If X *has finite first moment, then* $E(X) = M$.

b) If X *has finite second moment, then* $\mathrm{Cov}(X) = c\Sigma \otimes \Phi$, *where* $c = -2\psi'(0)$.

PROOF: Step 1. First, let us assume $M = O$ and $\Sigma \otimes \Phi = I_{pn}$. Then, $X \sim E_{p,n}(O, I_p \otimes I_n, \psi)$.

a) In view of Theorem 2.1.2, we have

$$(-I_p)X \sim E_{p,n}(O, I_p \otimes I_n, \psi).$$

Therefore, $\mathcal{E}(X) = \mathcal{E}(-X)$, and $\mathcal{E}(X) = O$.

b) Let $\mathbf{x} = \mathrm{vec}(X')$. Then $\mathbf{x} \sim E_{pn}(\mathbf{o}, I_{pn}, \psi)$. The characteristic function of $\mathbf{x}$ is $\phi_{\mathbf{x}}(\mathbf{t}) = \psi(\mathbf{t}'\mathbf{t})$, where $\mathbf{t} = (t_1, \ldots, t_{pn})'$. Then, $\dfrac{\partial \phi_{\mathbf{x}}(\mathbf{t})}{\partial t_i} = \dfrac{\partial \psi\left(\sum_{\ell=1}^{pn} t_\ell^2\right)}{\partial t_i} = 2t_i \psi'\left(\sum_{\ell=1}^{pn} t_\ell^2\right)$. So,

$$\frac{\partial^2 \phi_{\mathbf{x}}(\mathbf{t})}{\partial t_i^2} = 2\psi'\left(\sum_{\ell=1}^{pn} t_\ell^2\right) + 4t_i^2 \psi''\left(\sum_{\ell=1}^{pn} t_\ell^2\right)$$

and if $i \neq j$, then $\dfrac{\partial \phi_{\mathbf{x}}(\mathbf{t})}{\partial t_j \partial t_i} = 4t_i t_j \psi''\left(\sum_{\ell=1}^{pn} t_\ell^2\right)$. Therefore,

$$\left.\frac{\partial^2 \phi_{\mathbf{x}}(\mathbf{t})}{\partial t_i^2}\right|_{\mathbf{t}=0} = 2\psi'(0) \text{ and } \frac{\partial^2 \phi_{\mathbf{x}}(\mathbf{t})}{\partial t_j \partial t_i} = 0 \text{ if } i \neq j.$$

Thus, $\mathrm{Cov}(\mathbf{x}) = -2\psi'(0) I_{pn}$.

Step2. Now, let $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$. Let $\Sigma = AA'$ and $\Phi = BB'$ be the rank factorizations of Σ and Φ. Then, from Corollary 2.1.2.1, it follows that $Y = A^{-}(X - M)B'^{-} \sim E_{p_1,n_1}(O, I_{p_1} \otimes I_{n_1}, \psi)$ and $X = AYB'$. Using Step 1 ,we get the following results:

a) $\mathcal{E}(Y) = O$. Hence $\mathcal{E}(X) = AOB' + M = M$.

b) Let $\mathbf{x} = \mathrm{vec}(X')$, $\mathbf{y} = \mathrm{vec}(Y')$, and $\mathbf{m} = \mathrm{vec}(M')$. Then $\mathbf{x} = (A \otimes B)\mathbf{y} + \mathbf{m}$, and $\mathrm{Cov}(\mathbf{y}) = -2\psi'(0)I_{pn}$ and so

$$\begin{aligned}\mathrm{Cov}(\mathbf{x}) &= -2\psi'(0)(A \otimes B)\, I_{pn}(A' \otimes B') \\ &= -2\psi'(0)(AA') \otimes (BB') \\ &= -2\psi'(0)\Sigma \otimes \Phi. \blacksquare\end{aligned}$$

COROLLARY 2.4.1.1. *With the conditions of Theorem 2.4.1, the* i^{th} ($i = 1,2,\ldots,p$) *column of the matrix* X *has the covariance matrix* $\phi_{ii}\Sigma$ *and the* j^{th} *row* ($j = 1,\ldots,n$) *has the covariance matrix* $\sigma_{jj}\Phi$.

COROLLARY 2.4.1.2. *With the conditions of Theorem 2.4.1,*

$$\mathrm{Corr}(x_{ij}, x_{k\ell}) = \frac{\sigma_{ik}\,\phi_{j\ell}}{\sqrt{\sigma_{ii}\sigma_{kk}\phi_{jj}\phi_{\ell\ell}}},$$

that is, the correlations between two elements of the matrix X, *depend only on* Σ *and* Φ *but not on* ψ.

PROOF: From Theorem 2.4.1, we get $\mathrm{Cov}(x_{ij}, x_{k\ell}) = c\sigma_{ik}\phi_{j\ell}$, $\mathrm{Var}(x_{ij}) = c\sigma_{ii}\phi_{jj}$, and $\mathrm{Var}(x_{k\ell}) = c\sigma_{kk}\phi_{\ell\ell}$, where $c = -2\psi'(0)$. Therefore

$$\begin{aligned}\mathrm{Corr}(x_{ij}, x_{k\ell}) &= \frac{c\sigma_{ik}\phi_{j\ell}}{\sqrt{c^2\sigma_{ii}\phi_{jj}\sigma_{kk}\phi_{\ell\ell}}} \\ &= \frac{\sigma_{ik}\phi_{j\ell}}{\sqrt{\sigma_{ii}\phi_{jj}\sigma_{kk}\phi_{\ell\ell}}} \blacksquare\end{aligned}$$

2.5. STOCHASTIC REPRESENTATION

In Cambanis, Huang, and Simons (1981) the stochastic representation of vector variate elliptically contoured distribution was obtained using a result of Schoenberg (1938). This result was extended to m.e.c. distributions by Anderson and Fang (1982b). Shoenberg's result is given in the next theorem.

THEOREM 2.5.1. *Let* ψ *be a real function* ψ: $[0,\infty) \to \mathbb{R}$. *Then,* $\psi(\mathbf{t}'\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^k$ *is the characteristic function of a k-dimensional random variable* $\mathbf{x}$, *if and only if* $\psi(u) = \int_0^\infty \Omega_k(r^2u)\, dF(r)$, $u \geq 0$, *where* F *is a distribution function on* $[0,\infty)$ *and* $\Omega_k(\mathbf{t}'\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^k$ *is the characteristic function of the k-dimensional random variable* $\mathbf{u}_k$ *which is uniformly distributed on the unit sphere in* $\mathbb{R}^k$. *Moreover,* $F(r)$ *is the distribution function of* $r = (\mathbf{x}'\mathbf{x})^{\frac{1}{2}}$.

PROOF: Let us denote the unit sphere in $\mathbb{R}^k$ by S_k:

$$S_k = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^k;\ \mathbf{x}'\mathbf{x} = 1\},$$

and let A_k be the surface area of S_k i.e. $A_k = \dfrac{2\pi^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}\right)}$.

First, assume $\psi(u) = \int_0^\infty \Omega_k(r^2u)\, dF(r)$. Let r be a random variable with distribution function $F(r)$, and let $\mathbf{u}_k$ be independent of r and uniformly distributed on S_k. Define $\mathbf{x} = r\mathbf{u}_k$. Then, the characteristic function of $\mathbf{x}$ is

$$\begin{aligned}\phi_{\mathbf{x}}(\mathbf{t}) &= \mathcal{E}(\exp(i\mathbf{t}'\mathbf{x}))\\ &= \mathcal{E}(\exp(i\mathbf{t}'r\mathbf{u}_k))\\ &= \mathcal{E}\{\mathcal{E}(\exp(i\mathbf{t}'r\mathbf{u}_k) \mid r)\}\end{aligned}$$

$$= \int_0^\infty \mathcal{E}(\exp(it'ru_k) \mid r = y)\, dF(y)$$

$$= \int_0^\infty \phi_{u_k}(yt)\, dF(y)$$

$$= \int_0^\infty \Omega_k(y^2 t't)\, dF(y).$$

Therefore, $\psi(t't) = \int_0^\infty \Omega_k(y^2 t't)\, dF(y)$ is indeed the characteristic function of the k-dimensional random vector x. Moreover,
$F(y) = P(r \le y) = P((r^2)^{\frac{1}{2}} \le y) = P(((ru_k)'(ru_k))^{\frac{1}{2}} \le y) = P((x'x)^{\frac{1}{2}} \le y)$.

Conversely, assume $\psi(t't)$ is the characteristic function of a k-dimensional random vector x. Let $G(x)$ be the distribution function of x. Let $d\omega_k(t)$ denote the integration on S_k. We have $\psi(u) = \psi(ut't)$ for $t't = 1$, and therefore we can write

$$\psi(u) = \frac{1}{A_k} \int_{S_k} \psi(ut't)\, d\omega_k(t)$$

$$= \frac{1}{A_k} \int_{S_k} \phi_x(\sqrt{u}\, t)\, d\omega_k(t)$$

$$= \frac{1}{A_k} \int_{S_k} \int_{\mathbb{R}^m} \exp(i \sqrt{u}\, t'x)\, dG(x)\, d\omega_k(t)$$

$$= \int_{\mathbb{R}^m} \left(\frac{1}{A_k} \int_{S_k} \exp(i \sqrt{u}\, x't)\, d\omega_k(t) \right) dG(x)$$

$$= \int_{\mathbb{R}^m} \Omega_k((\sqrt{u}\, x)' (\sqrt{u}\, x))\, dG(x)$$

$$= \int_{\mathbb{R}^m} \Omega_k(ux'x)\, dG(x)$$

$$= \int_0^\infty \Omega_k(uy^2)\, dF(y),$$

where $F(y) = P((x'x)^{\frac{1}{2}} \leq y)$. ■

Now, we can derive the stochastic representation of a m.e.c. distribution.

THEOREM 2.5.2. *Let* X *be a* $p \times n$ *random matrix. Let* M *be* $p \times n$, Σ *be* $p \times p$, *and* Φ *be* $n \times n$ *constant matrices,* $\Sigma \geq O$, $\Phi \geq O$, $\text{rk}(\Sigma) = p_1$, $\text{rk}(\Phi) = n_1$. *Then,*

$$X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi) \tag{2.9}$$

if and only if

$$X \approx M + rAUB', \tag{2.10}$$

where U *is* $p_1 \times n_1$ *and* $\text{vec}(U')$ *is uniformly distributed on* $S_{p_1 n_1}$, r *is a nonnegative random variable,* r *and* U *are independent,* $\Sigma = AA'$, *and* $\Phi = BB'$ *are rank factorizations of* Σ *and* Φ. *Moreover,* $\psi(u) = \int_0^\infty \Omega_{p_1 n_1}(r^2 u)\, dF(r)$, $u \geq 0$, *where* $\Omega_{p_1 n_1}(t't)$, $t \in \mathbb{R}^{p_1 n_1}$ *denotes the characteristic function of* $\text{vec}(U')$, *and* $F(r)$ *denotes the distribution function of* r. *The expression,* $M + rAUB'$, *is called the stochastic representation of* X.

PROOF: First, assume $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$. Then, it follows from Corollary 2.1.2.1, that $Y = A^-(X - M)B'^- \sim E_{p_1,n_1}(O,I_{p_1} \otimes I_{n_1},\psi)$. Thus, $\mathbf{y} = \text{vec}(Y') \sim E_{p_1n_1}(\mathbf{o},I_{p_1n_1},\psi)$.

So, $\psi(\mathbf{t}'\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^{p_1n_1}$ is a characteristic function and from Theorem 2.6.1, we get

$$\psi(u) = \int_0^\infty \Omega_{p_1n_1}(y^2u)\, dF(y), \quad u \geq 0,$$

which means that $\mathbf{y} \approx r\mathbf{u}$, where r is nonnegative with distribution function $F(y)$, $\mathbf{u}$ is uniformly distributed on $S_{p_1n_1}$, and r and $\mathbf{u}$ are independent. Therefore, we can write $\mathbf{y} \approx r\mathbf{u}$, where $\mathbf{u} = \text{vec}(U')$. Now, using Corollary 2.1.2.1 again, we get

$$X \approx AYB' + M \approx M + rAUB'.$$

Conversely, suppose $X \approx M + rAUB'$. Let $\mathbf{u} = \text{vec}(U')$. Define $\psi(u) = \int_0^\infty \Omega_{p_1n_1}(y^2u)\, dF(y)$ where $F(y)$ is the distribution function of r, $u \geq 0$. Then, it follows from Theorem 2.6.1, that $\psi(\mathbf{t}'\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^{p_1n_1}$ is the characteristic function of $r\mathbf{u}$. So $r\mathbf{u} \sim E_{p_1n_1}(\mathbf{o},I_{p_1n_1},\psi)$ and hence,

$$rU \sim E_{p_1,n_1}(O,I_{p_1} \otimes I_{n_1},\psi).$$

Therefore,

$$X \approx M + rAUB' \sim E_{p,n}(M,(AA') \otimes (BB'),\psi) = E_{p,n}(M,\Sigma \otimes \Phi,\psi). \blacksquare$$

It may be noted that the stochastic representation is not uniquely defined. We can only say the following.

THEOREM 2.5.3. $M_1 + r_1A_1UB_1'$ *and* $M_2 + r_2A_2UB_2'$, *where* U *is*

$p_1 \times n_1$, *are two stochastic representations of the same* $p \times n$ *dimensional nondegenerate m.e.c. distribution if and only if* $M_1 = M_2$, *and there exist* $G \in O(p_1)$, $H \in O(n_1)$, *and positive constants* a, b, *and* c *such that* $ab = c$, $A_2 = aA_1G$, $B_2 = bB_1H$, *and* $r_2 = \frac{1}{c} r_1$.

PROOF: The "if" part is trivial. Conversely, let $X \approx M_1 + r_1A_1UB_1'$ and $X \approx M_2 + r_2A_2UB_2'$. Then

$$X \sim E_{p,n}(M_1,(A_1A') \otimes (B_1B_1'),\psi_1) \text{ and}$$

$$X \sim E_{p,n}(M_2,(A_2A_2') \otimes (B_2B_2'),\psi_2),$$

where $\psi_i(u) = \int_0^\infty \Omega_{p_1n_1}(y^2u)\, dF_i(y)$, and $F_i(y)$ denotes the distribution function of r_i, $i = 1,2$.

It follows, from Theorem 2.1.4, that $M_1 = M_2$, and there exist $a^2 > 0$, $b^2 > 0$, and $c^2 > 0$ such that $a^2b^2 = c^2$, $A_2A_2' = a^2A_1A_1'$, $B_2B_2' = b^2B_1B_1'$, and $\psi_2(z) = \psi_1\left(\frac{1}{c^2} z\right)$. Now, from Theorem 1.3.11, it follows that there exist $G \in O(p_1)$ and $H \in O(n_1)$ such that $A_2 = aA_1G$, and $B_2 = bB_1H$. Since, $\psi_2(z) = \psi_1\left(\frac{1}{c^2} z\right)$, we have

$$\begin{aligned} \psi_2(z) &= \int_0^\infty \Omega_{p_1n_1}(y^2u)\, dF_2(y) \\ &= \psi_1\left(\frac{z}{c^2}\right) \end{aligned}$$

$$= \int_0^\infty \Omega_{p_1 n_1}\left(y^2 \frac{z}{c^2}\right) dF_1(y)$$

$$= \int_0^\infty \Omega_{p_1 n_1}\left(\left(\frac{y}{c}\right)^2 z\right) dF_1(y)$$

$$= \int_0^\infty \Omega_{p_1, n_1}(t^2 z)\, dF_1(ct).$$

Therefore $F_2(y) = F_1(c\,y)$, and

$$P(r_2 < y) = P(r_1 < cy) = P\left(\frac{r_1}{c} < y\right).$$

Hence, $r_2 = \frac{1}{c} r_1$. ■

REMARK 2.5.1. It follows, from Theorem 2.5.2, that U does not depend on ψ. On the other hand, if p_1 and n_1 are fixed, ψ and R determine each other.

REMARK 2.5.2. Let $X \sim E_{p,n}(O, \Sigma \otimes \Phi, \psi)$ and $rAUB'$ be the stochastic representation of X. Then, $A^-XB'^- \approx rU$, and $\mathrm{tr}((A^-XB'^-)' (A^-XB'^-)) \approx \mathrm{tr}(r^2U'U)$. Now,

$$\begin{aligned} \mathrm{tr}((A^-XB'^-)' (A^-XB'^-)) &= \mathrm{tr}(B^-X'A'^-A^-XB'^-) \\ &= \mathrm{tr}(X'A'^-A^-XB'^-B^-) \\ &= \mathrm{tr}(X'\Sigma^-X\Phi^-). \end{aligned}$$

Here we used $A^{-\prime}A^- = \Sigma^-$, which follows from Theorem 1.3.23. On the other hand, $\mathrm{tr}(U'U) = 1$. Therefore, we get $r^2 \approx \mathrm{tr}(X'\Sigma^-X\Phi^-)$.

If an elliptically contoured random matrix is nonzero with probability one, then the terms of the stochastic representation can be obtained explicitly. First we introduce the following definition.

DEFINITION 2.5.1. Let X be a $p \times n$ matrix. Then its norm, denoted by $\|X\|$, is defined as $\|X\| = \left(\sum_{i=1}^{p} \sum_{j=1}^{n} x_{ij}^2 \right)^{\frac{1}{2}}$. That is, $\|X\| = (\mathrm{tr}(X'X))^{\frac{1}{2}}$, and if $n = 1$, then we have $\|\mathbf{x}\| = (\mathbf{x}'\mathbf{x})^{\frac{1}{2}}$.

The proof of the following theorem is based on Muirhead (1982).

THEOREM 2.5.4. *Let* $X \sim E_{p,n}(O, I_p \otimes I_n, \psi)$ *with* $P(X = O) = 0$. *Then,* $X = \|X\| \frac{X}{\|X\|}$, $P(\|X\| > 0) = 1$, $\mathrm{vec}\left(\frac{X}{\|X\|}\right)'$ *is uniformly distributed on* S_{pn}, *and* $\|X\|$ *and* $\frac{X}{\|X\|}$ *are independent. That is,* $\|X\| \frac{X}{\|X\|}$ *is the stochastic representation of* X.

PROOF: Since $X = O$ iff $\mathrm{tr}(X'X) = 0$, $P(\|X\| > 0) = 1$, follows so we can write $X = \|X\| \frac{X}{\|X\|}$. Define $\mathbf{x} = \mathrm{vec}(X')$. Then, $\mathbf{x} \sim E_{pn}(\mathbf{o}, I_{pn}, \psi)$ and $\|X\| = \|\mathbf{x}\|$. Hence, $\mathbf{x} = \|\mathbf{x}\| \frac{\mathbf{x}}{\|\mathbf{x}\|}$.

Let $T(\mathbf{x}) = \frac{\mathbf{x}}{\|\mathbf{x}\|}$, and $G \in O(pn)$. Then, we get $G\mathbf{x} \sim E_{pn}(\mathbf{o}, I_{pn}, \psi)$, so $\mathbf{x} \approx G\mathbf{x}$ and $T(\mathbf{x}) \approx T(G\mathbf{x})$. On the other hand,

$$T(G\mathbf{x}) = \frac{G\mathbf{x}}{\|G\mathbf{x}\|} = \frac{G\mathbf{x}}{\|\mathbf{x}\|} = GT(\mathbf{x})$$

hence, $T(\mathbf{x}) \approx GT(\mathbf{x})$. However, the uniform distribution is the only one on S_{pn} which is invariant under orthogonal transformation. So, $T(\mathbf{x})$ is uniformly distributed on S_{pn}.

Now, we define a measure μ on S_{pn}. Fix $B \subset \mathbb{R}_0^+$ Borel set. Let

$A \subset S_{pn}$ be a Borel set. Then,

$$\mu(A) = P(T(x) \in A \mid \|X\| \in B).$$

Since $\mu(\mathbb{R}^{pn}) = 1$, μ is a probability measure on S_{pn}.

Let $G \in O(pn)$. Then, $G^{-1}\mathbf{x} \approx \mathbf{x}$, and we have

$$\begin{aligned}
\mu(GA) &= P(T(\mathbf{x}) \in GA \mid \|\mathbf{x}\| \in B) \\
&= P(G^{-1}T(\mathbf{x}) \in A \mid \|\mathbf{x}\| \in B) \\
&= P(T(G^{-1}\mathbf{x}) \in A \mid \|\mathbf{x}\| \in B) \\
&= P(T(G^{-1}\mathbf{x}) \in A \mid \|G^{-1}\mathbf{x}\| \in B) \\
&= P(T(\mathbf{x}) \in A \mid \|\mathbf{x}\| \in B) \\
&= \mu(A).
\end{aligned}$$

Thus, $\mu(A)$ is a probability measure on S_{pn}, invariant under orthogonal transformation, therefore, it must be the uniform distribution. That is, it coincides with the distribution of $T(\mathbf{x})$. So, $\mu(A) = P(T(\mathbf{x}) \in A)$, from which it follows that

$$P(T(\mathbf{x}) \in A \mid \|\mathbf{x}\| \in B) = P(T(\mathbf{x}) \in A).$$

Therefore, $T(\mathbf{x})$ and $\|\mathbf{x}\|$ are independently distributed. Returning to the matrix notation, the proof is completed. ■

Muirhead (1982) has given the derivation of the p.d.f. of r in the case when $\mathbf{x} \sim E_p(\mathbf{o}, I_p, \psi)$ and $\mathbf{x}$ is absolutely continuous. Now for the elliptically contoured random matrices, the following theorem can be stated.

THEOREM 2.5.5. *Let* $X \sim E_{p,n}(O,\Sigma \otimes \Phi,\psi)$ *and* $rAUB'$ *be a stochastic representation of* X. *Assume* X *is absolutely continuous and has the p.d.f.*

$$f(X) = |\Sigma|^{-\frac{n}{2}} |\Phi|^{-\frac{p}{2}} h(tr(X'\Sigma^{-1}X\Phi^{-1})).$$

Then, r *is also absolutely continuous and has the p.d.f.*

$$g(r) = \frac{2\pi^{\frac{np}{2}}}{\Gamma\left(\frac{np}{2}\right)} r^{np-1} h(r^2), \quad r \geq 0.$$

PROOF: Step 1. First we prove the theorem for $n = 1$. Then, $x \sim E_p(o,\Sigma,\psi)$ and so

$$y = A^{-1}x \sim E_p(o,I_p,\psi).$$

Therefore y has the p.d.f. $h(y'y)$. Let us introduce polar coordinates:

$$y_1 = r \sin \theta_1 \sin \theta_2 \ldots \sin \theta_{p-2} \sin \theta_{p-1}$$

$$y_2 = r \sin \theta_1 \sin \theta_2 \ldots \sin \theta_{p-2} \cos \theta_{p-1}$$

$$y_3 = r \sin \theta_1 \sin \theta_2 \ldots \cos \theta_{p-2}$$

$$\vdots$$

$$y_{p-1} = r \sin \theta_1 \cos \theta_2$$

$$y_p = r \cos \theta_1,$$

where $r > 0$, $0 < \theta_i < \pi$, $i = 1,2,\ldots,p-2$, and $0 < \theta_{p-1} < 2\pi$. We want to express the p.d.f. of y in terms of $r,\theta_1,\ldots,\theta_{p-1}$. The Jacobian of the transformation $(y_1,y_2,\ldots,y_p) \to (r,\theta_1,\ldots,\theta_{p-1})$ is

$r^{p-1} \sin^{p-2} \theta_1 \sin^{p-3} \theta_2 \ldots \sin \theta_{p-2}$. On the other hand, $y'y = r^2$. Therefore, the p.d.f. of $(r,\theta_1,\ldots,\theta_{p-1})$ is $h(r^2)\, r^{p-1} \sin^{p-2} \theta_1 \sin^{p-3} \theta_2 \ldots \sin \theta_{p-3} \sin \theta_{p-2}$. Consequently, the p.d.f. of r is

$$g(r) = r^{p-1} h(r^2) \int_0^{2\pi} \int_0^{\pi} \ldots \int_0^{\pi} \sin^{p-2}\theta_1 \sin^{p-3}\theta_2 \ldots \sin^2\theta_{p-3} \sin \theta_{p-2}\, d\theta_1\, d\theta_2 \ldots d\theta_{p-2}\, d\theta_{p-1}$$

$$= r^{p-1} h(r^2) \frac{2\pi^{\frac{p}{2}}}{\Gamma\left(\frac{p}{2}\right)}.$$

Step 2. Now let $X \sim E_{p,n}(O,\Sigma \otimes \Phi,\psi)$, and $X \approx rAUB'$. Define $\mathbf{x} = \mathrm{vec}(X')$, and $\mathbf{u} = \mathrm{vec}(U')$. Then, $\mathbf{x} \sim E_{pn}(\mathbf{o},\Sigma \otimes \Phi,\psi)$, $\mathbf{x}$ has p.d.f. $\frac{1}{|\Sigma \otimes \Phi|} h(\mathbf{x}'(\Sigma \otimes \Phi)^{-1} \mathbf{x})$, and $\mathbf{x} \approx r(A \otimes B)\mathbf{u}$. Using Step 1 we get the following as the p.d.f. of r,

$$g(r) = r^{np-1} h(r^2) \frac{2\pi^{\frac{pn}{2}}}{\Gamma\left(\frac{pn}{2}\right)}. \blacksquare$$

The stochastic representation is a major tool in the study of m.e.c. distributions. It will often be used in further discussion.

Cambanis, Huang and Simons (1981), and Anderson and Fang (1987) derived the relationship between the stochastic representation of a multivariate elliptically contoured distribution and the stochastic representation of its marginals. This result is given in the next theorem.

THEOREM 2.5.6. *Let* $X \sim E_{p,n}(O,I_p \otimes I_n,\psi)$ *with stochastic representation* $X \approx rU$. *Let* X *be partitioned into*

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ \cdot \\ X_m \end{pmatrix},$$

where X_i *is* $p_i \times n$ *matrix,* $i = 1,\dots,m$. *Then,*

$$\begin{pmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ \cdot \\ X_m \end{pmatrix} \approx \begin{pmatrix} rr_1U_1 \\ rr_2U_2 \\ \cdot \\ \cdot \\ \cdot \\ rr_mU_m \end{pmatrix},$$

where r, $(r_1,r_2,\dots,r_m)$, $U_1,U_2,\dots,U_m$ *are independent,* $r_i \geq 0$, $i = 1,\dots,m$, $\sum_{i=1}^{m} r_i^2 = 1$,

$$(r_1^2,r_2^2,\dots,r_{m-1}^2) \sim D\left(\frac{p_1n}{2}, \frac{p_2n}{2}, \dots, \frac{p_{m-1}n}{2}; \frac{p_mn}{2}\right), \tag{2.11}$$

and $\mathrm{vec}(U_i')$ *is uniformly distributed on* S_{p_in}, $i = 1,2,\dots,m$.

PROOF: Since $X \approx rU$, we have

$$\begin{pmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ \cdot \\ X_m \end{pmatrix} \sim rU$$

where r and U are independent. Thus it suffices to prove that

$$U \approx \begin{pmatrix} r_1U_1 \\ r_2U_2 \\ \cdot \\ \cdot \\ \cdot \\ r_mU_m \end{pmatrix}.$$

Note that U does not depend on ψ, so we can choose $\psi(z) = \exp\left(-\frac{z}{2}\right)$, which means $X \sim N_{p,n}(O,I_p \otimes I_n)$. It follows that $X_i \sim N_{p_i,n}(O,I_{p_i} \otimes I_n)$ and X_i's are mutually independent, $i = 1,\ldots,m$.

Now, $U \approx \frac{X}{\|X\|} = \left(\frac{X_1'}{\|X\|}, \frac{X_2'}{\|X\|}, \ldots, \frac{X_m'}{\|X\|}\right)'$. From Theorem 2.5.4 it follows that $X_i = \|X_i\| \frac{X_i}{\|X_i\|}$, where $\|X_i\|$ and $\frac{X_i}{\|X_i\|}$ are independent and $\operatorname{vec}\left(\frac{X_i'}{\|X_i\|}\right) \sim \mathbf{u}_{p_in}$ which is uniformly distributed on S_{p_in}. Since, X_i's are independent, $\|X_i\|$ and $\frac{X_i}{\|X_i\|}$ are mutually independent, $i = 1,2,\ldots,m$. Therefore, we get

$$U \approx \left(\frac{\|X_1\|}{\|X\|}\frac{X_1'}{\|X_1\|}, \frac{\|X_2\|}{\|X\|}\frac{X_2'}{\|X_2\|}, \ldots, \frac{\|X_m\|}{\|X\|}\frac{X_m'}{\|X_m\|}\right).$$

Define $r_i = \frac{\|X_i\|}{\|X\|}$, and $U_i = \frac{X_i}{\|X_i\|}$, $i = 1,2,\ldots,m$. Since $\|X\| = \left(\sum_{i=1}^{m} \|X_i\|^2\right)^{\frac{1}{2}}$, r_i's are functions of $\|X_1\|,\|X_2\|,\ldots,\|X_m\|$. Hence, $(r_1,r_2,\ldots,r_m)$, $U_1,U_2,\ldots,U_m$ are independent. Moreover, $\|X_i\|^2 = \operatorname{tr}(X_i'X_i) \sim \chi^2_{p_in}$ and $\|X_i\|^2$'s are independent. Now, it is known that

$$\left(\frac{\|X_1\|^2}{\sum_{i=1}^{m} \|X_i\|^2}, \frac{\|X_2\|^2}{\sum_{i=1}^{m} \|X_i\|^2}, \ldots, \frac{\|X_{m-1}\|^2}{\sum_{i=1}^{m} \|X_i\|^2} \right) \sim D\left(\frac{p_1 n}{2}, \frac{p_2 n}{2}, \ldots, \frac{p_{m-1} n}{2}; \frac{p_m n}{2}\right)$$

(see Johnson and Kotz, 1972). Consequently, $(r_1^2, r_2^2, \ldots, r_{m-1}^2)$ has the distribution (2.11). ■

COROLLARY 2.5.6.1. *Let* $X \sim E_{p,n}(O, I_p \otimes I_n, \Psi\psi)$ *with stochastic representation* $X \approx rU$. *Let* X *be partitioned into*

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix},$$

where X_1 *is* $q \times n$ *matrix,* $1 \leq q < p$. *Then,* $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \approx \begin{pmatrix} rr_1U_1 \\ rr_2U_2 \end{pmatrix}$, *where* r, (r_1, r_2), U_1, U_2 *are independent,* $r_i \geq 0$, $i = 1,2$, $r_1^2 + r_2^2 = 1$, *and* $r_1^2 \sim B\left(\frac{qn}{2}, \frac{(p-q)n}{2}\right)$. *Also* $\mathrm{vec}(U_1')$ *is uniformly distributed on* S_{qn} *and* $\mathrm{vec}(U_2')$ *is uniformly distributed on* $S_{(p-q)n}$.

2.6. CONDITIONAL DISTRIBUTIONS

First, we derive the conditional distribution for the vector variate elliptically contoured distribution. We will follow the lines of Cambanis, Huang, and Simons (1981). The following lemma will be needed in the proof.

LEMMA 2.6.1. *Let* x *and* y *be one-dimensional nonnegative random variables. Assume that* y *is absolutely continuous with probability density function* $g(y)$. *Denote the distribution function of* x *by* $F(x)$. *Define* $z = xy$. *Then,* z *is absolutely continuous on* $\mathbb{R}^+$ *with p.d.f.*

$$h(z) = \int_{(0,\infty)} \frac{1}{x} g\left(\frac{z}{x}\right) dF(x). \tag{2.12}$$

If $F(0) = 0$, *then* z *is absolutely continuous on* $\mathbb{R}_0^+$, *and if* $F(0) > 0$, *then* z *has an atom of size* $F(0)$ *at zero. Moreover, a conditional distribution of* x *given* z *is*

$$P(x \le x_0 \mid z = z_0) = \begin{cases} \dfrac{1}{h(z_0)} \displaystyle\int_{(0,w]} \frac{1}{x_0} g\left(\frac{z_0}{x_0}\right) dF(x_0) & \textit{if } x_0 \ge 0,\ z_0 > 0, \textit{ and } h(z) \ne 0 \\ 1 & \textit{if } x_0 \ge 0, \textit{ and } z_0 = 0 \\ & \textit{or } x_0 \ge 0,\ z_0 > 0 \textit{ and } h(z_0) = 0 \\ 0 & \textit{if } x_0 < 0. \end{cases} \tag{2.13}$$

PROOF: $P(0 < z \le z_0) = P(0 < xy \le z_0)$

$$= \int_{(0,\infty)} P(o < xy \le z_0 \mid x = x_0)\, dF(x_0)$$

$$= \int_{(0,\infty)} P(o < y \le \frac{z_0}{x})\, dF(x_0)$$

$$= \int_{(0,\infty)} \int_{(0,\frac{z_0}{x_0}]} g(y)\, dy\, dF(x_0).$$

Let $t = x_0 y$. Then, $y = \frac{t}{x_0}$, $dy = \frac{1}{x_0} dt$ and

$$\int_{(0,\infty)} \int_{(0,\frac{z_0}{x_0}]} g(y)\, dy\, dF(x_0) = \int_{(0,\infty)} \int_{(0,z_0]} \frac{1}{x_0} g\left(\frac{t}{x_0}\right) dt\, dF(x_0)$$

$$= \int_{(0,z_0]} \int_{(0,\infty)} \frac{1}{x_0} g\left(\frac{t}{x_0}\right) dF(x_0)\, dt$$

and this proves (2.12).

Since, y is absolutely continuous, $P(y = 0) = 0$. Hence,

$$P(\chi_{\{0\}}(z) = \chi_{\{0\}}(x)) = 1. \tag{2.14}$$

Therefore, if $F(0) = 0$, then $P(z = 0) = 0$, and so z is absolutely continuous on $\mathbb{R}_0^+$. If $F(0) > 0$, then $P(z = 0) = F(0)$ and thus z has an atom of size $F(0)$ at zero.

Now, we prove (2.13). Since $x \geq 0$, we have $P(x \leq x_0) = 0$ if $x_0 < 0$. Hence, $P(x \leq x_0 \mid z) = 0$ if $x_0 < 0$. If $x_0 \geq 0$, we have to prove that the function $P(x \leq x_0 \mid z)$ defined under (2.13) satisfies

$\int_{[0,r]} P(x \leq x_0 \mid z)\, dH(z) = P(x \leq x_0, z \leq r)$, where $H(z)$ denotes the distribution function of z and $r \geq 0$. Now,

$$\begin{aligned}\int_{[0,r]} P(x \leq x_0 \mid z)\, dH(z) &= P(x \leq x_0 \mid z = 0)\, H(0) + \int_{(0,r]} P(x \leq x_0 \mid z)\, dH(z) \\ &= H(0) + \int_{(0,r]} \frac{1}{h(z)} \left(\int_{(0,x_0]} \frac{1}{x}\, g\left(\frac{z}{x}\right) dF(x) \right) h(z)\, dz \\ &= H(0) + \int_{(0,x_0]} \int_{(0,r]} \frac{1}{x} g\left(\frac{z}{x}\right) dz\, dF(x).\end{aligned}$$

Let $u = \frac{z}{x}$. Then, $J(u \to z) = x$, and so

$$\int_{(0,r]} \frac{1}{x} g\left(\frac{z}{x}\right) dz = \int_{(0,\frac{r}{x}]} g(u)\, du = P\left(0 < y \leq \frac{r}{x}\right).$$

Hence,

$$H(0) + \int_{(0,x_0]}\int_{(0,r]} \frac{1}{x} g\left(\frac{z}{x}\right) dz\, dF(x) = H(0) + \int_{(0,x_0]} P(o < y \le \frac{r}{x} \mid x = x_0)\, dF(x_0)$$

$$= H(0) + P(0 < x \le x_0, 0 < y \le \frac{r}{x})$$

$$= H(0) + P(0 < x \le x_0, 0 < xy \le r)$$

$$= P(z = 0) + P(0 \le x \le x_0, 0 < z \le r)$$

$$= P(x = 0, z = 0) + P(0 < x \le x_0, 0 < z \le r)$$

$$= P(0 \le x \le x_0, 0 \le z \le r)$$

$$= P(x \le x_0, z \le r),$$

where we used (2.14). ■

Now, we obtain the conditional distribution for spherical distributions.

THEOREM 2.6.1. *Let* $\mathbf{x} \sim E_p(\mathbf{o}, I_p, \psi)$ *with stochastic representation* $r\mathbf{u}$. *Let us partition* $\mathbf{x}$ *as* $\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}$, *where* $\mathbf{x}_1$ *is q-dimensional* $(1 \le q < p)$. *Then, the conditional distribution of* $\mathbf{x}_1$ *given* $\mathbf{x}_2$ *is* $(\mathbf{x}_1 \mid \mathbf{x}_2) \sim E_q(\mathbf{o}, I_q, \psi_{\|\mathbf{x}_2\|^2})$, *and the stochastic representation of* $(\mathbf{x}_1 \mid \mathbf{x}_2)$ *is* $r_{\|\mathbf{x}_2\|^2}\mathbf{u}_1$, *where* $\mathbf{u}_1$ *is q-dimensional. The distribution of* $r_{\|\mathbf{x}_2\|^2}$ *is given by*

$$a) \qquad P(r_{a^2} \le y) = \frac{\int_{(a,\sqrt{a^2+y^2}]} (w^2 - a^2)^{\frac{q}{2}-1} w^{-(p-2)}\, dF(w)}{\int_{(a,\infty)} (w^2 - a^2)^{\frac{q}{2}-1} w^{-(p-2)}\, dF(w)} \qquad (2.15)$$

for $y \geq 0$ *if* $a > 0$ *and* $F(a) < 1$,

$$b) \qquad P(r_{a^2} = 0) = 1 \text{ if } a = 0 \text{ or } F(a) = 1. \tag{2.16}$$

Here F *denotes the distribution function of* r.

PROOF: From Corollary 2.5.6.1, we have the representation

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \approx \begin{pmatrix} rr_1 u_1 \\ rr_2 u_2 \end{pmatrix}.$$

Using the independence of r, r_1, u_1 and u_2, we get

$$(x_1 | x_2) \approx (rr_1 u_1 | rr_2 u_2 = x_2)$$

$$= (rr_1 u_1 | r(1 - r_1^2)^{\frac{1}{2}} u_2 = x_2)$$

$$= (rr_1 | r(1 - r_1^2)^{\frac{1}{2}} u_2 = x_2) u_1,$$

and defining $r_0 = (rr_1 | r(1 - r_1^2)^{\frac{1}{2}} u_2 = x_2)$, we see that r and u_1 are independent; therefore, $(x_1 | x_2)$ has a spherical distribution.

Next, we show that

$$(rr_1 | r(1 - r_1^2)^{\frac{1}{2}} u_2 = x_2) \approx ((r^2 - \|x_2\|)^{\frac{1}{2}} | r(1 - r_1^2)^{\frac{1}{2}} = \|x_2\|).$$

If $r(1 - r_1^2)^{\frac{1}{2}} u_2 = x_2$, then $(r(1 - r_1^2)^{\frac{1}{2}} u_2)' r(1 - r_1^2)^{\frac{1}{2}} u_2 = \|x_2\|^2$ and therefore,

$r^2(1 - r_1^2) = \|x_2\|^2$. Hence, we get $r^2 - r^2 r_1^2 = \|x_2\|^2$, thus $r^2 r_1^2 = r^2 - \|x_2\|^2$ and

$rr_1 = (r^2 - \|\mathbf{x}_2\|^2)^{\frac{1}{2}}$. Therefore,

$$(rr_1 \mid r(1 - r_1^2)^{\frac{1}{2}} \mathbf{u}_2 = \mathbf{x}_2) \approx ((r^2 - \|\mathbf{x}_2\|^2)^{\frac{1}{2}} \mid r(1 - r_1^2)^{\frac{1}{2}} \mathbf{u}_2 = \mathbf{x}_2).$$

If $\mathbf{x}_2 = \mathbf{o}$, then $\|\mathbf{x}_2\| = 0$, and using the fact that $\mathbf{u}_1 \neq \mathbf{o}$ we get

$$\begin{aligned}((r^2 - \|\mathbf{x}_2\|^2)^{\frac{1}{2}} \mid r(1 - r_1^2)^{\frac{1}{2}} \mathbf{u}_2 = \mathbf{x}_2) &= ((r^2 - \|\mathbf{x}_2\|^2)^{\frac{1}{2}} \mid r(1 - r_1^2)^{\frac{1}{2}} = 0) \\ &= ((r^2 - \|\mathbf{x}_2\|^2)^{\frac{1}{2}} \mid r(1 - r_1^2)^{\frac{1}{2}} = \|\mathbf{x}_2\|).\end{aligned}$$

If $\mathbf{x}_2 \neq \mathbf{o}$, then we can write

$$\begin{aligned}&((r^2 - \|\mathbf{x}_2\|^2)^{\frac{1}{2}} \mid r(1 - r_1^2)^{\frac{1}{2}} \mathbf{u}_2 = \mathbf{x}_2) \\ &= \left((r^2 - \|\mathbf{x}_2\|^2)^{\frac{1}{2}} \mid r(1 - r_1^2)^{\frac{1}{2}} \mathbf{u}_2 = \|\mathbf{x}_2\| \frac{\mathbf{x}_2}{\|\mathbf{x}_2\|} \right) \\ &= \left((r^2 - \|\mathbf{x}_2\|^2)^{\frac{1}{2}} \mid r(1 - r_1^2)^{\frac{1}{2}} = \|\mathbf{x}_2\| \text{ and } \mathbf{u}_2 = \frac{\mathbf{x}_2}{\|\mathbf{x}_2\|} \right) \\ &= ((r^2 - \|\mathbf{x}_2\|^2)^{\frac{1}{2}} \mid r(1 - r_1^2)^{\frac{1}{2}} = \|\mathbf{x}_2\|)\end{aligned}$$

where we used the fact that r, r_1, and $\mathbf{u}_2$ are independent. Since

$1 - r_1^2 \sim B\left(\frac{p-q}{2}, \frac{q}{2}\right)$ its p.d.f. is

$$b(t) = \frac{\Gamma\left(\frac{p}{2}\right)}{\Gamma\left(\frac{q}{2}\right)\Gamma\left(\frac{p-q}{2}\right)} t^{\frac{p-q}{2}-1} (1-t)^{\frac{q}{2}-1}, \quad 0 < t < 1.$$

Hence, the p.d.f. of $(1 - r_1^2)^{\frac{1}{2}}$ is

$$g(y) = \frac{\Gamma\left(\frac{p}{2}\right)}{\Gamma\left(\frac{q}{2}\right)\Gamma\left(\frac{p-q}{2}\right)} (y^2)^{\frac{p-q}{2}-1} (1-y^2)^{\frac{q}{2}-1} 2y$$

$$= \frac{2\Gamma\left(\frac{p}{2}\right)}{\Gamma\left(\frac{q}{2}\right)\Gamma\left(\frac{p-q}{2}\right)} y^{p-q-1} (1-y^2)^{\frac{q}{2}-2}, \quad 0 < y < 1.$$

Using Lemma 2.6.1 we obtain a conditional distribution of r given $r(1 - r_1^2)^{\frac{1}{2}} = a$.

$P(r \le u \mid r(1 - r_1^2)^{\frac{1}{2}} = a)$

$$= \begin{cases} \dfrac{\displaystyle\int_{(a,u]} \frac{1}{w} \frac{2\Gamma\left(\frac{p}{2}\right)}{\Gamma\left(\frac{q}{2}\right)\Gamma\left(\frac{p-q}{2}\right)} \left(\frac{a}{w}\right)^{p-q-1} \left(1 - \frac{a^2}{w^2}\right)^{\frac{q}{2}-1} dF(w)}{h(a)} & \text{if } u \ge 0,\ a > 0 \text{ and } h(a) \ne 0 \\ 1 & \text{if } u \ge 0, \text{ and } a = 0 \text{ or} \\ & u \ge 0,\ a > 0 \text{ and } h(a) = 0 \\ 0 & \text{if } u < 0, \end{cases} \tag{2.17}$$

where

$$h(a) = \int_{(a,\infty)} \frac{1}{w} \frac{2\Gamma\left(\frac{p}{2}\right)}{\Gamma\left(\frac{q}{2}\right)\Gamma\left(\frac{p-q}{2}\right)} \left(\frac{a}{w}\right)^{p-q-1} \left(1 - \frac{a^2}{w^2}\right)^{\frac{q}{2}-1} dF(w).$$

Now,

$$\frac{\displaystyle\int_{(a,u]} \frac{1}{w} \frac{2\Gamma\left(\frac{p}{2}\right)}{\Gamma\left(\frac{q}{2}\right)\Gamma\left(\frac{p-q}{2}\right)} \left(\frac{a}{w}\right)^{p-q-1} \left(1 - \frac{a^2}{w^2}\right)^{\frac{q}{2}-1} dF(w)}{h(a)}$$

$$= \frac{\displaystyle\int_{(a,u]} (w^2 - a^2)^{\frac{q}{2}-1}\, w^{-(1+p-q-1+q-2)}\, dF(w)}{\displaystyle\int_{(a,\infty)} (w^2 - a^2)^{\frac{q}{2}-1}\, w^{-(1+p-q-1+q-2)}\, dF(w)}$$

$$= \frac{\displaystyle\int_{(a,u]} (w^2 - a^2)^{\frac{q}{2}-1}\, w^{-(p-2)}\, dF(w)}{\displaystyle\int_{(a,\infty)} (w^2 - a^2)^{\frac{q}{2}-1}\, w^{-(p-2)}\, dF(w)}. \tag{2.18}$$

We note that $h(a) = 0$ if and only if

$$\int_{(a,\infty)} (w^2 - a^2)^{\frac{q}{2}-1}\, w^{-(p-2)}\, dF(w) = 0, \tag{2.19}$$

and since $(w^2 - a^2)^{\frac{q}{2}-1} w^{-(p-2)} > 0$ for $w > a$, we see that (2.19) is equivalent to $F(a) = 1$. Therefore, $h(a) = 0$ if and only if $F(a) = 1$.

a) If $a > 0$ and $F(a) < 1$, then for $r \geq 0$ we have

$$P(r_{a2} \le y) = P((r^2 - a^2)^{\frac{1}{2}} \le y \mid r(1 - r_1^2)^{\frac{1}{2}} = a)$$

$$= P((r \le (y^2 + a^2)^{\frac{1}{2}} \mid r(1 - r_1^2)^{\frac{1}{2}} = a). \tag{2.20}$$

From (2.17), (2.18), and (2.20) we have

$$P(r_{a2} \le y) = P((r \le (y^2 + a^2)^{\frac{1}{2}} \mid r(1 - r_1^2)^{\frac{1}{2}} = a)$$

$$= \frac{\int_{(a,\sqrt{a^2+y^2}]} (w^2 - a^2)^{\frac{q}{2}-1} w^{-(p-2)} dF(w)}{\int_{(a,\infty)} (w^2 - a^2)^{\frac{q}{2}-1} w^{-(p-2)} dF(w)}.$$

b) If $r \ge 0$ and $a = 0$ or $r \ge 0$, $a > 0$ and $F(a) = 1$, then from (2.17) we get

$$P(r \le (y^2 + a^2)^{\frac{1}{2}} \mid r(1 - r_1^2)^{\frac{1}{2}} = a) = 1.$$

Take $y = 0$, then we get

$$P(r_{a^2} \le 0) = P(r \le a \mid r(1 - r_1^2)^{1/2} = a) = 1. \tag{2.21}$$

Now, since $r_{a2} \ge 0$, (2.21) implies $P(r_{a2} = 0) = 1$. ■

In order to derive the conditional distribution for the multivariate elliptical distribution we need an additional lemma.

LEMMA 2.6.2. *Let* $\mathbf{x} \sim E_p(\mathbf{m},\Sigma,\psi)$ *and partition* $\mathbf{x}, \mathbf{m}, \Sigma$ *as* $\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}$, $\mathbf{m} = \begin{pmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{pmatrix}$, *and* $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, *where* $\mathbf{x}_1, \mathbf{m}_1$ *are q-dimensional vectors and* Σ_{11} *is* $q \times q$, $1 \le q < p$. *Let* $\mathbf{y} \sim E_p(\mathbf{o},I_p,\psi)$ *and partition* $\mathbf{y}$ *as* $\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}$, *where* $\mathbf{y}_1$ *is*

q-dimensional. Define $\Sigma_{11\cdot 2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-}\Sigma_{21}$ *and let* $\Sigma_{11\cdot 2} = AA'$ *and* $\Sigma_{22} = A_2A_2'$ *be rank factorizations of* $\Sigma_{11\cdot 2}$ *and* Σ_{22}. *Then*

$$\begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \approx \begin{pmatrix} \mathbf{m}_1 + A\mathbf{y}_1 + \Sigma_{12}\Sigma_{22}^{-}A_2\mathbf{y}_2 \\ \mathbf{m}_2 + A_2\mathbf{y}_2 \end{pmatrix}.$$

PROOF: Since $\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} \sim E_p(\mathbf{o},I_p,\psi)$, we have

$$\begin{pmatrix} \mathbf{m}_1 + A\mathbf{y}_1 + \Sigma_{12}\Sigma_{22}^{-}A_2\mathbf{y}_2 \\ \mathbf{m}_2 + A_2\mathbf{y}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{pmatrix} + \begin{pmatrix} A & \Sigma_{12}\Sigma_{22}^{-}A_2 \\ O & A_2 \end{pmatrix} \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}$$

$$\sim E_p\left(\begin{pmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{pmatrix}, \begin{pmatrix} A & \Sigma_{12}\Sigma_{22}^{-}A_2 \\ O & A_2 \end{pmatrix} \begin{pmatrix} A' & O \\ A_2'\Sigma_{22}^{-}\Sigma_{21} & A_2' \end{pmatrix}, \psi \right)$$

$$= E_p\left(\mathbf{m}, \begin{pmatrix} AA' + \Sigma_{12}\Sigma_{22}^{-}A_2A_2'\Sigma_{22}^{-}\Sigma_{21} & \Sigma_{12}\Sigma_{22}^{-}A_2A_2' \\ A_2A_2'\Sigma_{22}^{-}\Sigma_{21} & A_2A_2' \end{pmatrix}, \psi \right)$$

$$= E_p\left(\mathbf{m}, \begin{pmatrix} \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-}\Sigma_{21} + \Sigma_{12}\Sigma_{22}^{-}\Sigma_{22}\Sigma_{22}^{-}\Sigma_{21} & \Sigma_{12}\Sigma_{22}^{-}\Sigma_{22} \\ \Sigma_{22}\Sigma_{22}^{-}\Sigma_{21} & \Sigma_{22} \end{pmatrix}, \psi \right). \tag{2.22}$$

Now we prove that $\Sigma_{12}\Sigma_{22}^{-}\Sigma_{22} = \Sigma_{12}$. If Σ_{22} is of the form $\Sigma_{22} = \begin{pmatrix} L & O \\ O & O \end{pmatrix}$, where L is a nonsingular, diagonal $s \times s$ matrix, then Σ_{12} must be of the form $\Sigma_{12} = (K,O)$ where K is $q \times s$. Indeed, otherwise there would be numbers i and j such that
$1 \le i \le q$ and $q + s \le j \le p$ and $\sigma_{ij} \ne 0$, $\sigma_{jj} = 0$. Since $\Sigma_{22} \ge O$, we must have $\begin{vmatrix} \sigma_{ii} & \sigma_{ij} \\ \sigma_{ji} & \sigma_{jj} \end{vmatrix} \ge 0$. However with $\sigma_{ij} \ne 0$, $\sigma_{jj} = 0$, we have $\begin{vmatrix} \sigma_{ii} & \sigma_{ij} \\ \sigma_{ji} & \sigma_{jj} \end{vmatrix} = -\sigma_{ij}^2 < 0$

which is a contradiction. Therefore, $\Sigma_{12} = (K,O)$.

Let Σ_{22}^{-} be partitioned as

$$\Sigma_{22}^{-} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

By the definition of a generalized inverse matrix, we must have

$$\begin{pmatrix} L & O \\ O & O \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} L & O \\ O & O \end{pmatrix} = \begin{pmatrix} L & O \\ O & O \end{pmatrix},$$

which gives

$$\begin{pmatrix} LAL & O \\ O & O \end{pmatrix} = \begin{pmatrix} L & O \\ O & O \end{pmatrix}.$$

So $LAL = L$, and since L is nonsingular, we get $A = L^{-1}$. Thus $\Sigma_{22}^{-} = \begin{pmatrix} L^{-1} & B \\ C & D \end{pmatrix}$. Then,

$$\begin{aligned} \Sigma_{12}\Sigma_{22}^{-}\Sigma_{22} &= (K,O)\begin{pmatrix} L^{-1} & B \\ C & D \end{pmatrix}\begin{pmatrix} L & O \\ O & O \end{pmatrix} \\ &= (K,O)\begin{pmatrix} I_s & O \\ CL & O \end{pmatrix} \\ &= (K,O) \\ &= \Sigma_{12}. \end{aligned}$$

If Σ_{22} is not of the form $\Sigma_{22} = \begin{pmatrix} L & O \\ O & O \end{pmatrix}$, then there exists a $G \in O(p - q)$ such that $G\Sigma_{22}G' = \begin{pmatrix} L & O \\ O & O \end{pmatrix}$. Now, define

$$\Sigma^* = \begin{pmatrix} I_q & O \\ O & G \end{pmatrix}\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\begin{pmatrix} I_q & O \\ O & G' \end{pmatrix}$$

$$= \begin{pmatrix} \Sigma_{11} & \Sigma_{12}G' \\ G\Sigma_{21} & G\Sigma_{22}G' \end{pmatrix}.$$

Then, we must have

$$\Sigma_{12}G'(G\Sigma_{22}G')^{-}\,(G\Sigma_{22}G') = \Sigma_{12}G'.$$

That is, $\Sigma_{12}G'G\Sigma_{22}^{-}G'G\Sigma_{22}G' = \Sigma_{12}G'$, which is equivalent to $\Sigma_{12}\Sigma_{22}^{-}\Sigma_{22} = \Sigma_{12}$. Using $\Sigma_{12}\Sigma_{22}^{-}\Sigma_{22} = \Sigma_{12}$ in (2.21), we have

$$\begin{pmatrix} m_1 + Ay_1 + \Sigma_{12}\Sigma_{22}^{-}A_2y_2 \\ m_2 + A_2y_2 \end{pmatrix} \sim E_p\left(m, \begin{pmatrix} \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-}\Sigma_{21} + \Sigma_{12}\Sigma_{22}^{-}\Sigma_{21} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{pmatrix}, \psi \right)$$

$$= E_p\left(m, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \psi \right)$$

$$= E_p(m, \Sigma, \psi)$$

which is the distribution of x. ■

Next, we give the conditional distribution of the multivariate elliptical distribution in two different forms.

THEOREM 2.6.2. *Let* $\mathbf{x} \sim E_p(\mathbf{m},\Sigma,\psi)$ *with stochastic representation* $r\mathbf{A}\mathbf{u}$. *Let* F *be the distribution function of* r. *Partition* $\mathbf{x}$, $\mathbf{m}$, *and* Σ *as*

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}, \quad \mathbf{m} = \begin{pmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{pmatrix}, \quad \text{and } \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

where $\mathbf{x}_1, \mathbf{m}_1$ *are* $q \times 1$, *and* Σ_{11} *is* $q \times q$, $1 \le q < p$. *Assume* $\mathrm{rk}(\Sigma_{22}) \ge 1$. *Then, a conditional distribution of* $\mathbf{x}_1$ *given* $\mathbf{x}_2$ *is*
$(\mathbf{x}_1 | \mathbf{x}_2) \sim E_q(\mathbf{m}_1 + \Sigma_{12}\Sigma_{22}^{-}(\mathbf{x}_2 - \mathbf{m}_2), \Sigma_{11\cdot 2}, \psi_{q(\mathbf{x}_2)})$, *where*

$\Sigma_{11\cdot 2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-}\Sigma_{21}$, $q(\mathbf{x}_2) = (\mathbf{x}_2 - \mathbf{m}_2)'\,\Sigma_{22}^{-}(\mathbf{x}_2 - \mathbf{m}_2)$, *and*

$$\psi_{q(\mathbf{x}_2)}(u) = \int_0^\infty \Omega_q(r^2 u)\, dF_{q(\mathbf{x}_2)}(r), \tag{2.23}$$

where

$$a) \quad F_{q(\mathbf{x}_2)}(r) = \frac{\displaystyle\int_{(\sqrt{q(\mathbf{x}_2)},\, \sqrt{q(\mathbf{x}_2) + r^2}\,]} (w^2 - q(\mathbf{x}_2))^{\frac{q}{2}-1}\, w^{-(p-2)}\, dF(w)}{\displaystyle\int_{(\sqrt{q(\mathbf{x}_2)},\infty)} (w^2 - q(\mathbf{x}_2))^{\frac{q}{2}-1}\, w^{-(p-2)}\, dF(w)} \tag{2.24}$$

for $r \ge 0$ *if* $q(\mathbf{x}_2) > 0$ *and* $F(\sqrt{q(\mathbf{x}_2)}) < 1$, *and*

$$b) \quad F_{q(\mathbf{x}_2)}(r) = 1 \ \textit{for } r \ge 0 \ \textit{if } q(\mathbf{x}_2) = 0 \ \textit{or } F(\sqrt{q(\mathbf{x}_2)}) = 1. \tag{2.25}$$

PROOF: From Lemma 2.6.2, we get

$$\begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \sim \begin{pmatrix} \mathbf{m}_1 + \mathbf{A}\mathbf{y}_1 + \Sigma_{12}\Sigma_{22}^{-}\mathbf{A}_2\mathbf{y}_2 \\ \mathbf{m}_2 + \mathbf{A}_2\mathbf{y}_2 \end{pmatrix},$$

where $AA' = \Sigma_{11\cdot 2}$ and $A_2A_2' = \Sigma_{22}$ are rank factorizations of $\Sigma_{11\cdot 2}$, Σ_{22}, and $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \sim E_p o, I_p, \psi)$. Thus,

$$\begin{aligned}(x_1 \mid x_2) &\approx (m_1 + Ay_1 + \Sigma_{12}\Sigma_{22}^{-}A_2y_2 \mid m_2 + A_2y_2 = x_2) \\ &= m_1 + \Sigma_{12}\Sigma_{22}^{-}(A_2y_2 \mid m_2 + A_2y_2 = x_2) + A(y_1 \mid m_2 + A_2y_2 = x_2) \\ &= m_1 + \Sigma_{12}\Sigma_{22}^{-}(A_2y_2 \mid A_2y_2 = x_2 - m_2) + A(y_1 \mid A_2y_2 = x_2 - m_2) \\ &= m_1 + \Sigma_{12}\Sigma_{22}^{-}(x_2 - m_2) + A(y_1 \mid A_2^{-}A_2y_2 = A_2^{-}(x_2 - m_2)). \qquad (2.26)\end{aligned}$$

Now $A_2^{-}A_2 = I_{rk(\Sigma_{22})}$, and hence we get

$$(y_1 \mid A_2^{-}A_2y_2 = A_2^{-}(x_2 - m_2)) = (y_1 \mid y_2 = A_2^{-}(x_2 - m_2)). \qquad (2.27)$$

From Theorem 2.6.1, we get

$$(y_1 \mid y_2 = A_2^{-}(x_2 - m_2)) \sim E_q(o, I_q, \psi_{q(x_2)}),$$

where

$$\begin{aligned}q(x_2) &= (A_2^{-}(x_2 - m_2))' A_2^{-}(x_2 - m_2) \\ &= (x_2 - m_2)' A_2^{-\prime} A_2^{-}(x_2 - m_2) \\ &= (x_2 - m_2)' \Sigma_{22}^{-}(x_2 - m_2)\end{aligned}$$

and $\psi_{q(x_2)}(r)$ is defined by (2.23), (2.24) and (2.25). Thus,

$$A(y_1 \mid y_2 = A_2^{-}(x_2 - m_2)) \sim E_q(o, AA', \psi_{q(x_2)})$$

$$= E_q(\mathbf{o}, \Sigma_{11\cdot 2}, \psi_{q(\mathbf{x}_2)}). \tag{2.28}$$

Finally, from (2.26), (2.27), and (2.28) we get

$$(\mathbf{x}_1 | \mathbf{x}_2) \sim E_q(\mathbf{m}_1 + \Sigma_{12}\Sigma_{22}^{-}(\mathbf{x}_2 - \mathbf{m}_2), \Sigma_{11\cdot 2}, \psi_{q(\mathbf{x}_2)}). \blacksquare$$

Another version of the conditional distribution is given in the following theorem.

THEOREM 2.6.3. *Let* $\mathbf{x} \sim E_p(\mathbf{m}, \Sigma, \psi)$ *with stochastic representation* $r\mathbf{A}\mathbf{u}$. *Let* F *be the distribution function of* r. *Partition* $\mathbf{x}, \mathbf{m}, \Sigma$ *as*

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}, \quad \mathbf{m} = \begin{pmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{pmatrix}, \quad \text{and} \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

where $\mathbf{x}_1, \mathbf{m}_1$ *are q-dimensional vectors, and* Σ_{11} *is* $q \times q$, $1 \le q < p$. *Assume* $\mathrm{rk}(\Sigma_{22}) \ge 1$. *Let* S *denote the subspace of* $\mathbb{R}^{p-q}$ *defined by the columns of* Σ_{22}; *that is,* $\mathbf{y} \in S$, *if there exists* $\mathbf{a} \in \mathbb{R}^{p-q}$ *such that* $\mathbf{y} = \Sigma_{22}\mathbf{a}$.

Then, a conditional distribution of $\mathbf{x}_1$ *given* $\mathbf{x}_2$ *is*

a) $(\mathbf{x}_1 | \mathbf{x}_2) \sim E_q(\mathbf{m}_1 + \Sigma_{12}\Sigma_{22}^{-}(\mathbf{x}_2 - \mathbf{m}_2), \Sigma_{11\cdot 2}, \psi_{q(\mathbf{x}_2)})$

for $\mathbf{x}_2 \in \mathbf{m}_2 + S$, *where* $q(\mathbf{x}_2) = (\mathbf{x}_2 - \mathbf{m}_2)' \Sigma_{22}^{-}(\mathbf{x}_2 - \mathbf{m}_2)$ *and* $\psi_{q(\mathbf{x}_2)}$ *is defined by (2.23), (2.24) and (2.25).*

b) $(\mathbf{x}_1 | \mathbf{x}_2) = \mathbf{m}_1$ *for* $\mathbf{x}_2 \notin \mathbf{m}_2 + S$.

PROOF: It suffices to prove that $P(\mathbf{x}_2 \notin \mathbf{M}_2 + S) = 0$ since $(\mathbf{x}_1 | \mathbf{x}_2)$ can be arbitrarily defined for $\mathbf{x}_2 \in B$ where B is of measure zero. However, $P(\mathbf{x}_2 \notin \mathbf{m}_2 + S) = o$ is equivalent to $P(\mathbf{x}_2 \in \mathbf{m}_2 + S) = 1$; that is, $P(\mathbf{x}_2 - \mathbf{m}_2 \in S) = 1$. Now, $\mathbf{x}_2 \sim E_{p-q}(\mathbf{m}_2, \Sigma_{22}, \psi)$ and so $\mathbf{x}_2 - \mathbf{m}_2 \sim E_{p-q}(\mathbf{o}, \Sigma_{22}, \psi)$.

Let $k = \mathrm{rk}(\Sigma_{22})$. Let $\mathbf{G} \in O(p-q)$ such that $\mathbf{G}\Sigma_{22}\mathbf{G}' = \begin{pmatrix} \mathbf{L} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}$, where $\mathbf{L}$ is a diagonal and nonsingular $k \times k$ matrix and define $\mathbf{y} = \mathbf{G}(\mathbf{x}_2 - \mathbf{m}_2)$. Then,

$$\mathbf{y} \sim E_{p-q}\left(\mathbf{o}, \begin{pmatrix} L & O \\ O & O \end{pmatrix}, \psi\right). \tag{2.29}$$

Partition $\mathbf{y}$ as $\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}$, where $\mathbf{y}_1$ is $k \times 1$. We have

$$P(\mathbf{x}_2 - \mathbf{m}_2 \in S) = P(\mathbf{x}_2 - \mathbf{m}_2 = \Sigma_{22}\mathbf{a} \text{ with } \mathbf{a} \in \mathbb{R}^{p-q})$$

$$= P(G(\mathbf{x}_2 - \mathbf{m}_2) = G\Sigma_{22}G'G\mathbf{a} \text{ with } \mathbf{a} \in \mathbb{R}^{p-q})$$

$$= P\left(\mathbf{y} = \begin{pmatrix} L & O \\ O & O \end{pmatrix}\mathbf{b} \text{ with } \mathbf{b} \in \mathbb{R}^{p-q}\right)$$

$$= P\left(\mathbf{y} = \begin{pmatrix} L & O \\ O & O \end{pmatrix}\begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix} \text{ with } \mathbf{b}_1 \in \mathbb{R}^k,\ \mathbf{b}_2 \in \mathbb{R}^{p-q-k}\right)$$

$$= P\left(\mathbf{y} = \begin{pmatrix} L\mathbf{b}_1 \\ \mathbf{o} \end{pmatrix} \text{ with } \mathbf{b}_1 \in \mathbb{R}^k\right)$$

$$= P\left(\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{c} \\ \mathbf{o} \end{pmatrix} \text{ with } \mathbf{c} \in \mathbb{R}^k\right)$$

$$= P(\mathbf{y}_2 = \mathbf{o}).$$

Now, it follows from (2.29) that $\mathbf{y}_2 \sim E_{p-q-k}(\mathbf{o}, O, \psi)$ and so $P(\mathbf{y}_2 = \mathbf{o}) = 1$. Therefore, $P(\mathbf{x}_2 - \mathbf{m}_2 \in S) = 1$. ■

Now we can derive the conditional distribution for m.e.c. distributions.

THEOREM 2.6.4. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ *with stochastic representation* $rAUB'$. *Let* F *be the distribution function of* r. *Partition* X, M, *and* Σ *as*

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix},\quad M = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix},\quad \textit{and}\quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

where X_1 *is* $q \times n$, M_1 *is* $q \times n$, *and* Σ_{11} *is* $q \times q$, $1 \le q < p$. *Assume*

$rk(\Sigma_{22}) \geq 1$.

Let S *denote the subspace of* $\mathbb{R}^{(p-q)n}$ *defined by the columns of* $\Sigma_{22} \otimes \Phi$; *that is,* $\mathbf{y} \in S$, *if there exists* $\mathbf{b} \in \mathbb{R}^{(p-q)n}$ *such that* $\mathbf{y} = (\Sigma_{22} \otimes \Phi)\mathbf{b}$. *Then, a conditional distribution of* X_1 *given* X_2 *is*

1) $(X_1 | X_2) \sim E_{q,n}(M_1 + \Sigma_{12}\Sigma_{22}^{-}(X_2 - M_2), \Sigma_{11\cdot 2} \otimes \Phi, \psi_{q(x_2)})$,

for $\text{vec}(X_2') \in \text{vec}(M_2') + S$, *where* $q(X_2) = \text{tr}((X_2 - M_2)' \Sigma_{22}^{-}(X_2 - M_2)\Phi^{-})$,

$$\psi_{q(x_2)}(u) = \int_0^{\infty} \Omega_{qn}(r^2u)\, dF_{q(x_2)}(r), \tag{2.30}$$

with

$$a) \quad F_{a^2}(r) = \frac{\int_{(a,\sqrt{a^2+r^2}]} (w^2 - a^2)^{\frac{qn}{2}-1} w^{-(pn-2)}\, dF(w)}{\int_{(a,\infty)} (w^2 - a^2)^{\frac{qn}{2}-1} w^{-(pn-2)}\, dF(w)} \tag{2.31}$$

for $r \geq 0$ *if* $a > 0$ *and* $F(a) < 1$, and

$$b) \quad F_{a^2}(r) = 1 \text{ for } r \geq 0 \text{ if } a = 0 \text{ or } F(a) = 1. \tag{2.32}$$

2) $(X_1 | X_2) = M_1$ *for* $\text{vec}(X_2') \notin \text{vec}(M_2') + S$.

PROOF: Define $\mathbf{x} = \text{vec}(X')$, $\mathbf{x}_1 = \text{vec}(X_1')$, $\mathbf{x}_2 = \text{vec}(X_2')$, $\mathbf{m} = \text{vec}(M')$, $\mathbf{m}_1 = \text{vec}(M_1')$, and $\mathbf{m}_2 = \text{vec}(M_2')$. Then $\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}$ and $\mathbf{x} \sim E_{pn}(\mathbf{m}, \Sigma \otimes \Phi, \psi)$.

Now apply Theorem 2.7.3.

1) If $\mathbf{x}_2 \in \text{vec}(M_2') + S$, we have

$$\begin{aligned}(\mathbf{x}_1 | \mathbf{x}_2) \sim E_{qn}(&\mathbf{m}_1 + (\Sigma_{12} \otimes \Phi)(\Sigma_{22} \otimes \Phi)^{-}(\mathbf{x}_2 - \mathbf{m}_2),\\ &(\Sigma_{11} \otimes \Phi) - (\Sigma_{12} \otimes \Phi)(\Sigma_{22} \otimes \Phi)^{-}(\Sigma_{21} \otimes \Phi), \psi_{q(x_2)})\end{aligned} \tag{2.33}$$

where

$$q(\mathbf{x}_2) = (\mathbf{x}_2 - \mathbf{m}_2)' (\Sigma_{22} \otimes \Phi)^- (\mathbf{x}_2 - \mathbf{m}_2)$$

$$= (\mathrm{vec}((\mathbf{X}_2 - \mathbf{M}_2)'))' (\Sigma_{22}^- \otimes \Phi^-)(\mathrm{vec}((\mathbf{X}_2 - \mathbf{M}_2)')$$

$$= \mathrm{tr}((\mathbf{X}_2 - \mathbf{M}_2)' \Sigma_{22}^-(\mathbf{X}_2 - \mathbf{M}_2)\Phi^-)).$$

From (2.23) and (2.24) we get (2.30) and (2.31). Since $\mathbf{x}_2 \in \mathbf{m}_2 + S$, there exists $\mathbf{b} \in \mathbb{R}^{(p-q)n}$ such that $\mathbf{x}_2 - \mathbf{m}_2 = (\Sigma_{22} \otimes \Phi)\mathbf{b}$. Then, we have

$$(\Sigma_{12} \otimes \Phi)(\Sigma_{22} \otimes \Phi)^- (\mathbf{x}_2 - \mathbf{m}_2) = (\Sigma_{12} \otimes \Phi)(\Sigma_{22} \otimes \Phi)^- (\Sigma_{22} \otimes \Phi)\mathbf{b}$$

$$= (\Sigma_{12}\Sigma_{22}^-\Sigma_{22} \otimes \Phi\Phi^-\Phi)\mathbf{b}$$

$$= (\Sigma_{12}\Sigma_{22}^-\Sigma_{22} \otimes \Phi)\mathbf{b}$$

$$= (\Sigma_{12}\Sigma_{22}^- \otimes I_n)(\Sigma_{22} \otimes \Phi)\mathbf{b}$$

$$= (\Sigma_{12}\Sigma_{22}^- \otimes I_n)(\mathbf{x}_2 - \mathbf{m}_2).$$

We also have

$$(\Sigma_{11} \otimes \Phi) - (\Sigma_{12} \otimes \Phi)(\Sigma_{22} \otimes \Phi)^- (\Sigma_{21} \otimes \Phi)$$

$$= (\Sigma_{11} \otimes \Phi) - (\Sigma_{12}\Sigma_{22}^-\Sigma_{21}) \otimes \Phi$$

$$= (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^-\Sigma_{21}) \otimes \Phi.$$

Therefore, (2.33) can be written as

$$(x_1 \mid x_2) \sim E_{qn}(m_1 + (\Sigma_{12}\Sigma_{22}^{-} \otimes I_n)(x_2 - m_2), \Sigma_{11\cdot 2} \otimes \Phi, \psi_{q(x_2)}).$$

Hence,

$$X_1 \mid X_2 \sim E_{q,n}(M_1 + \Sigma_{12}\Sigma_{22}^{-}(X_2 - M_2), \Sigma_{11\cdot 2} \otimes \Phi, \psi_{q(X_2)}).$$

2) If $x_2 \notin \text{vec}(M_2') + S$, we get $(x_1 \mid x_2) = m_1$, so $(X_1 \mid X_2) = M_1$. ■

COROLLARY 2.6.4.1. *With the notation of Theorem 2.6.4, we have*

$$1 - F(w) = K_{a^2} \int_{(\sqrt{w^2-a^2},\infty)} (r^2 + a^2)^{\frac{pn}{2}-1} r^{-(qn-2)}\, dF_{a^2}(r), \quad w \geq a,$$

where $K_{a^2} = \int_{(a,\infty)} (w^2 - a^2)^{\frac{qn}{2}-1} w^{-(pn-2)}\, dF(w).$

PROOF: From (2.31) we get,

$$dF_{a^2}(r) = \frac{1}{K_{a^2}} (r^2)^{\frac{qn}{2}-1} (r^2 + a^2)^{-\frac{pn-2}{2}} \frac{dw}{dr}\, dF(w),$$

where $r^2 + a^2 = w^2$. Hence,

$$dF(w) = K_{a^2}\, r^{-(qn-2)} (r^2 + a^2)^{\frac{pn}{2}-1} \frac{dr}{dw}\, dF_{a^2}(r),$$

where $a < w \leq \sqrt{a^2 + r^2}$. Therefore,

$$1 - F(w) = K_{a^2} \int_{(\sqrt{w^2-a^2},\infty)} r^{-(qn-2)} (r^2 + a^2)^{\frac{pn}{2}-1}\, dF_{a^2}(r), \quad w \geq a. \blacksquare$$

THEOREM 2.6.5. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ *with stochastic representation* $rAUB'$. *Let* F *be the distribution function of* r *and* X, M, Σ *be partitioned as in Theorem 2.6.4.*

1) *If* $\mathrm{vec}(X_2') \in \mathrm{vec}(M_2') + S$, *where* S *is defined in Theorem 2.6.4, and*

   a) *if* X *has finite first moment, then*
   $\mathcal{E}(X_1 | X_2) = M_1 + \Sigma_{12}\Sigma_{22}^{-}(X_2 - M_2)$.

   b) *if* X *has finite second moment, then* $\mathrm{Cov}(X_1 | X_2) = c_1 \Sigma_{11 \cdot 2} \otimes \Phi$, *where* $c_1 = -2\psi'_{q(X_2)}(0)$, *and* $\psi_{q(X_2)}$ *is defined by (2.30), (2.31) and (2.32).*

2) *If* $\mathrm{vec}(X_2') \notin \mathrm{vec}(M_2') + S$, *then* $\mathcal{E}(X_1 | X_2) = M_1$, $\mathrm{Cov}(X_1 | X_2) = O$.

PROOF: It follows from Theorem 2.4.1 and Theorem 2.6.4. ■

The next theorem shows that if the distribution of X is absolutely continuous, then the constant c_1 in Theorem 2.6.5 can be obtained in a simple way. This was shown by Chu (1973), but his proof applies only to a subclass of absolutely continuous distributions. The following proof, however, works for all absolutely continuous distributions.

THEOREM 2.6.6. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ *and* X, M, Σ *be partitioned as in Theorem 2.6.4. Assume the distribution of* X *is absolutely continuous and it has finite second moment.*

Let $f_2(X_2) = \dfrac{1}{|\Sigma_{22}|^{\frac{n}{2}} |\Phi|^{\frac{pq}{2}}} h_2(\mathrm{tr}((X_2 - M_2)' \Sigma_{22}^{-1}(X_2 - M_2)\Phi^{-1}))$ *be the p.d.f. of the submatrix* X_2. *Then,* $\mathrm{Cov}(X_1 | X_2) = \dfrac{\int_r^{\infty} h_2(z)dz}{2h_2(r)} \Sigma_{11 \cdot 2} \otimes \Phi$, *where* $r = \mathrm{tr}((X_2 - M_2)' \Sigma_{22}^{-1}(X_2 - M_2)\Phi^{-1})$.

PROOF: Step 1. First we prove the theorem for the case $n = 1$, $\mathbf{m} = \mathbf{o}$. From Theorem 2.6.4 and Theorem 2.6.5, we conclude that

$\mathrm{Cov}(x_1 \mid x_2) = c_1\Sigma_{11\cdot2}$, where c_1 is determined by p, q, ψ and $x_2'\Sigma_{22}^{-1}x_2$. Hence, c_1 does not depend on Σ_{11} and Σ_{12}. Thus without loss of generality, we can assume that $\Sigma_{11} = I_q$ and $\Sigma_{12} = O$. Then $\Sigma_{11\cdot2} = I_q$. Let $x_1 = (x_1, x_2 \ldots x_q)'$, and $f_1(x_1,x_2) = |\Sigma_{22}|^{-\frac{1}{2}} h_1(x_1^2 + x_2'\Sigma_{22}^{-1}x_2)$ be the joint p.d.f. of x_1 and x_2. Then,

$$c_1 = \mathrm{Var}(x_1 \mid x_2) = \frac{\int_{-\infty}^{\infty} x_1^2 f_1(x_1,x_2)dx_1}{f_2(x_2)}$$

$$= \frac{\int_{-\infty}^{\infty} x_1^2 h_1(x_1^2 + x_2'\Sigma_{22}^{-1}x_2)dx_1}{h_2(x_2'\Sigma_{22}^{-1}x_2)}$$

$$= 2\frac{\int_{0}^{\infty} x_1^2 h_1(x_1^2 + x_2'\Sigma_{22}^{-1}x_2)dx_1}{h_2(x_2'\Sigma_{22}^{-1}x_2)}. \tag{2.34}$$

Now, $f_2(x_2) = \int_{-\infty}^{\infty} f_1(x_1,x_2)dx_1$, hence

$$h_2(x_2'\Sigma_{22}^{-1}x_2) = \int_{-\infty}^{\infty} h_1(x_1^2 + x_2'\Sigma_{22}^{-1}x_2)dx_1$$

$$= 2 \int_0^\infty h_1(x_1^2 + x_2'\Sigma_{22}^{-1}x_2)dx_1.$$

So, $h_2(z) = 2 \int_0^\infty h_1(x_1^2 + z)dx_1$, for $z \ge 0$. Hence, for $u \ge 0$, we get

$$\int_u^\infty h_2(z)dz = 2 \int_u^\infty \int_0^\infty h_1(x_1^2 + z)dx_1dz$$

$$= 2 \int_0^\infty \int_0^\infty \chi(z \ge u)h_1(x_1^2 + z)dx_1dz.$$

Let $w = \sqrt{x_1^2 + z - u}$. Then, $w^2 = x_1^2 + z - u$ and so $J(z \to w) = 2w$. Hence,

$$\int_u^\infty h_2(z)dz = 2 \int_0^\infty \int_0^\infty \chi(w^2 - x_1^2 + u \ge u)h_1(w^2 + u)2w\, dx_1dw$$

$$= 4 \int_0^\infty \int_0^\infty \chi(w^2 \ge x_1^2)\, wh_1(w^2 + u)dx_1\, dw$$

$$= 4 \int_0^\infty \int_0^w wh_1(w^2 + u)dx_1dw$$

$$= 4 \int_0^\infty (wh_1(w^2 + u) \int_0^w dx_1)dw$$

$$= 4 \int_0^\infty wh_1(w^2 + u)wdw$$

$$= 4 \int_0^\infty w^2h_1(w^2 + u)dw. \tag{2.35}$$

Now from (2.34) and (2.35), we get

$$c_1 = \frac{\int_u^\infty h_2(z)dz}{2h_2(u)}, \quad \text{where } u = x_2'\Sigma_{22}^{-1}x_2.$$

Step 2. Next, let $n = 1$, $\mathbf{m} \neq \mathbf{0}$, and $\mathbf{y} = \mathbf{x} - \mathbf{m}$. Then,

$$\mathrm{Cov}(y_1 \mid y_2) = \frac{\int_u^\infty h_2(z)dz}{2h_2(u)} \Sigma_{11\cdot 2},$$

where $u = y_2'\Sigma_{22}^{-1}y_2$. Therefore,

$$\begin{aligned}\mathrm{Cov}(x_1 \mid x_2) &= \mathrm{Cov}(y_1 + m_1 \mid y_2 = x_2 - m_2)\\ &= \frac{\int_u^\infty h_2(z)dz}{2h_2(u)} \Sigma_{11\cdot 2},\end{aligned}$$

where $u = (x_2 - m_2)' \Sigma_{22}^{-1}(x_2 - m_2)$.

Step 3. Finally, let $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$. Define $\mathbf{x} = \mathrm{vec}(X')$. Now, for $\mathbf{x}$ we can use Step 2. Therefore, $\mathrm{Cov}(X_1 \mid X_2) = \frac{\int_r^\infty h_2(z)dz}{2h_2(r)} \Sigma_{11\cdot 2}$, where $r = \mathrm{tr}((X_2 - M_2)' \Sigma_{22}^{-1}(X_2 - M_2)\Phi^{-1})$. ■

2.7. EXAMPLES

In this section we give some examples of the elliptically contoured distributions. We also give a method to generate elliptically contoured distributions.

2.7.1. ONE-DIMENSIONAL CASE

Let $p = n = 1$. Then, the class $E_1(m,\sigma,\psi)$, coincides with the class of one-dimensional distributions which are symmetric about a point. More precisely, $x \sim E_1(m,\sigma,\psi)$ if and only if $P(x \le r) = P(x \ge m - r)$ for every $r \in \mathbb{R}$. Some examples are: uniform, normal, Cauchy, double exponential, Student's t-distribution, and the distribution with the p.d.f.

$$f(x) = \frac{\sqrt{2}}{\pi\sigma\left(1 + \left(\frac{x}{\sigma}\right)^4\right)}, \quad \sigma > 0.$$

2.7.2. VECTOR VARIATE CASE

The definitions and results here are taken from Fang, Kotz and Ng (1990). Let $p > 1$ and $n = 1$.

(i) MULTIVARIATE UNIFORM DISTRIBUTION

The p-dimensional random vector $\mathbf{u}$ is said to have a multivariate uniform distribution if it is uniformly distributed on the unit sphere in $\mathbb{R}^p$.

THEOREM 2.7.2.1. *Let* $\mathbf{x} = (x_1,x_2,\ldots,x_p)'$ *have a p-variate uniform distribution. Then the p.d.f. of* $(x_1,x_2,\ldots,x_{p-1})$ *is*

$$\frac{\Gamma\left(\frac{p}{2}\right)}{\pi^{\frac{p}{2}}}\left(1 - \sum_{i=1}^{p-1} x_i^2\right)^{-\frac{1}{2}}, \quad \sum_{i=1}^{p-1} x_i^2 < 1.$$

PROOF: It follows from Theorem 2.5.6 that

$(x_1^2, x_2^2, \ldots, x_{p-1}^2) \sim D\left(\frac{1}{2}, \ldots, \frac{1}{2}; \frac{1}{2}\right)$. Hence, the p.d.f. of $(x_1^2, x_2^2, \ldots, x_{p-1}^2)$ is

$$\frac{\Gamma\left(\frac{p}{2}\right)}{\left(\Gamma\left(\frac{1}{2}\right)\right)^p} \prod_{i=1}^{p-1} (x_i^2)^{-\frac{1}{2}} \left(1 - \sum_{i=1}^{p-1} x_i^2\right)^{-\frac{1}{2}}.$$

Since the Jacobian of the transformation $(x_1^2, x_2^2, \ldots, x_{p-1}^2) \to (|x_1|, |x_2|, \ldots, |x_{p-1}|)$ is $2^{p-1} \prod_{i=1}^{p-1} |x_i|$, the p.d.f. of $(|x_1|, |x_2|, \ldots, |x_{p-1}|)$ is

$$\frac{2^{p-1} \Gamma\left(\frac{p}{2}\right)}{\pi^{\frac{p}{2}}} \left(1 - \sum_{i=1}^{p-1} x_i^2\right)^{-\frac{1}{2}}.$$

Now because of the spherical symmetry of x, the p.d.f. of $(x_1, x_2, \ldots, x_{p-1})$ is the p.d.f. of $(|x_1|, |x_2|, \ldots, |x_{p-1}|)$ divided by 2^{p-1}; that is

$$\frac{\Gamma\left(\frac{p}{2}\right)}{\pi^{\frac{p}{2}}} \left(1 - \sum_{i=1}^{p-1} x_i^2\right)^{-\frac{1}{2}}, \quad \sum_{i=1}^{p-1} x_i^2 < 1.$$

(ii) SYMMETRIC KOTZ TYPE DISTRIBUTION

The p-dimensional random vector x is said to have a symmetric Kotz type distribution with parameters $q, r, s \in \mathbb{R}$, μ: p-dimensional vector, Σ: $p \times p$ matrix, $r > 0$, $s > 0$, $2q + p > 2$, and $\Sigma > O$ if its p.d.f. is

$$f(x) = \frac{sr^{\frac{2q+p-2}{2s}}\,\Gamma\left(\frac{p}{2}\right)}{\pi^{\frac{p}{2}}\,\Gamma\left(\frac{2q+p-2}{2s}\right)|\Sigma|^{\frac{1}{2}}}\,[(x-\mu)'\,\Sigma^{-1}(x-\mu)]^{q-1})\,\exp\{-r[(x-\mu)'\,\Sigma^{-1}(x-\mu)]^{s}\}\,.$$

As a special case, take $q = s = 1$ and $r = \frac{1}{2}$. Then, we get the multivariate normal distribution with p.d.f.

$$f(x) = \frac{1}{(2\pi)^{\frac{p}{2}}\,|\Sigma|^{\frac{1}{2}}}\exp\left\{-\frac{(x-\mu)'\,\Sigma^{-1}(x-\mu)}{2}\right\},$$

and its characteristic function is

$$\phi_x(t) = \exp(it'\mu)\,\exp\left(-\frac{1}{2}\,t'\Sigma t\right). \tag{2.36}$$

The multivariate normal distribution is denoted by $N_p(\mu,\Sigma)$.

Remark 2.7.1. The distribution, $N_p(\mu,\Sigma)$, can be defined by its characteristic function (2.36). Then, Σ does not have to be positive definite; it suffices to assume that $\Sigma \geq O$.

(iii) Symmetric Multivariate Pearson Type II Distribution.

The p-dimensional random vector x is said to have a symmetric multivariate Pearson type II distribution with parameters $q \in \mathbb{R}$, μ: p-dimensional vector, Σ: $p \times p$ matrix with $q > -1$, and $\Sigma > O$ if its p.d.f. is

$$f(x) = \frac{\Gamma\left(\frac{p}{2}+q+1\right)}{\pi^{\frac{p}{2}}\,\Gamma(q+1)\,|\Sigma|^{\frac{1}{2}}}\,(1-(x-\mu)'\,\Sigma^{-1}(x-\mu))^{q},$$

where $(\mathbf{x} - \mu)' \Sigma^{-1}(\mathbf{x} - \mu) \le 1$.

(iv) SYMMETRIC MULTIVARIATE PEARSON TYPE VII DISTRIBUTION.

The p-dimensional random vector $\mathbf{x}$ is said to have a symmetric multivariate Pearson type VII distribution with parameters $q, r \in \mathbb{R}$, μ: p-dimensional vector, Σ: $p \times p$ matrix with $r > 0$, $q > \frac{p}{2}$, and $\Sigma > O$ if its p.d.f. is

$$f(\mathbf{x}) = \frac{\Gamma(q)}{(\pi r)^{\frac{p}{2}} \Gamma\left(q - \frac{p}{2}\right) |\Sigma|^{\frac{1}{2}}} \left(1 + \frac{(\mathbf{x} - \mu)'\Sigma^{-1}(\mathbf{x} - \mu)}{r}\right)^{-q}. \tag{2.37}$$

As a special case, when $q = \frac{p+r}{2}$, $\mathbf{x}$ is said to have a multivariate t-distribution with r degrees of freedom and it is denoted by $Mt_p(r,\mu,\Sigma)$.

THEOREM 2.7.2. *The class of symmetric multivariate Pearson type VII distributions coincides with the class of multivariate t-distributions.*

PROOF: Clearly, the multivariate t-distribution is Pearson type VII distribution. We ony have to prove that all Pearson type VII distributions are multivariate t-distributions.

Assume $\mathbf{x}$ has p.d.f. (2.37). Define $r_0 = 2\left(q - \frac{p}{2}\right)$ and $\Sigma_0 = \Sigma \frac{r}{r_0}$. Then, $Mt_p(r_0,\mu,\Sigma_0)$ is the same distribution as the one with p.d.f. (2.37). ■

The special case of multivariate t-distribution when $r = 1$; that is, $Mt_p(1,\mu,\Sigma)$ is called multivariate Cauchy distribution.

(v) SYMMETRIC MULTIVARIATE BESSEL DISTRIBUTION.

The p-dimensional random vector $\mathbf{x}$ is said to have a symmetric multivariate Bessel distribution with parameters $q, r \in \mathbb{R}$, μ: p-dimensional vector, Σ: $p \times p$ matrix with $r > 0$, $q > -\frac{p}{2}$, and $\Sigma > O$ if its p.d.f. is

$$f(x) = \frac{[(x-\mu)'\,\Sigma^{-1}(x-\mu)]^{\frac{q}{2}}}{2^{q+p-1}\,\pi^{\frac{p}{2}}\,r^{p+q}\,\Gamma\left(q+\frac{p}{2}\right)\,|\Sigma|^{\frac{1}{2}}}\,K_q\left(\frac{[(x-\mu)'\,\Sigma^{-1}(x-\mu)]^{\frac{1}{2}}}{r}\right),$$

where $K_q(z)$ is the modified Bessel function of the third kind; that is $K_q(z) = \frac{\pi}{2}\,\frac{I_{-q}(z) - I_q(z)}{\sin(q\pi)}$, $|\arg(z)| < \pi$, q is integer and

$$I_q(z) = \sum_{k=0}^{\infty} \frac{1}{k!\,\Gamma(k+q+1)}\left(\frac{z}{2}\right)^{q+2k}, \quad |z| < \infty, \quad |\arg(z)| < \pi.$$

If $q = 0$ and $r = \frac{\sigma}{\sqrt{2}}$, $\sigma > 0$, then x is said to have a multivariate Laplace distribution.

(vi) ELLIPTICALLY SYMMETRIC LOGISTIC DISTRIBUTION.

The p-dimensional random vector x is said to have an elliptically symmetric logistic distribution with parameters μ: p-dimensional vector, Σ: $p \times p$ matrix with $\Sigma > O$ if its p.d.f. is

$$f(x) = \frac{c}{|\Sigma|^{\frac{1}{2}}}\;\frac{\exp(-(x-\mu)'\,\Sigma^{-1}(x-\mu))}{(1+\exp(-(x-\mu)'\,\Sigma^{-1}(x-\mu)))^2}$$

with $c = \dfrac{\pi^{\frac{p}{2}}}{\Gamma\left(\frac{p}{2}\right)} \displaystyle\int_0^{\infty} z^{\frac{p}{2}-1}\,\frac{e^{-z}}{(1+e^{-z})^2}\,dz.$

(vii) SYMMETRIC MULTIVARIATE STABLE LAW.

The p-dimensional random vector x is said to follow a symmetric multivariate stable law with parameters $q,r \in \mathbb{R}$, μ: p-dimensional vector,

Σ: $p \times p$ matrix with $0 < q \leq 1$, $r > 0$, and $\Sigma \geq O$ if its characteristic function is

$$\phi_x(t) = \exp(it'\mu - r(t'\Sigma t)^q).$$

2.7.3. GENERAL MATRIX VARIATE CASE

The matrix variate elliptically contoured distributions listed here are the matrix variate versions of the multivariate distributions given in Section 2.7.2. Let $p \geq 1$ and $n \geq 1$.

(i) MATRIX VARIATE SYMMETRIC KOTZ TYPE DISTRIBUTION.

The $p \times n$ random matrix X is said to have a matrix variate symmetric Kotz type distribution with parameters $q,r,s \in \mathbb{R}$, M: $p \times n$, Σ: $p \times p$, Φ: $n \times n$ with $r > 0$, $s > 0$, $2q + pn > 2$, $\Sigma > O$, and $\Phi > O$ if its p.d.f. is

$$f(X) = \frac{s\, r^{\frac{2q+pn-2}{2s}}\, \Gamma\left(\frac{pn}{2}\right)}{\pi^{\frac{pn}{2}}\, \Gamma\left(\frac{2q+pn-2}{2s}\right) |\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} [\mathrm{tr}(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}]^{q-1}$$

$$\cdot \exp\{-r[\mathrm{tr}(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}]^s\}.$$

If we take $q = s = 1$ and $r = \frac{1}{2}$, we obtain the p.d.f. of the absolutely continuous matrix variate normal distribution:

$$f(X) = \frac{1}{(2\pi)^{\frac{pn}{2}} |\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} \mathrm{etr}\left\{-\frac{(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}}{2}\right\}.$$

The characteristic function of this distribution is

$$\phi_X(T) = \mathrm{etr}(iT'M)\,\mathrm{etr}\left(-\frac{1}{2}T'\Sigma T\Phi\right). \tag{2.38}$$

REMARK 2.7.2. If we define $N_{p,n}(M,\Sigma \otimes \Phi)$ through its characteristic function (2.38), then $\Sigma > O$ is not required, instead it suffices to assume that $\Sigma \geq O$.

(ii) MATRIX VARIATE PEARSON TYPE II DISTRIBUTION.

The $p \times n$ random matrix X is said to have a matrix variate symmetric Pearson type II distribution with parameters $q \in \mathbb{R}$, M: $p \times n$, Σ: $p \times p$, Φ: $n \times n$ with $q > -1$, $\Sigma > O$, and $\Phi > O$ if its p.d.f. is

$$f(X) = \frac{\Gamma\left(\frac{pn}{2} + q + 1\right)}{\pi^{\frac{pn}{2}} \Gamma(q+1) |\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} (1 - \mathrm{tr}((X-M)' \Sigma^{-1}(X-M)\Phi^{-1}))^q,$$

where $\mathrm{tr}((X-M)' \Sigma^{-1}(X-M)\Phi^{-1})) \leq 1$.

(iii) MATRIX VARIATE SYMMETRIC PEARSON TYPE VII DISTRIBUTION.

The $p \times n$ random matrix X is said to have a matrix variate symmetric Pearson type VII distribution with parameters $q,r \in \mathbb{R}$, M: $p \times n$, Σ: $p \times p$, Φ: $n \times n$ with $r > 0$, $q > \frac{pn}{2}$, $\Sigma > O$, and $\Phi > O$ if its p.d.f. is

$$f(X) = \frac{\Gamma(q)}{(\pi r)^{\frac{pn}{2}} \Gamma\left(q - \frac{pn}{2}\right) |\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} \left(1 + \frac{\mathrm{tr}((X-M)' \Sigma^{-1}(X-M)\Phi^{-1})}{r}\right)^{-q}.$$

Particularly, when $q = \frac{pn+r}{2}$, X is said to have a matrix variate t-distribution with r degrees of freedom and it is denoted by $Mt_{p,n}(r,M,\Sigma \otimes \Phi)$. It follows, from Theorem 2.7.2, that the class of matrix variate symmetric Pearson type VII distributions coincides with the class of matrix variate t-distributions.

When $r = 1$, in the definition of matrix variate t-distribution, that is, $Mt_{p,n}(1,M,\Sigma \otimes \Phi)$, then X is said to have a matrix variate Cauchy distribution.

(iv) MATRIX VARIATE SYMMETRIC BESSEL DISTRIBUTION.

The $p \times n$ random matrix X is said to have a matrix variate symmetric Bessel distribution with parameters $q,r \in \mathbb{R}$, M: $p \times n$, Σ: $p \times p$, Φ: $n \times n$ with $r > 0$, $q > -\frac{pn}{2}$, $\Sigma > O$, and $\Phi > O$ if its p.d.f. is

$$f(X) = \frac{[\mathrm{tr}(X - M)' \Sigma^{-1}(X - M)]^{\frac{1}{2}}}{2^{q+pn-1} \pi^{\frac{pn}{2}} r^{pn+q} \Gamma\left(q + \frac{pn}{2}\right) |\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} K_q\left(\frac{[\mathrm{tr}(X - M)' \Sigma^{-1}(X - M)]^{\frac{1}{2}}}{r}\right),$$

where $K_q(z)$ is the modified Bessel function of the third kind as defined in 2.7.2(v). For $q = 0$ and $r = \frac{\sigma}{\sqrt{2}}$, $\sigma > 0$, this distribution is known as the matrix variate Laplace distribution.

(v) MATRIX VARIATE SYMMETRIC LOGISTIC DISTRIBUTION.

The $p \times n$ random matrix X is said to have a matrix variate symmetric logistic distribution with parameters M: $p \times n$, Σ: $p \times p$, Φ: $n \times n$ with $\Sigma > O$, and $\Phi > O$ if its p.d.f. is

$$f(X) = \frac{c}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} \frac{\mathrm{etr}(-(X - M)' \Sigma^{-1}(X - M)\Phi^{-1})}{(1 + \mathrm{etr}(-(X - M)' \Sigma^{-1}(X - M)\Phi^{-1}))^2}$$

where $c = \dfrac{\pi^{\frac{pn}{2}}}{\Gamma\left(\frac{pn}{2}\right) \int_0^\infty z^{\frac{pn}{2}-1} \frac{e^{-z}}{(1 + e^{-z})^2} dz}$.

(vi) MATRIX VARIATE SYMMETRIC STABLE LAW.

The $p \times n$ random matrix X is said to follow a matrix variate symmetric stable law with parameters $q,r \in \mathbb{R}$, M: $p \times n$, Σ: $p \times p$, Φ: $n \times n$ with $0 < q \leq 1$, $r > 0$, $\Sigma \geq O$, and $\Phi \geq O$ if its characteristic function is

$$\phi_X(T) = \text{etr}(iT'M - r(T'\Sigma T\Phi)^q).$$

2.7.4. GENERATING ELLIPTICALLY CONTOURED DISTRIBUTIONS

If we have a m.e.c. distribution, based on it we can easily generate other m.e.c. distributions. For vector variate elliptical distributions, this is given in Muirhead (1982).

THEOREM 2.7.3. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ *have the p.d.f.* $f(X) = \dfrac{1}{|\Sigma|^{\frac{n}{2}}|\Phi|^{\frac{p}{2}}} h(\text{tr}((X-M)'\Sigma^{-1}(X-M)\Phi^{-1}))$. *Suppose* $G(z)$ *is a distribution function on* $(0,\infty)$. *Let*

$$g(X) = \frac{1}{|\Sigma|^{\frac{n}{2}}|\Phi|^{\frac{p}{2}}} \int_0^\infty z^{-\frac{pn}{2}} h\left(\frac{1}{z}\text{tr}((X-M)'\Sigma^{-1}(X-M)\Phi^{-1})\right) dG(z).$$

Then, $g(X)$ *is also the p.d.f. of a m.e.c. distribution.*

PROOF: Clearly, $g(X) \geq 0$. Moreover,

$$\int_{\mathbb{R}^{p\times n}} g(X)\, dX$$

$$= \int_{\mathbb{R}^{p\times n}} \frac{1}{|\Sigma|^{\frac{n}{2}}|\Phi|^{\frac{p}{2}}} \int_0^\infty z^{-\frac{pn}{2}} h\left(\frac{1}{2}\text{tr}((X-M)'\Sigma^{-1}(X-M)\Phi^{-1})\right) dG(z)\, dX$$

$$= \int_0^\infty \int_{\mathbb{R}^{p\times n}} \frac{1}{|z\Sigma|^{\frac{n}{2}}|\Phi|^{\frac{p}{2}}} h(\text{tr}((X-M)'(z\Sigma)^{-1}(X-M)\Phi^{-1}))\, dX\, dG(z)$$

$$= \int_0^1 1 \, dG(z) = 1.$$

Hence, $g(X)$ is a p.d.f. Let $r(w) = \int_0^\infty z^{-\frac{pn}{2}} h\left(\frac{w}{z}\right) dG(z)$. Then,

$$g(X) = \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} r(\mathrm{tr}((X - M)' \Sigma^{-1}(X - M)\Phi^{-1})).$$

Therefore, $g(X)$ is the p.d.f. of an elliptically contoured distribution. ■

COROLLARY 2.7.3.1. *Let* $h(u) = (2\pi)^{-\frac{pn}{2}} \exp\left(-\frac{u}{2}\right)$ *in Theorem 2.7.3. Then, for any distribution function* $G(z)$ *on* $(0,\infty)$,

$$g(X) = \frac{1}{(2\pi)^{\frac{pn}{2}} |\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} \int_0^\infty z^{-\frac{pn}{2}} \mathrm{etr}\left(\frac{-1}{2z}(X - M)' \Sigma^{-1}(X - M)\Phi^{-1}\right) dG(z)$$

defines the p.d.f. of a m.e.c. distribution. In this case, the distribution of X *is said to be a mixture of normal distributions.*

In particular, if $G(1) = 1 - \varepsilon$ and $G(\sigma^2) = \varepsilon$ with $0 < \varepsilon < 1$, $\sigma^2 > 0$, we obtain the ε-contaminated matrix variate normal distribution. It has the p.d.f.

$$f(X) = \frac{1}{(2\pi)^{\frac{pn}{2}} |\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} [(1 - \varepsilon) \,\mathrm{etr}(-\frac{1}{2}(X - M)' \Sigma^{-1}(X - M)\Phi^{-1})$$

$$+ \frac{\varepsilon}{\sigma^{pn}} \mathrm{etr}(-\frac{1}{2\sigma^2}(X - M)' \Sigma^{-1}(X - M)\Phi^{-1})].$$

CHAPTER 3

PROBABILITY DENSITY FUNCTION AND EXPECTED VALUES

3.1. PROBABILITY DENSITY FUNCTION

The m.e.c. probability density function has some interesting properties which will be given in this chapter. These results are taken from Kelker (1970), Cambanis, Huang, and Simons (1981), Fang, Kotz, and Ng (1990), and Gupta and Varga (1992b).

The first remarkable property is that the marginal distributions of a m.e.c. distribution are absolutely continuous unless the original distribution has an atom of positive weight at zero. Even if the original distribution has an atom of positive weight at zero, the marginal density is absolutely continuous outside zero. This is shown for multivariate elliptical distributions in the following theorem due to Cambanis, Huang, and Simons (1981).

THEOREM 3.1.1. *Let* $\mathbf{x} \sim E_p(\mathbf{o}, \mathbf{I}_p, \psi)$ *have the stochastic representation* $\mathbf{x} \approx r\mathbf{u}$. *Let* $F(r)$ *be the distribution function of* r. *Assume* $\mathbf{x}$ *is partitioned as* $\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}$, *where* $\mathbf{x}_1$ *is q-dimensional,* $1 \le q < p$. *Let* $\mathbf{x}_1$ *have the stochastic representation* $\mathbf{x}_1 \approx r_1\mathbf{u}_1$. *Then, the distribution of* r_1 *has an atom of weight* $F(0)$ *at zero and it is absolutely continuous on* $(0,\infty)$ *with p.d.f.*

$$g_q(s) = \frac{2\Gamma\left(\frac{p}{2}\right)s^{q-1}}{\Gamma\left(\frac{q}{2}\right)\Gamma\left(\frac{p-q}{2}\right)} \int_s^\infty r^{-(p-2)} (r^2 - s^2)^{\frac{p-q}{2}-1} \, dF(r), \quad 0 < s < \infty. \tag{3.1}$$

PROOF. From Corollary 2.5.6.1, it follows that $r_1 \approx r r_0$, where

$r_0^2 \sim B\left(\frac{q}{2}, \frac{p-q}{2}\right)$. Therefore, $P(r_1 = 0) = P(r = 0) = F(0)$ and

$$P(0 < r_1 \le t) = P(0 < rr_0 \le t)$$

$$= \int_{(0,\infty)} P(0 < rr_0 \le t)\, dF(r)$$

$$= \int_{(0,\infty)} P\left(0 < r_0 \le \frac{t}{r}\right) dF(r)$$

$$= \int_{(0,\infty)} P\left(0 < r_0^2 \le \frac{t^2}{r^2}\right) dF(r)$$

$$= \int_{(0,\infty)} \frac{\Gamma\left(\frac{p}{2}\right)}{\Gamma\left(\frac{q}{2}\right)\Gamma\left(\frac{p-q}{2}\right)} \int_{\left(0,\min\left(1,\frac{t^2}{r^2}\right)\right]} x^{\frac{q}{2}-1}(1-x)^{\frac{p-q}{2}-1}\, dx\, dF(r)$$

$$= \int_{(0,\infty)} \frac{\Gamma\left(\frac{p}{2}\right)}{\Gamma\left(\frac{q}{2}\right)\Gamma\left(\frac{p-q}{2}\right)} \int_{\left(0,\min\left(1,\frac{t^2}{r^2}\right)\right]} x^{\frac{q}{2}-1}(1-x)^{\frac{p-q}{2}-1}\, dx\, dF(r)\,.$$

Let $x = \frac{s^2}{r^2}$. Then, $J(x \to s) = \frac{r^2}{2s}$ and we have

$$\int_{(0,\infty)} \frac{\Gamma\left(\frac{p}{2}\right)}{\Gamma\left(\frac{q}{2}\right)\Gamma\left(\frac{p-q}{2}\right)} \int_{\left(0,\min\left(1,\frac{t^2}{r^2}\right)\right]} x^{\frac{q}{2}-1}(1-x)^{\frac{p-q}{2}-1}\, dx\, dF(r)$$

$$= \int_{(0,\infty)} \frac{\Gamma\left(\frac{p}{2}\right)}{\Gamma\left(\frac{q}{2}\right)\Gamma\left(\frac{p-q}{2}\right)} \int_{(0,\min(r,t)]} \frac{2s}{r^2}\left(\frac{s}{r}\right)^{q-2}\left(1-\frac{s^2}{r^2}\right)^{\frac{p-q}{2}-1} ds\, dF(r)$$

$$= \frac{2\Gamma\left(\frac{p}{2}\right)}{\Gamma\left(\frac{q}{2}\right)\Gamma\left(\frac{p-q}{2}\right)} \int_{(0,\infty)} \int_{(0,\min(r,t)]} s^{q-1}\, r^{-2-q+2-(p-q)+2} (r^2-s^2)^{\frac{p-q}{2}-1} ds\, dF(r)$$

$$= \int_{(0,t]} \frac{2\Gamma\left(\frac{p}{2}\right) s^{q-1}}{\Gamma\left(\frac{q}{2}\right)\Gamma\left(\frac{p-q}{2}\right)} \int_{[s,\infty)} r^{-(p-2)} (r^2-s^2)^{\frac{p-q}{2}-1} dF(r)\, ds,$$

from which (3.1) follows. ■

COROLLARY 3.1.1.1. *Let* $\mathbf{x} \sim E_p(\mathbf{o}, I_p, \psi)$ *and assume that* $P(\mathbf{x} = \mathbf{o}) = 0$. *Let* $\mathbf{x} \approx r\mathbf{u}$ *be the stochastic representation of* $\mathbf{x}$ *and* $F(r)$ *the distribution function of* r. *Partition* $\mathbf{x}$ *into* $\mathbf{x} = \begin{pmatrix}\mathbf{x}_1\\ \mathbf{x}_2\end{pmatrix}$, *where* $\mathbf{x}_1$ *is q-dimensional,* $1 \le q < p$. *Then,* $\mathbf{x}_1$ *is absolutely continuous with p.d.f.*

$$f_q(\mathbf{y}) = \frac{\Gamma\left(\frac{p}{2}\right)}{\pi^{\frac{q}{2}}\Gamma\left(\frac{p-q}{2}\right)} \int_{(\mathbf{y}'\mathbf{y})^{\frac{1}{2}}}^{\infty} r^{-(p-2)} (r^2 - \mathbf{y}'\mathbf{y})^{\frac{p-q}{2}-1} dF(r). \tag{3.2}$$

PROOF. From Theorem 3.1.1, it follows that $\mathbf{x}_1$ is absolutely continuous and if $r_1\mathbf{u}_1$ is the stochastic representation of $\mathbf{x}_1$ then r_1 has the p.d.f.

$$g_q(s) = \frac{2\Gamma\left(\frac{p}{2}\right) s^{q-1}}{\Gamma\left(\frac{q}{2}\right)\Gamma\left(\frac{p-q}{2}\right)} \int_s^{\infty} r^{-(p-2)} (r^2-s^2)^{\frac{p-q}{2}-1} dF(r). \tag{3.3}$$

Since, x_1 also has a m.e.c. distribution, its p.d.f. is of the form

$$f_q(y) = h_q(y'y). \tag{3.4}$$

From Theorem 2.5.5, it follows that

$$g_q(s) = \frac{2\pi^{\frac{q}{2}}}{\Gamma\left(\frac{q}{2}\right)} s^{q-1} h_q(s^2), \quad s \geq 0, \tag{3.5}$$

and from (3.3) and (3.5), it follows that

$$h_q(s^2) = \frac{\Gamma\left(\frac{p}{2}\right)}{\pi^{\frac{q}{2}} \Gamma\left(\frac{p-q}{2}\right)} \int_s^{\infty} r^{-(p-2)} (r^2 - s^2)^{\frac{p-q}{2}-1} \, dF(r). \tag{3.6}$$

Now, (3.2) follows from (3.4) and (3.6). ■

The following result was given by Fang, Kotz, and Ng (1990).

THEOREM 3.1.2. *Let* $x \sim E_p(o, I_p, \psi)$ *with p.d.f.* $f(x) = h(x'x)$. *Let* x *be partitioned as* $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, *where* x_1 *is q-dimensional,* $1 \leq q < p$. *Then,* x_1 *is absolutely continuous and its p.d.f. is*

$$f_q(y) = \frac{\pi^{\frac{p-q}{2}}}{\Gamma\left(\frac{p-q}{2}\right)} \int_{y'y}^{\infty} (u - y'y)^{\frac{p-q}{2}-1} h(u) \, du. \tag{3.7}$$

PROOF. Let ru be the stochastic representation of x and F be the distribution function of r. Then, from Theorem 2.5.5, we get the p.d.f. of r as

$$g(r) = \frac{2\pi^{\frac{p}{2}}}{\Gamma\left(\frac{p}{2}\right)} r^{p-1} h(r^2). \tag{3.8}$$

From (3.2) and (3.8), we have

$$f_q(y) = \frac{\Gamma\left(\frac{p}{2}\right)}{\pi^{\frac{q}{2}}\Gamma\left(\frac{p-q}{2}\right)} \int_{(y'y)^{\frac{1}{2}}}^{\infty} r^{-(p-2)} (r^2 - y'y)^{\frac{p-q}{2}-1} \frac{2\pi^{\frac{p}{2}}}{\Gamma\left(\frac{p}{2}\right)} r^{p-1} h(r^2)dr$$

$$= \frac{2\pi^{\frac{p-q}{2}}}{\Gamma\left(\frac{p-q}{2}\right)} \int_{(y'y)^{\frac{1}{2}}}^{\infty} r(r^2 - y'y)^{\frac{p-q}{2}-1} h(r^2)\, dr.$$

Let $u = r^2$. Then, $J(r \to u) = \frac{1}{2r}$ and we have

$$f_q(y) = \frac{\pi^{\frac{p-q}{2}}}{\Gamma\left(\frac{p-q}{2}\right)} \int_{y'y}^{\infty} (u - y'y)^{\frac{p-q}{2}-1} h(u)\, du. \blacksquare$$

Marginal densities have certain continuity and differentiability properties. These are given in the following theorems. The first theorem is due to Kelker (1970).

THEOREM 3.1.3. *Let* $x \sim E_p(o, I_p, \psi)$, $1 < p$, *with* p.d.f. $f(x) = h(x'x)$. *Let* x *be partitioned as* $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, *where* x_1 *is* $p - 1$ *dimensional. Let* x_1 *have the p.d.f.* $f_{p-1}(y) = h_{p-1}(y'y)$. *If* $h_{p-1}(z)$ *is bounded in a neighborhood of* z_0 *then* $h_{p-1}(z)$ *is continuous at* $z = z_0$.

PROOF. From (3.7), we get

$$h_{p-1}(z) = \frac{\pi^{\frac{1}{2}}}{\Gamma\left(\frac{1}{2}\right)} \int_z^{\infty} (u - z)^{-\frac{1}{2}} h(u)\, du.$$

Thus,

$$h_{p-1}(z) = \int_z^{\infty} (u - z)^{-\frac{1}{2}} h(u)\, du. \tag{3.9}$$

Choose any $\eta > 0$. There exist $k > 0$ and $K > 0$ such that if $|z - z_0| < k$ then $h(z) < K$. Let $\varepsilon = \min\left(k, \frac{\eta^2}{64}\right)$. Further let δ be such that $0 < \delta < \varepsilon$, and $z_0 < z < z_0 + \delta$. Then, we have

$$|h_{p-1}(z) - h_{p-1}(z_0)| = \left| \int_z^{\infty} (u - z)^{-\frac{1}{2}} h(u)\, du - \int_{z_0}^{\infty} (u - z_0)^{-\frac{1}{2}} h(u)\, du \right|$$

$$= \left| \int_{z_0+\varepsilon}^{\infty} ((u - z)^{-\frac{1}{2}} - (u - z_0)^{-\frac{1}{2}})\, h(u)\, du + \int_z^{z_0+\varepsilon} ((u - z)^{-\frac{1}{2}} - (u - z_0)^{-\frac{1}{2}})\, h(u)\, du \right.$$

$$\left. - \int_{z_0}^{z} (u - z_0)^{-\frac{1}{2}} h(u)\, du \right|$$

$$\leq \int_{z_0+\varepsilon}^{\infty} ((u - z)^{-\frac{1}{2}} - (u - z_0)^{-\frac{1}{2}})\, h(u)\, du$$

$$+ \int_z^{z_0+\varepsilon} ((u - z)^{-\frac{1}{2}} - (u - z_0)^{-\frac{1}{2}})\, h(u)\, du + \int_{z_0}^{z} (u - z_0)^{-\frac{1}{2}} h(u)\, du. \tag{3.10}$$

Now,

$$\int_{z}^{z_0+\varepsilon} ((u - z)^{-\frac{1}{2}} - (u - z_0)^{-\frac{1}{2}})\, h(u)\, du$$

$$\leq \int_{z}^{z_0+\varepsilon} (u - z)^{-\frac{1}{2}}\, h(u)\, du$$

$$\leq K \int_{z}^{z_0+\varepsilon} (u - z)^{-\frac{1}{2}}\, du$$

$$\leq 2K(\varepsilon + z_0 - z)^{\frac{1}{2}} \leq 2\sqrt{\varepsilon} \tag{3.11}$$

and

$$\int_{z_0}^{z} (u - z_0)^{-\frac{1}{2}}\, h(u)\, du \leq 2K(z - z_0)^{\frac{1}{2}}$$

$$\leq 2\sqrt{\varepsilon}. \tag{3.12}$$

Furthermore,

$$\int_{z_0+\varepsilon}^{\infty} ((u - z)^{-\frac{1}{2}} - (u - z_0)^{-\frac{1}{2}})\, h(u)\, du$$

$$\leq \int_{z_0+\varepsilon}^{\infty} ((u - (z_0 + \delta))^{-\frac{1}{2}} - (u - z_0)^{-\frac{1}{2}}) h(u) du. \tag{3.13}$$

We have

$$\lim_{\delta\to 0} [(u - (z_0 + \delta))^{-\frac{1}{2}} - (u - z_0)^{-\frac{1}{2}}] = 0 \text{ for } u \geq z_0 + \varepsilon,$$

and

$$0 \le (u - (z_0 + \delta))^{-\frac{1}{2}} - (u - z_0)^{-\frac{1}{2}}$$

$$\le (u - (z_0 + \varepsilon))^{-\frac{1}{2}} - (u - z_0)^{-\frac{1}{2}}$$

$$\le (u - (z_0 + \varepsilon))^{-\frac{1}{2}} \text{ for } u \ge z_0 + \varepsilon.$$

Since,

$$\int_{z_0+\varepsilon}^{\infty} (u - (z_0 + \varepsilon))^{-\frac{1}{2}} h(u)\, du = h_{p-1}(z_0 + \varepsilon) < \infty,$$

we can use the dominated Lebesgue convergence theorem to get

$$\lim_{\delta \to 0} \int_{z_0+\varepsilon}^{\infty} ((u + z)^{-\frac{1}{2}} - (u - z_0)^{-\frac{1}{2}})\, h(u)du = 0.$$

Therefore, there exists $\upsilon > 0$ such that if $0 < \delta < \upsilon$ then

$$\int_{z_0+\varepsilon}^{\infty} ((u - z)^{-\frac{1}{2}} - (u - z_0)^{-\frac{1}{2}})\, h(u)du < \frac{\eta}{2}. \tag{3.14}$$

Hence, if $\delta = \min(\varepsilon, \upsilon)$ then for $z_0 < z < z_0 + \delta$, (3.10)-(3.14) give

$$| h_{p-1}(z) - h_{p-1}(z_0) | \le 4\sqrt{\varepsilon} + \frac{\eta}{2} \le \eta.$$

Therefore, $h_{p-1}(z)$ is continuous from the right at z_0. In a similar way, it can be proved that $h_{p-1}(z)$ is continuous from the left at z_0. ■

The next theorem shows that the $p - 2$ dimensional marginal density of a p-dimensional absolutely continuous multivariate elliptical distribution is differentiable. This theorem is due to Kelker (1970).

THEOREM 3.1.4. *Let* $\mathbf{x} \sim E_p(\mathbf{o}, I, \psi)$, $2 < p$, *with p.d.f.* $f(\mathbf{x}) = h(\mathbf{x}'\mathbf{x})$. *Let* $\mathbf{x}$ *be partitioned as* $\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}$, *where* $\mathbf{x}_1$ *is* $p - 2$ *dimensional. Let* $\mathbf{x}_1$ *have the p.d.f.* $f_{p-2}(\mathbf{y}) = h_{p-2}(\mathbf{y}'\mathbf{y})$. *Then,* $h_{p-2}(z)$ *is differentiable and*

$$h'_{p-2}(z) = -\pi h(z). \tag{3.15}$$

PROOF. From Theorem 3.1.2, we get

$$h_{p-2}(z) = \pi \int_z^{\infty} h(u)\, du, \quad z \geq 0. \tag{3.16}$$

Hence, $h'_{p-2}(z) = -\pi h(z)$. ∎

REMARK 3.1.1. Theorem 3.1.4 shows that if $p \geq 3$ and the one-dimensional marginal density of a p-dimensional absolutely continuous spherical distribution is known, then all the marginal densities and also the density of the parent distribution can be obtained easily. In fact, if
$f_j(\mathbf{y}) = h_j(\mathbf{y}'\mathbf{y})$, $\mathbf{y} \in \mathbb{R}^j$, denotes the p.d.f. of the j-dimensional marginal, then from (3.15) we get

$$h_{2j+1}(z) = \left(-\frac{1}{\pi}\right)^j h_1^{(j)}(z), \quad z \geq 0.$$

From (3.9), we have

$$h_2(z) = \int_z^{\infty} (u - z)^{-\frac{1}{2}} h_3(u)\, du.$$

From $h_2(z)$, we can obtain the other marginals using

$$h_{2j}(z) = \left(-\frac{1}{\pi}\right)^{j-1} h_2^{(j-1)}(z), \quad z \geq 0.$$

REMARK 3.1.2. Assume $\mathbf{x} \sim E_p(\mathbf{o}, I_p, \psi)$ is absolutely continuous, $p \geq 3$. If $f_j(\mathbf{y}) = h_j(\mathbf{y}'\mathbf{y})$, $\mathbf{y} \in \mathbb{R}^j$, denotes the p.d.f. of the j-dimensional marginal then from (3.16) it follows that $h_j(z)$, $z \geq 0$ is nondecreasing for $j = 1, 2, \ldots, p-2$. Since $f_1(z) = h_1(z^2)$, the p.d.f. of the one-dimensional marginal, is nondecreasing on $(-\infty, 0]$ and nonincreasing on $[0, \infty)$.

The results given in this chapter so far referred to the vector variate elliptically contoured distributions. Their extension to the matrix variate case is straightforward since $X \sim E_{p,n}(O, I_p \otimes I_n, \psi)$ is equivalent to $\mathbf{x} = \text{vec}(X') \sim E_{pn}(\mathbf{o}, I_{pn}, \psi)$. The following theorems can be easily derived.

THEOREM 3.1.5. *Let* $X \sim E_{p,n}(O, I_{pn}, \psi)$ *have the stochastic representation* $X \approx rU$. *Let* $F(r)$ *be the distribution function of* r. *Assume* X *is partitioned as* $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, *where* X_1 *is* $q \times n$. *Let* X_1 *have the stochastic representation* $X_1 \approx r_1 U_1$. *Then, the distribution of* r_1 *has an atom of weight* $F(0)$ *at zero and it is absolutely continuous on* $(0, \infty)$ *with p.d.f.*

$$g_{q,n}(s) = \frac{2\Gamma\left(\frac{pn}{2}\right) s^{qn-1}}{\Gamma\left(\frac{qn}{2}\right)\Gamma\left(\frac{(p-q)n}{2}\right)} \int_s^{\infty} r^{-(pn-2)} (r^2 - s^2)^{\frac{(p-q)n}{2}-1} dF(r),$$

$$0 < s < \infty.$$

COROLLARY 3.1.5.1. *With the notation of Theorem 3.1.5 we get* $P(X_1 = O) = P(X = O)$.

THEOREM 3.1.6. *Let* $X \sim E_{p,n}(O, I_p \otimes I_n, \psi)$ *and assume that* $P(X = O) = 0$. *Let* $X \approx rU$ *be the stochastic representation of* X *and* $F(r)$ *the distribution function* *of* r. *Partition* X *into* $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, *where* X_1 *is* $q \times n$, $1 \le q < p$. *Then,* X_1 *is absolutely continuous with p.d.f.*

$$f_{q,n}(Y) = \frac{\Gamma\left(\frac{pn}{2}\right)}{\pi^{\frac{qn}{2}} \Gamma\left(\frac{(p-q)n}{2}\right)} \int_{(tr(Y'Y))^{\frac{1}{2}}}^{\infty} r^{-(pn-2)} (r^2 - tr(Y'Y))^{\frac{(p-q)n}{2}-1} dF(r).$$

THEOREM 3.1.7. *Let* $X \sim E_{p,n}(O, I_p \otimes I_n, \psi)$ *with p.d.f.* $f(X) = h(tr(X'X))$. *Let* X *be partitioned as* $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, *where* X_1 *is* $q \times n$, $1 \le q < p$. *Then,* X_1 *is absolutely continuous with p.d.f.*

$$f_q(Y) = \frac{\pi^{\frac{(p-q)n}{2}}}{\Gamma\left(\frac{(p-q)n}{2}\right)} \int_{tr(Y'Y)} (u - tr(Y'Y))^{\frac{p-q}{2}-1} h(u)\, du.$$

THEOREM 3.1.8. *Let* $X \sim E_{p,n}(O, I_p \otimes I_n, \psi)$, $1 < p$, $1 < n$, *with p.d.f.* $f(X) = h(tr(X'X))$. *Let* X *be partitioned as* $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, *where* X_1 *is* $q \times n$, $1 \le q < p$. *Let* X_1 *have the p.d.f.* $f_{q,n}(Y) = h_{q,n}(tr(Y'Y))$. *Then,* $h_{q,n}(z)$ *is differentiable and*

i) *if* $(p - q)n$ *is even,* $(p - q)n = 2j$, *then*

$$h_{q,n}^{(j)}(z) = (-\pi)^j h(z),$$

ii) *if* $(p - q)n$ *is odd,* $(p - q)n = 2j + 1$, *then*

$$\frac{\partial^j}{\partial z^j}\left(\int_z^{\infty} (u - z)^{-\frac{1}{2}} h'_{q,n}(u) du \right) = (-\pi)^{j+1} h(z).$$

THEOREM 3.1.9. *Let* $X \sim E_{p,n}(O, I_p \otimes I_n, \psi)$, $pn \ge 3$ *with p.d.f.*

$f(X) = h(\mathrm{tr}(X'X))$. *Let* $f_1(y) = h_1(y^2)$ *denote the p.d.f. of a one-dimensional marginal of* X. *Then, if we know* $h_1(z)$, *we can obtain* $h(z)$ *in the following way.*

i) *If* pn *is odd,* $pn = 2j + 1$, *then*

$$h(z) = \frac{h_1^{(j)}(z)}{(-\pi)^j}.$$

ii) *If* pn *is even,* $pn = 2j$, *then*

$$h(z) = \left(\frac{-1}{\pi}\right)^j \frac{\partial^{j-1}}{\partial z^{j-1}} \int_z^{\infty} (u - z)^{-\frac{1}{2}} h_1'(u)\, du.$$

THEOREM 3.1.10. *Assume* $X \sim E_{p,n}(O, I_p \otimes I_n, \psi)$ *is absolutely continuous,* $p > 1$, $n > 1$. *If* $f_{j,n}(Y) = h_{j,n}(\mathrm{tr}(Y'Y))$, $Y \in \mathbb{R}^{j\times n}$, *denotes the p.d.f. of the* $j \times n$ *dimensional marginal, then* $h_j(z)$, $z \geq 0$ *is nondecreasing for* $j = 1,2,\ldots,p-1$. *Moreover, the p.d.f. of the one-dimensional marginal is nondecreasing on* $(-\infty, 0]$ *and nonincreasing on* $[0, \infty)$.

3.2. MORE ON EXPECTED VALUES

The stochastic representation can be used to compute the expected values of matrix variate functions when the underlying distribution is elliptically contoured. In order to simplify the expressions for the expected values, we need the following theorem which is a special case of Berkane and Bentler (1986a).

THEOREM 3.2.1. *Let* $\phi(t) = \psi(t^2)$ *be a characteristic function of a one-dimensional elliptical distribution. Assume the distribution has finite* m^{th} *moment. Then,*

$$\phi^{(m)}(0) = \begin{cases} \dfrac{m!}{\left(\frac{m}{2}\right)!}\psi^{\left(\frac{m}{2}\right)}(0) & \text{if } m \text{ is even} \\ 0 & \text{if } m \text{ is odd.} \end{cases} \tag{3.17}$$

PROOF. If the m^{th} moment exists, say μ_m then, $\phi^{(m)}(t)$ also exists and $\phi^{(m)}(0) = i^m\mu_m$. First we prove that

$$\phi^{(k)}(t) = \sum_{n=\left[\frac{k+1}{2}\right]}^{k} K_n^k\, t^{2n-k}\, \psi^{(n)}(t^2), \tag{3.18}$$

where $0 \le k \le m$, $[z]$ denotes the integer part of z, and K_n^k is a constant not depending on ψ. We prove this by induction.

If $k = 0$, we have $\phi(t) = \psi(t^2)$, so $K_0^0 = 1$. If $k = 1$ we get $\phi'(t) = 2t\psi'(t^2)$, and the statement is true in this case, with $K_1^1 = 2$. Assume the theorem is true up to $\ell < m$. Then, for $k = \ell + 1$ we get

$$\begin{aligned}\phi^{(k)}(t) &= \sum_{n=\left[\frac{\ell+1}{2}\right]}^{\ell} (2n-\ell)K_n^{\ell}\, t^{2n-\ell-1}\,\psi^{(n)}(t^2) \\ &\quad + \sum_{n=\left[\frac{\ell+1}{2}\right]}^{\ell} 2K_n^{\ell}\, t^{2n-\ell+1}\,\psi^{(n+1)}(t^2) \\ &= \sum_{n=\left[\frac{k}{2}\right]}^{k-1} (2n-k+1)K_n^{k-1}\, t^{2n-k}\,\psi^{(n)}(t^2) \\ &\quad + \sum_{n=\left[\frac{k}{2}\right]+1}^{k} 2K_{n-1}^{k-1}\, t^{2n-k}\,\psi^{(n)}(t^2)\end{aligned}$$

$$= \left(2\left[\frac{k}{2}\right] - k + 1\right) K^{k-1}_{\left[\frac{k}{2}\right]} t^{2\left[\frac{k}{2}\right] - k} \psi^{\left(\left[\frac{k}{2}\right]\right)}(t^2)$$

$$+ \sum_{n=\left[\frac{k}{2}\right]+1}^{k-1} (2K^{k-1}_{n-1} + (2n - k + 1)K^{k-1}_{n}) t^{2n-k} \psi^{(n)}(t^2)$$

$$+ 2K^{k-1}_{k-1} t^{2k-k} \psi^{(k)}(t^2). \tag{3.19}$$

We have to distinguish between two cases.

a) k even.

Then, $\left[\frac{k+1}{2}\right] = \left[\frac{k}{2}\right]$ and $2\left[\frac{k}{2}\right] - k + 1 = 1$. Hence, (3.19) gives

$$\phi^{(k)}(t) = \sum_{n=\left[\frac{k+1}{2}\right]}^{k} K^{k}_{n} t^{2n-k} \psi^{(n)}(t^2), \text{ with}$$

$$K^k_n = \begin{cases} K^{k-1}_{\left[\frac{k+1}{2}\right]} & \text{if } n = \left[\frac{k+1}{2}\right] \\ 2K^{k-1}_{n-1} + (2n - k + 1)K^{k-1}_{n} & \text{if } \left[\frac{k+1}{2}\right] < n < k \\ 2K^{k-1}_{k-1} & \text{if } n = k. \end{cases}$$

b) k odd.

Then, $\left[\frac{k+1}{2}\right] = \left[\frac{k}{2}\right] + 1$ and $2\left[\frac{k}{2}\right] - k + 1 = 0$. From (3.19) we get

$$\phi^{(k)}(t) = \sum_{n=\left[\frac{k+1}{2}\right]}^{k} K^{k}_{n} t^{2n-k} \psi^{(n)} (t^2), \text{ with}$$

$$K_n^k = \begin{cases} 2K_{n-1}^{k-1} + (2n - k + 1)K_n^{k-1} & \text{if } \left[\frac{k+1}{2}\right] \le n < k \\ 2K_{k-1}^{k-1} & \text{if } n = k. \end{cases}$$

Hence (3.18) is established. Taking $k = m$ and $t = 0$ in (3.18), we get

$$\phi^{(m)}(0) = \begin{cases} K_{\frac{m}{2}}^{m} \psi^{\left(\frac{m}{2}\right)}(0) & \text{if } m \text{ is even} \\ 0 & \text{if } m \text{ is odd.} \end{cases} \tag{3.20}$$

If m is even, $m = 2s$ say, then we have $\phi^{(2s)}(0) = i^{2s}\mu_{2s}$, that is $\phi^{(2s)}(0) = (-1)^s \mu_{2s}$. Let $x \sim N_1(0,1)$, then $\mu_{2s} = \frac{(2s)!}{2^s s!}$, $\psi(z) = \exp\left(-\frac{z}{2}\right)$, and $\psi^{(s)}(z) = \left(-\frac{1}{2}\right)^s \exp\left(-\frac{z}{2}\right)$, from which $\psi^{(s)}(0) = \left(-\frac{1}{2}\right)^s$ follows. Therefore, $\phi^{(2s)}(0) = K_s^{2s}\psi^{(s)}(0) = K_s^{2s}\left(-\frac{1}{2}\right)^s$. Thus, we get

$$K_s^{2s}\left(-\frac{1}{2}\right)^s = (-1)^s \frac{(2s)!}{2^s s!},$$

from which it follows that $K_s^{2s} = \frac{(2s)!}{s!}$. Comparing this with (3.20), we obtain (3.17). ■

COROLLARY 3.2.1.1. *Let $\phi(t) = \psi(t^2)$ be as in Theorem 3.2.1, then*

i) $\phi(0) = 1$,

ii) *if $\psi(t)$ is differentiable,* $\phi'(0) = 0$,

iii) *if $\psi(t)$ has second derivative,* $\phi''(0) = 2\psi'(0)$,

iv) *if $\psi(t)$ has third derivative,* $\phi'''(0) = 0$,

v) *if $\psi(t)$ has fourth derivative,* $\phi^{iv}(0) = 12\psi''(0)$.

COROLLARY 3.2.1.2. *Let* $\phi(t) = \exp\left(-\frac{t^2}{2}\right)$. *Then,*

$$\phi^{(m)}(0) = \begin{cases} \left(-\frac{1}{2}\right)^{\frac{m}{2}} \dfrac{m!}{\left(\frac{m}{2}\right)!} & \text{if } m \text{ is even} \\ 0 & \text{if } m \text{ is odd.} \end{cases}$$

For example,

i) $\phi(0) = 1$,

ii) $\phi'(0) = 0$,

iii) $\phi''(0) = -1$,

iv) $\phi'''(0) = 0$,

v) $\phi^{iv}(0) = 3$.

Now, we can derive the following results.

THEOREM 3.2.2. *Let* $X \sim E_{p,n}(O, \Sigma \otimes \Phi, \psi)$ *with* $P(X = O) = 0$. *Let* $q = rk(\Sigma)$, $m = rk(\Phi)$ *and* $rAUB'$ *be the stochastic representation of* X. *Assume* $Y \sim N_{p,n}(O, \Sigma \otimes \Phi)$. *Let* $\mathcal{F}$ *be a subset of the* $p \times n$ *real matrices such that if* $Z \in \mathbb{R}^{p\times n}$, $Z \in \mathcal{F}$, *and* $a > 0$, *then* $aZ \in \mathcal{F}$ *and* $P(X \notin \mathcal{F}) = P(Y \notin \mathcal{F}) = 0$. *Let* $K(Z)$ be a function defined on $\mathcal{F}$ *such that if* $Z \in \mathcal{F}$ *and* $a > 0$, *then* $K(aZ) = a^k K(Z)$ *where* $k > -qm$. *Assuming* $\mathcal{E}(K(X))$ *and* $\mathcal{E}(K(Y))$ *exist, we get*

a) $$\mathcal{E}(K(X)) = \mathcal{E}(K(Y)) \frac{\mathcal{E}(r^k)\, \Gamma\left(\frac{qm}{2}\right)}{2^{\frac{k}{2}}\, \Gamma\left(\frac{qm+k}{2}\right)},$$

b) *if* X *has the p.d.f.* $f(X) = \dfrac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(tr(X'\Sigma^{-1}X\Phi^{-1}))$ *then*

$$\mathcal{E}(K(X)) = \mathcal{E}(K(Y)) \frac{\pi^{\frac{np}{2}} \int_0^\infty z^{np+k-1} h(z^2) dz}{2^{\frac{k}{2}-1}\, \Gamma\left(\frac{np+k}{2}\right)},$$

c) *if* $k = 0$, *then* $\mathcal{E}(K(X)) = \mathcal{E}(K(Y))$,

d) *if* k *is a positive even integer, then*

$\mathcal{E}(K(X)) = \mathcal{E}(K(Y))(-2)^{\frac{k}{2}} \psi^{(\frac{k}{2})}(0)$,

e) *if* k *is a positive odd integer and* $K(aZ) = a^k K(Z)$ *holds for all* $a \neq 0$, *then* $\mathcal{E}(K(X)) = O$.

PROOF. a) $K(X)$ and $K(Y)$ are defined if $X \in \mathcal{F}$ and $Y \in \mathcal{F}$. Since $P(X \notin \mathcal{F}) = P(Y \notin \mathcal{F}) = 0$, we see that $K(X)$ and $K(Y)$ are defined with probability one.

Let r_2AU_2B' be the stochastic representation of Y. From the conditions of the theorem, it follows that if $aZ \in \mathcal{F}$ and $a > 0$, then $Z \in \mathcal{F}$. Since, $P(X = O) = 0$, we have $P(r = 0) = 0$, and from $P(rAU_1B' \in \mathcal{F}) = 1$, we get $P(AUB' \in \mathcal{F}) = 1$. Thus $K(AUB')$ is defined with probability one. Moreover, $P(K(rAUB') = r^k K(AUB')) = 1$. Therefore $\mathcal{E}(K(rAUB') = \mathcal{E}(r^k K(AUB'))$. Since, r and U are independent, we get

$$\mathcal{E}(K(AUB')) = \mathcal{E}(r^k)\, \mathcal{E}(K(AUB')). \tag{3.21}$$

Similarly,

$$\mathcal{E}(K(r_2AU_2B')) = \mathcal{E}(r^k)\, \mathcal{E}(K(AU_2B')). \tag{3.22}$$

However, $AUB' \approx AU_2B'$, hence

$$\mathcal{E}(K(AUB')) = \mathcal{E}(K(AU_2B')). \tag{3.23}$$

Now Theorem 2.5.5 shows that r_2 has the p.d.f.

$$q_2(r_2) = \frac{1}{2^{\frac{qm}{2}-1}\Gamma\left(\frac{qm}{2}\right)} r_2^{qm-1} \exp\left\{-\frac{r_2^2}{2}\right\}, \quad r_2 \geq 0.$$

Therefore,

$$\mathcal{E}(r_2^k) = \int_0^\infty \frac{1}{2^{\frac{qm}{2}-1}\Gamma\left(\frac{qm}{2}\right)} r_2^{qm+k-1} \exp\left\{-\frac{r_2^2}{2}\right\} dr_2.$$

Let $z = r_2^2$. Then $\frac{dr_2}{dz} = \frac{1}{2\sqrt{z}}$ and we have

$$\begin{aligned}
\mathcal{E}(r_2^k) &= \frac{1}{2^{\frac{qm}{2}-1}\Gamma\left(\frac{qm}{2}\right)} \int_0^\infty z^{\frac{qm+k-1}{2}} \exp\left\{-\frac{z}{2}\right\} \frac{1}{2\sqrt{z}} dz \\
&= \frac{1}{2^{\frac{qm}{2}}\Gamma\left(\frac{qm}{2}\right)} \int_0^\infty z^{\frac{qm+k}{2}-1} \exp\left\{-\frac{z}{2}\right\} dz \\
&= \frac{2^{\frac{k}{2}}\Gamma\left(\frac{qm+k}{2}\right)}{\Gamma\left(\frac{qm}{2}\right)}. \qquad (3.24)
\end{aligned}$$

Now, from (3.22) and (3.24) we get

$$\begin{aligned}
\mathcal{E}(K(\Sigma^{\frac{1}{2}}U_2\Phi^{\frac{1}{2}})) &= \frac{\mathcal{E}(K(Y))}{\mathcal{E}(r_2^k)} \\
&= \frac{\Gamma\left(\frac{qm}{2}\right)}{2^{\frac{k}{2}}\Gamma\left(\frac{qm+k}{2}\right)} \mathcal{E}(K(Y)). \qquad (3.25)
\end{aligned}$$

Using (3.21), (3.23) and (3.25) we get

$$\mathcal{E}(K(X)) = \mathcal{E}(K(Y)) \frac{\mathcal{E}(r^k)\, \Gamma\left(\frac{qm}{2}\right)}{2^{\frac{k}{2}}\, \Gamma\left(\frac{qm+k}{2}\right)}. \tag{3.26}$$

b) Using Theorem 2.5.5, the p.d.f. of r is

$$q_1(r) = \frac{2\pi^{\frac{np}{2}}}{\Gamma\left(\frac{np}{2}\right)} r^{np-1}\, h(r^2), \quad r \geq 0.$$

Therefore,

$$\mathcal{E}(r^k) = \int_0^\infty \frac{2\pi^{\frac{np}{2}}}{\Gamma\left(\frac{np}{2}\right)} z^{np+k-1}\, h(z^2) dz. \tag{3.27}$$

From (3.26) and (3.27), we get

$$\mathcal{E}(K(X)) = \mathcal{E}(K(Y)) \frac{\pi^{\frac{np}{2}} \int_0^\infty z^{np+k-1}\, h(z^2) dz}{2^{\frac{k}{2}-1}\, \Gamma\left(\frac{k+np}{2}\right)}.$$

c) This is a special case of (a).

d) It follows from part (a), that $\mathcal{E}(K(X)) = c_{q,m}(\psi,m)\mathcal{E}(K(Y))$ where $c_{q,m}(\psi,k)$ is a constant depending on q, m, k and ψ only. So, in order to determine $c_{q,m}(\psi,k)$, we can choose $X \sim E_{q,m}(O, I_q \otimes I_m, \psi)$, $Y \sim N_{q,m}(O, I_q \otimes I_m)$ and $K(Z) = z_{11}^k$ where z_{11} is the $(1,1)^{th}$ element of the

$q \times m$ matrix Z. Then $K(aZ) = a^k K(Z)$, $a > 0$, is obviously satisfied. Now, $x_{11}^2 \leq \mathrm{tr}(X'X)$ and hence $|x_{11}|^k \leq (\mathrm{tr}(X'X))^{\frac{k}{2}}$. Here $r \approx (\mathrm{tr}(X'X))^{\frac{1}{2}}$ and since r^k is integrable, $(\mathrm{tr}(X'X))^{\frac{k}{2}}$ is also integrable over $[0,\infty)$. Therefore, $\mathcal{E}(x_{11}^k)$ exists. Similarly, $\mathcal{E}(y_{11}^k)$ also exists.

Hence, we can write

$$\mathcal{E}(x_{11}^k) = c_{q,m}(\psi,k)\ \mathcal{E}(y_{11}^k). \tag{3.28}$$

However, $x_{11} \sim E_1(0,1,\psi)$ and $y_{11} \sim N_1(0,1)$. Then, from Theorem 3.2.1, it follows that $\mathcal{E}(x_{11}^k) = \dfrac{k!}{i^k \left(\frac{k}{2}\right)!} \psi^{\left(\frac{k}{2}\right)}(0)$ and from Corollary 3.2.1.2 we get that $\mathcal{E}(y_{11}^k) = \dfrac{k!}{i^k \left(\frac{k}{2}\right)!} \left(-\frac{1}{2}\right)^{\frac{k}{2}}$. Hence $c_{q,m}(\psi,k) = (-2)^{\frac{k}{2}}\, \psi^{\left(\frac{k}{2}\right)}(0)$.

e) Take $a = -1$. Then, we have $K(-Z) = (-1)^k K(Z)$ and since k is odd, we get $K(-Z) = -K(Z)$. However, $X \approx -X$ and so $K(X) \approx K(-X)$. Therefore, $\mathcal{E}(K(X)) = \mathcal{E}(K(-X)) = -\mathcal{E}(K(X))$ and hence $\mathcal{E}(K(X)) = O$. ■

In the next theorem we examine expected values of functions which are defined on the whole $p \times n$ dimensional real space. In contrast to Theorem 3.2.3, here we do not require that the underlying distribution assign probability zero to the zero matrix.

THEOREM 3.2.3. *Let* $X \sim E_{p,n}(O,\Sigma \otimes \Phi,\psi)$. *Let* $q = \mathrm{rk}(\Sigma)$, $m = \mathrm{rk}(\Phi)$, *and* $rAUB'$ *be the stochastic representation of* X. *Assume* $Y \sim N_{p,n}(O,\Sigma \otimes \Phi)$. *Let* $K(Z)$ *be a function defined on* $\mathbb{R}^{p\times n}$ *such that if* $Z \in \mathbb{R}^{p\times n}$ *and* $a \geq 0$ *then* $K(aZ) = a^k K(Z)$ *where* $k > -qm$. *Assume* $\mathcal{E}(K(X))$ *and* $\mathcal{E}(K(Y))$ *exist. Then,* *a)* $$\mathcal{E}(K(X)) = \mathcal{E}(K(Y)) \frac{\mathcal{E}(r^k)\,\Gamma\left(\frac{qm}{2}\right)}{2^{\frac{k}{2}}\,\Gamma\left(\frac{qm+k}{2}\right)},$$

b) *if* $k = 0$ *then* $\mathcal{E}(K(X)) = \mathcal{E}(K(Y))$,

c) *if* k *is a positive even integer, then*

$$\mathcal{E}(K(X)) = \mathcal{E}(K(Y))(-2)^{\frac{k}{2}}\,\psi^{(\frac{k}{2})}(0),$$

d) *if* k *is a positive odd integer and* $K(aZ) = a^k K(Z)$ *holds for all* $a \neq 0$, *then* $\mathcal{E}(K(X)) = O$.

PROOF. Let r_2AU_2B' be the stochastic representation of Y. Then, we have

$$\mathcal{E}(K(rAUB')) = \mathcal{E}(r^k)\,\mathcal{E}(K(AUB')) \tag{3.29}$$

and

$$\mathcal{E}(K(r_2AU_2B')) = \mathcal{E}(r_2^k)\,\mathcal{E}(K(AU_2B')). \tag{3.30}$$

However (3.29) and (3.30) are the same as (3.21) and (3.22). Therefore, the proof can be completed in exactly the same way as the proof of Theorem 3.2.3. ■

Since moments of normal random variables are well known, using Theorem 3.2.4, we can obtain the moments of m.e.c. distributions.

THEOREM 3.2.4. *Let* $X \sim E_{p,n}(O, I_p \otimes I_n, \psi)$. *Let* rU *be the stochastic representation of* X. *Then, provided the moments exist, we get*

$$\mathcal{E}\left(\prod_{i=1}^{p}\prod_{j=1}^{n} x_{ij}^{2s_{ij}}\right) = \mathcal{E}(r^{2s})\,\frac{\Gamma\left(\frac{np}{2}\right)}{\pi^{\frac{np}{2}}\,\Gamma\left(\frac{np}{2}+s\right)}\prod_{i=1}^{p}\prod_{j=1}^{n}\Gamma\left(\frac{1}{2}+s_{ij}\right), \qquad (3.31)$$

where s_{ij} *are nonnegative integers and* $s = \sum_{i=1}^{p}\sum_{j=1}^{n} s_{ij}$. *We also have*

$$\mathcal{E}\left(\prod_{i=1}^{p}\prod_{j=1}^{n} x_{ij}^{2s_{ij}}\right) = \psi^{(s)}(0)\,\frac{(-1)^s\,2^{2s}}{\pi^{\frac{np}{2}}}\prod_{i=1}^{p}\prod_{j=1}^{n}\Gamma\left(\frac{1}{2}+s_{ij}\right). \qquad (3.32)$$

PROOF. We use Theorem 3.2.4. Let $Z \in \mathbb{R}^{p\times n}$ and $K(Z) = \prod_{i=1}^{p}\prod_{j=1}^{n} x_{ij}^{2s_{ij}}$. If $a \geq 0$, then

$$K(aZ) = a^{2s}K(Z). \qquad (3.33)$$

If $Y \sim N_{p,n}(O, I_p \otimes I_n)$, then the elements of Y are independently and identically distributed standard normal variables, hence,

$$\mathcal{E}\left(\prod_{i=1}^{p}\prod_{j=1}^{n} y_{ij}^{2s_{ij}}\right) = \prod_{i=1}^{p}\prod_{j=1}^{n} \mathcal{E}(y_{ij}^{2s_{ij}}), \qquad (3.34)$$

and

$$\mathcal{E}(y_{ij}^{2s_{ij}}) = \frac{(2s_{ij})!}{s_{ij}!\,2^{s_{ij}}}$$

$$= 2^{s_{ij}}\,\frac{\Gamma\left(\frac{1}{2}\right)}{\pi^{\frac{1}{2}}}\,\frac{1}{2}\,\frac{3}{2}\cdots\frac{2s_{ij}-1}{2}$$

$$= \frac{2^{s_{ij}}\Gamma\left(\frac{1}{2}+s_{ij}\right)}{\pi^{\frac{1}{2}}}. \tag{3.35}$$

Now, from (3.33), (3.34), (3.35) and part (a) of Theorem 3.2.4 we obtain (3.31).

On the other hand, (3.33), (3.34), (3.35), and part (c) of Theorem 3.2.4 yield (3.32). ■

The formula (3.31) is given in Fang, Kotz, and Ng (1990).

THEOREM 3.2.5. *Let* $X \sim E_{p,n}(O,\Sigma \otimes \Phi,\psi)$. *Then, provided the left-hand sides exist,*

i) $\mathcal{E}(x_{i_1j_1}) = 0$.

ii) $\mathcal{E}(x_{i_1j_1}x_{i_2j_2}) = -2\psi'(0)\sigma_{i_1i_2}\phi_{j_1j_2}$.

iii) $\mathcal{E}(x_{i_1j_1}x_{i_2j_2}x_{i_3j_3}) = 0$.

iv) $\mathcal{E}(x_{i_1j_1}x_{i_2j_2}x_{i_3j_3}x_{i_4j_4}) = 4\psi''(0)(\sigma_{i_1i_2}\phi_{j_1j_2}\sigma_{i_3i_4}\phi_{j_3j_4} + \sigma_{i_1i_3}\phi_{j_1j_3}\sigma_{i_2i_4}\phi_{j_2j_4} + \sigma_{i_1i_4}\phi_{j_1j_4}\sigma_{i_2i_3}\phi_{j_2j_3})$.

PROOF. Step 1. Let $Y \sim N_{p,n}(O,I_p \otimes I_n)$. Then, the elements of Y are independent, standard normal variables, so

$$\mathcal{E}(y_{i_1j_1}) = 0,$$
$$\mathcal{E}(y_{i_1j_1}y_{i_2j_2}) = \delta_{i_1i_2}\delta_{j_1j_2},$$
$$\mathcal{E}(y_{i_1j_1}y_{i_2j_2}y_{i_3j_3}) = 0,$$
$$\mathcal{E}(y_{i_1j_1}y_{i_2j_2}y_{i_3j_3}y_{i_4j_4}) = \delta_{i_1i_2}\delta_{j_1j_2}\delta_{i_3i_4}\delta_{j_3j_4} + \delta_{i_1i_3}\delta_{j_1j_3}\delta_{i_2i_4}\delta_{j_2j_4} + \delta_{i_1i_4}\delta_{j_1j_4}\delta_{i_2i_3}\delta_{j_2j_3}.$$

Step 2. Let $X \sim N_{p,n}(O,\Sigma \otimes \Phi)$. Let $q = rk(\Sigma)$, $m = rk(\Phi)$ and $rAUB'$ be the stochastic representation of X. Then, we can write $X = AYB'$, where $Y \sim N_{q,m}(O,I_q \otimes I_m)$, $x_{ij} = \sum_{\ell=1}^{n}\sum_{k=1}^{p} a_{ik}y_{k\ell}b_{\ell j}$, $\Sigma = AA'$, and $\Phi = BB'$.

Using the result of Step 1, we get

$$\mathcal{E}(x_{i_1j_1}) = \mathcal{E}\left(\sum_{\ell=1}^{n}\sum_{k=1}^{p} a_{i_1k}y_{k\ell}b_{\ell j_1}\right)$$

$$= \sum_{\ell,k} a_{i_1k}\,\mathcal{E}(y_{k\ell})b_{\ell j_1}$$

$$= 0.$$

$$\mathcal{E}(x_{i_1j_1}x_{i_2j_2}) = \mathcal{E}\left(\left(\sum_{\ell=1}^{n}\sum_{k=1}^{p} a_{i_1k}y_{k\ell}b_{\ell j_1}\right)\left(\sum_{t=1}^{n}\sum_{s=1}^{p} a_{i_2s}y_{st}b_{tj_2}\right)\right)$$

$$= \sum_{\ell,k,t,s} a_{i_1k}b_{\ell j_1}a_{i_2s}b_{tj_2}\,\mathcal{E}(y_{k\ell}y_{st})$$

$$= \sum_{\ell,k,t,s} a_{i_1k}b_{\ell j_1}a_{i_2s}b_{tj_2}\delta_{ks}\delta_{\ell t}$$

$$= \sum_{\ell,k} a_{i_1k}a_{i_2k}b_{\ell j_1}b_{\ell j_2}$$

$$= \sigma_{i_1i_2}\phi_{j_1j_2}.$$

$$\mathcal{E}(x_{i_1j_1}x_{i_2j_2}x_{i_3j_3}) = \mathcal{E}\left(\left(\sum_{\ell=1}^{n}\sum_{k=1}^{p} a_{i_1k}y_{k\ell}b_{\ell j_1}\right)\left(\sum_{t=1}^{n}\sum_{s=1}^{p} a_{i_2s}y_{st}b_{tj_2}\right)\right.$$

$$\left.\cdot\left(\sum_{q=1}^{n}\sum_{r=1}^{p} a_{i_3r}y_{rq}b_{qj_3}\right)\right)$$

$$= \sum_{\ell,k,t,s,q,r} a_{i_1k}b_{\ell j_1}a_{i_2s}b_{tj_2}a_{i_3r}b_{qj_3}\cdot E(y_{k\ell}y_{st}y_{rq})$$

$$= 0.$$

$$\mathcal{E}(x_{i_1j_1}x_{i_2j_2}x_{i_3j_3}x_{i_4j_4}) = \mathcal{E}\left(\left(\sum_{\ell=1}^{n}\sum_{k=1}^{p} a_{i_1k}y_{k\ell}b_{\ell j_1}\right)\left(\sum_{t=1}^{n}\sum_{s=1}^{p} a_{i_2s}y_{st}b_{tj_2}\right)\right.$$
$$\left.\cdot\left(\sum_{q=1}^{n}\sum_{r=1}^{p} a_{i_3r}y_{rq}b_{qj_3}\right)\left(\sum_{w=1}^{n}\sum_{u=1}^{p} a_{i_4u}y_{uw}b_{wj_4}\right)\right)$$
$$= \sum_{\ell,k,t,s,q,r,w,u} a_{i_1k}b_{\ell j_1}a_{i_2s}b_{tj_2}$$
$$\cdot\, a_{i_3r}b_{qj_3}a_{i_4u}b_{wj_4}\mathcal{E}(y_{k\ell}y_{st}y_{rq}y_{uw})$$
$$= \sum_{\ell,k,t,s,q,r,w,u} a_{i_1k}b_{\ell j_1}a_{i_2s}b_{tj_2}$$
$$\cdot\, a_{i_3r}b_{qj_3}a_{i_4u}b_{wj_4}\cdot(\delta_{ks}\delta_{\ell t}\delta_{ru}\delta_{qw}$$
$$+ \delta_{kr}\delta_{\ell q}\delta_{su}\delta_{tw} + \delta_{ku}\delta_{\ell w}\delta_{sr}\delta_{tq})$$
$$= \sum_{k,\ell,r,q} a_{i_1k}a_{i_2k}b_{\ell j_1}b_{\ell j_2}a_{i_3r}a_{i_4r}b_{qj_3}b_{qj_4}$$
$$+ \sum_{k,\ell,s,t} a_{i_1k}a_{i_3k}b_{\ell j_1}b_{\ell j_3}a_{i_2s}a_{i_4s}b_{tj_2}b_{tj_4}$$
$$+ \sum_{k,\ell,s,t} a_{i_4k}a_{i_4u}b_{\ell j_1}b_{\ell j_4}a_{i_2s}a_{i_3s}b_{tj_2}b_{tj_3}$$
$$= \sigma_{i_1i_2}\phi_{j_1j_2}\sigma_{i_3i_4}\phi_{j_3j_4} + \sigma_{i_1i_3}\phi_{j_1j_3}\sigma_{i_2i_4}\phi_{j_2j_4}$$
$$+ \sigma_{i_1i_4}\phi_{j_1j_4}\sigma_{i_2i_3}\phi_{j_2j_3}.$$

Step 3. Let $X \sim E_{p,n}(O,\Sigma \otimes \Phi,\psi)$. Define, on the set of $p \times n$ dimensional matrices, the functions

$$K_{i_1,j_1}(Z) = z_{i_1j_1}$$

$$K_{i_1,j_1,i_2,j_2}(Z) = z_{i_1j_1}z_{i_2j_2}$$

$$K_{i_1,j_1,i_2,j_2,i_3,j_3}(Z) = z_{i_1j_1}z_{i_2j_2}z_{i_3j_3}$$

$$K_{i_1,j_1,i_2,j_2,i_3,j_3,i_4,j_4}(Z) = z_{i_1j_1}z_{i_2j_2}z_{i_3j_3}z_{i_4j_4}.$$

Now, the results follow from Theorem 3.2.4. ■

THEOREM 3.2.6. *Let* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$. *Then provided the left-hand sides exist,*

i) $\mathcal{E}(x_{i_1j_1}) = m_{i_1j_1}$.

ii) $\mathcal{E}(x_{i_1j_1}x_{i_2j_2}) = -2\psi'(0)\sigma_{i_1i_2}\phi_{j_1j_2} + m_{i_1j_1}m_{i_2j_2}$.

iii) $\mathcal{E}(x_{i_1j_1}x_{i_2j_2}x_{i_3j_4}) = -2\psi'(0)(\sigma_{i_ii_2}\phi_{j_1j_2}m_{i_3j_3} + \sigma_{i_1i_3}\phi_{j_1j_3}m_{i_2j_2} + \sigma_{i_2i_3}\phi_{j_2j_3}m_{i_1j_1}) + m_{i_1j_1}m_{i_1j_2}m_{i_3j_3}$.

iv)
$$\begin{aligned}\mathcal{E}(x_{i_1j_1}x_{i_2j_2}x_{i_3j_3}x_{i_4j_4}) &= 4\psi''(0)(\sigma_{i_1i_2}\phi_{j_1j_2}\sigma_{i_3i_4}\phi_{j_3j_4} + \sigma_{i_1i_3}\phi_{j_1j_3}\sigma_{i_2i_4}\phi_{j_2j_4}\\ &+ \sigma_{i_1i_4}\phi_{j_1j_4}\sigma_{i_2i_3}\phi_{j_2j_3}) - 2\psi'(0)(m_{i_1j_1}m_{i_2j_2}\sigma_{i_3i_4}\phi_{j_3j_4}\\ &+ m_{i_1j_1j_{i_3j_3}}\sigma_{i_2i_4}\phi_{j_2j_4} + m_{i_1j_1}m_{i_4j_4}\sigma_{i_2i_3}\phi_{j_2j_3}\\ &+ m_{i_2j_2}m_{i_3j_3}\sigma_{i_1i_4}\phi_{j_1j_4} + m_{i_2j_2}m_{i_4j_4}\sigma_{i_1i_3}\phi_{j_1j_3}\\ &+ m_{i_3j_3}m_{i_4j_4}\sigma_{i_1i_2}\phi_{j_1j_2}) + m_{i_1j_1}m_{i_2j_2}m_{i_3j_3}m_{i_4}m_{j_4}.\end{aligned}$$

PROOF. Let $Y = X - M$. Then, $Y \sim E_{p,n}(O,\Sigma \otimes \Phi,\psi)$ and using Theorem 3.2.6, we obtain the first four moments of Y. Therefore,

i) $\mathcal{E}(x_{i_1j_1}) = \mathcal{E}(y_{i_1j_1} + m_{i_1j_1}) = m_{i_1j_1}$.

ii)
$$\begin{aligned}\mathcal{E}(x_{i_1j_1}x_{i_2j_2}) &= \mathcal{E}((y_{i_1j_1} + m_{i_1j_1})(y_{i_2j_2} + m_{i_2j_2}))\\ &= \mathcal{E}(y_{i_1j_1}y_{i_2j_2}) + \mathcal{E}(y_{i_1j_1})m_{i_2j_2} + \mathcal{E}(y_{i_2j_2})m_{i_1j_1}\\ &\quad + m_{i_1j_1}m_{i_2j_2}\\ &= -2\psi'(0)\sigma_{i_1i_2}\phi_{j_1j_2} + m_{i_1j_1}m_{i_2j_2}.\end{aligned}$$

iii)
$$\begin{aligned}\mathcal{E}(x_{i_1j_1}x_{i_2j_2}x_{i_3j_3}) &= \mathcal{E}((y_{i_1j_1} + m_{i_1j_1})(y_{i_2j_2} + m_{i_2j_2})(y_{i_3j_3} + m_{i_3j_3})\\ &= \mathcal{E}(y_{i_1j_1}y_{i_2j_2}y_{i_3j_3}) + \mathcal{E}(y_{i_1j_1}y_{i_2j_2})m_{i_3j_3}\\ &\quad + \mathcal{E}(y_{i_1j_1}y_{i_3j_3})m_{i_2j_2} + \mathcal{E}(y_{i_2j_2}y_{i_3j_3})m_{i_1j_1}\\ &\quad + m_{i_1j_1}m_{i_2j_2}m_{i_3j_3}\\ &= -2\psi'(0)(\sigma_{i_1i_2}\phi_{j_1j_2}m_{i_3j_3} + \sigma_{i_1i_3}\phi_{j_1j_3}m_{i_2j_2}\\ &\quad + \sigma_{i_2i_3}\phi_{j_2j_3}m_{i_1j_1}) + m_{i_1j_1}m_{i_2j_2}m_{i_3j_3}.\end{aligned}$$

iv)
$$\begin{aligned}\mathcal{E}(x_{i_1j_1}x_{i_2j_2}x_{i_3j_3}x_{i_4j_4}) &= \mathcal{E}((y_{i_1j_1} + m_{i1j_1})(y_{i_2j_2} + m_{i_2j_2})\\ &\quad \cdot (y_{i_3j_3} + m_{i_3j_3})(y_{i_4j_4} + m_{i_4j_4}))\\ &= \mathcal{E}(y_{i_1j_1}y_{i_2j_2}y_{i_3j_3}y_{i_4j_4}) + \mathcal{E}(y_{i_1j_1}y_{i_2j_2}y_{i_3j_3})m_{i_4j_4}\end{aligned}$$

$$
\begin{aligned}
&+ \mathcal{E}(y_{i_1j_1}y_{i_2j_2}y_{i_4j_4})m_{i_3j_3} + \mathcal{E}(y_{i_1j_1}y_{i_3j_3}y_{i_4j_4})m_{i_2j_2} \\
&+ \mathcal{E}(y_{i_2j_2}y_{i_3j_3}y_{i_4j_4})m_{i_1j_1} \\
&+ \mathcal{E}(y_{i_1j_1}y_{i_2j_2})m_{i_3j_4}m_{i_4j_4} \\
&+ \mathcal{E}(y_{i_1j_1}y_{i_3j_3}))m_{i_2j_2}m_{i_4j_4} \\
&+ \mathcal{E}(y_{i_1j_12}y_{i_4j_4})m_{i_2j_2}m_{i_3j_3} \\
&+ \mathcal{E}(y_{i_2j_2}y_{i_3j_3})m_{i_1j_1}m_{i_4j_4} \\
&+ \mathcal{E}(y_{i_2j_2}y_{i_4j_4})m_{i_1j_1}m_{i_3j_3} \\
&+ \mathcal{E}(y_{i_3j_3}y_{i_4j_4})m_{i_1j_1}m_{i_2j_4} + m_{i_1j_1}m_{i_2j_2}m_{i_3j_3}m_{i_4j_4} \\
= {}& 4\psi''(0)(\sigma_{i_1i_2}\phi_{j_1j_2}\sigma_{i_3i_4}\phi_{j_3j_4} + \sigma_{i_1i_3}\phi_{j_1j_3}\sigma_{i_2i_4}\phi_{j_2j_4} \\
&+ \sigma_{i_1i_4}\phi_{j_1j_4}\sigma_{i_2i_3}\phi_{j_2j_3}) \\
&- 2\psi'(0)(m_{i_1j_1}m_{i_2j_2}\sigma_{i_3i_4}\phi_{j_3j_4} \\
&+ m_{i_1j_1}m_{i_3j_3}\sigma_{i_2i_4}\sigma_{j_2j_4} + m_{i_1j_1}m_{i_4j_4}\,\sigma_{i_2i_3}\phi_{j_2j_3} \\
&+ m_{i_2j_2}m_{i_3j_3}\sigma_{i_1i_4}\phi_{j_1j_4} + m_{i_2j_2}m_{i_4j_4}\sigma_{i_1i_3}\phi_{j_1j_3} \\
&+ m_{i_3j_3}m_{i_4j_4}\sigma_{i_1i_2}\phi_{j_1j_2}) + m_{i_1j_1}m_{i_2j_2}m_{i_3j_3}m_{i_4j_4}. \blacksquare
\end{aligned}
$$

REMARK 3.2.1. The derivation of (i) and (ii) of Theorem 3.2.7 provides another proof of Theorem 2.4.1.

THEOREM 3.2.7. *Let* $X \sim E_{p,n}(M,\Sigma\otimes\Phi,\psi)$ *with finite second order moments. Let* $c_0 = -2\psi'(0)$. *Then, for any constant matrix* A, *if the expressions on the left-hand sides are defined, we have*

i) $\mathcal{E}(XAX) = c_0\Sigma A'\Phi + MAM$,

ii) $\mathcal{E}(XAX') = c_0\Sigma\mathrm{tr}(A'\Phi) + MAM'$,

iii) $\mathcal{E}(X'AX) = c_0\Phi\mathrm{tr}(\Sigma A') + M'AM$,

iv) $\mathcal{E}(X'AX') = c_0\Phi A'\Sigma + M'AM'$.

PROOF. i) $\mathcal{E}(XAX)_{ij} = \mathcal{E}\left(\sum_{\ell=1}^{p}\sum_{k=1}^{n} x_{ik}a_{k\ell}x_{\ell j}\right)$

$$= \sum_{\ell,k}(c_0\sigma_{ij}\phi_{kj} + m_{ik}m_{j\ell})a_{k\ell}$$

$$= c_0\left(\sum_{\ell,k}\sigma_{i\ell}a_{k\ell}\phi_{kj}\right) + \left(\sum_{\ell,k} m_{ik}a_{k\ell}m_{\ell j}\right)$$

$$= (c_0\Sigma\, A'\Phi + MAM)_{ij}.$$

ii) $$\mathcal{E}(XAX')_{ij} = \mathcal{E}\left(\sum_{\ell=1}^{p}\sum_{k=1}^{n} x_{ik}a_{k\ell}x_{j\ell}\right)$$

$$= \sum_{\ell,k} (c_0\sigma_{ij}\phi_{kj} + m_{ik}m_{j\ell})a_{k\ell}$$

$$= c_0\left(\sigma_{ij}\sum_{\ell,k}\phi_{k\ell}a_{k\ell}\right) + \sum_{\ell,k} m_{ik}a_{k\ell}m_{j\ell}$$

$$= (c_0\Sigma\,\mathrm{tr}(\Phi A) + MAM')_{ij}.$$

iii) From Theorem 2.1.3, it follows that $X' \sim E_{n,p}(M',\Phi \otimes \Sigma,\psi)$. Using (ii), we get $\mathcal{E}(X'AX) = c_0\Phi\,\mathrm{tr}(A'\Sigma) + M'AM$.

iv) Since, $X' \sim E_{n,p}(M',\Phi \otimes \Sigma,\psi)$, from (i) we have $\mathcal{E}(X'AX') = c_0\Phi A'\Sigma + M'AM'$. ∎

THEOREM 3.2.8. *Let* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ *with finite second order moment. Let* $c_0 = -2\psi'(0)$. *Then for any constant matrices* A, B, *if the expressions on the left-hand sides are defined, we have*

i) $\mathcal{E}(X\,\mathrm{tr}(AX)) = c_0\Sigma A'\Phi + M\,\mathrm{tr}(AM)$,

ii) $\mathcal{E}(X\,\mathrm{tr}(AX')) = c_0\Sigma A\Phi + M\,\mathrm{tr}(AM')$,

iii) $\mathcal{E}(X'\,\mathrm{tr}(AX)) = c_0\Phi A\Sigma + M'\,\mathrm{tr}(AM)$,

iv) $\mathcal{E}(X'\,\mathrm{tr}(AX')) = c_0\Phi A'\Sigma + M'\,\mathrm{tr}(AM')$,

v) $\mathcal{E}(\mathrm{tr}(XAXB)) = c_0\,\mathrm{tr}(\Sigma A'\Phi B) + \mathrm{tr}(MAMB)$,

vi) $\mathcal{E}(\mathrm{tr}(XAX'B)) = c_0\,\mathrm{tr}(\Sigma B)\,\mathrm{tr}(\Phi A') + \mathrm{tr}(MAM'B)$,

vii) $\mathcal{E}(\mathrm{tr}(XA)\,\mathrm{tr}(XB)) = c_0\,\mathrm{tr}(\Sigma B'\Phi A) + \mathrm{tr}(MA)\,\mathrm{tr}(MB)$,

viii) $\mathcal{E}(\mathrm{tr}(XA)\,\mathrm{tr}(X'B)) = c_0\,\mathrm{tr}(\Sigma B\Phi A) + \mathrm{tr}(MA)\,\mathrm{tr}(M'B)$.

PROOF.

i) $$\mathcal{E}(X\,\mathrm{tr}(AX))_{ij} = \mathcal{E}(x_{ij}\sum_{k=1}^{n}\sum_{\ell=1}^{p} a_{k\ell}x_{\ell k})$$

$$= \sum_{\ell,k} (c_0\sigma_{i\ell}\phi_{jk} + m_{ij}m_{\ell k})a_{k\ell}$$

$$= c_0\left(\sum_{\ell,k} \sigma_{i\ell} a_{k\ell} \phi_{kj}\right) + m_{ij} \sum_{\ell,k} a_{k\ell} m_{\ell k}$$

$$= (c_0 \Sigma A' \Phi + M\ \mathrm{tr}(AM))_{ij}.$$

ii) $\mathcal{E}(X\ \mathrm{tr}(AX')) = \mathcal{E}(X\ \mathrm{tr}(XA'))$

$$= \mathcal{E}(X\ \mathrm{tr}(A'X))$$

$$= c_0 \Sigma A \Phi + M\ \mathrm{tr}(A'M)$$

$$= c_0 \Sigma A \Phi + M\ \mathrm{tr}(AM').$$

iii) $X' \sim E_{n,p}(M', \Phi \otimes \Sigma, \psi)$.

$\mathcal{E}(X'\ \mathrm{tr}(AX)) = \mathcal{E}(X'\ \mathrm{tr}(A'X'))$

$$= c_0 \Phi A \Sigma + M'\ \mathrm{tr}(A'M')$$

$$= c_0 \Phi A \Sigma + M'\ \mathrm{tr}(AM).$$

iv) $X' \sim E_{n,p}(M', \Phi \otimes \Sigma, \psi)$.

$\mathcal{E}(X'\ \mathrm{tr}(AX')) = c_0 \Phi A' \Sigma + M'\ \mathrm{tr}(AM')$.

v) $\mathcal{E}(\mathrm{tr}(XAXB)) = \mathcal{E}\left(\sum_{i=1}^{p} \sum_{j=1}^{n} \sum_{k=1}^{p} \sum_{\ell=1}^{n} x_{ij} a_{jk} x_{k\ell} b_{\ell i}\right)$

$$= \sum_{i,j,k,\ell} (c_0 \sigma_{ik} \phi_{j\ell} + m_{ij} m_{k\ell}) a_{jk} b_{\ell i}$$

$$= c_0\left(\sum_{i,j,k,\ell} \sigma_{ik} a_{jk} \phi_{j\ell} b_{\ell i}\right) + \sum_{i,j,k,\ell} m_{ij} a_{jk} m_{k\ell} b_{\ell i}$$

$$= c_0\ \mathrm{tr}(\Sigma A' \Phi B) + \mathrm{tr}(MAMB).$$

vi) $\mathcal{E}(\mathrm{tr}(XAX'B)) = \mathcal{E}\left(\sum_{i=1}^{p} \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{\ell=1}^{p} x_{ij} a_{jk} x_{\ell k} b_{\ell i}\right)$

$$= \sum_{i,j,k,\ell} (c_0 \sigma_{i\ell} \phi_{jk} + m_{ij} m_{\ell k}) a_{jk} b_{\ell i}$$

$$= c_0\left(\sum_{i,\ell} \sigma_{i\ell} b_{\ell i}\right)\left(\sum_{j,k} \phi_{jk} a_{jk}\right)$$

$$+ \sum_{i,j,k,\ell} m_{ij} a_{jk} m_{\ell k} b_{\ell i}$$

$$= c_0\ \mathrm{tr}(\Sigma B)\ \mathrm{tr}(\Phi A) + \mathrm{tr}(MAM'B)$$

$$= c_0 \operatorname{tr}(\Sigma B) \operatorname{tr}(\Phi A') + \operatorname{tr}(MAM'B).$$

vii) $$\mathcal{E}(\operatorname{tr}(XA)\operatorname{tr}(XB)) = \mathcal{E}\left(\left(\sum_{i=1}^{p}\sum_{\ell=1}^{n} x_{i\ell}a_{\ell i}\right)\left(\sum_{j=1}^{p}\sum_{k=1}^{n} x_{jk}b_{kj}\right)\right)$$

$$= \sum_{i,j,k,\ell} (c_0\sigma_{ij}\phi_{\ell k} + m_{i\ell}m_{jk})a_{\ell i}b_{kj}$$

$$= c_0 \sum_{i,j,k,\ell} \sigma_{ij}b_{kj}\phi_{\ell k}a_{\ell i} + \left(\sum_{i,\ell} m_{i\ell}a_{\ell i}\right)\left(\sum_{j,k} m_{jk}b_{kj}\right)$$

$$= c_0 \operatorname{tr}(\Sigma B'\Phi A) + \operatorname{tr}(MA)\operatorname{tr}(MB).$$

viii) $$\mathcal{E}(\operatorname{tr}(XA)\operatorname{tr}(X'B)) = \mathcal{E}(\operatorname{tr}(XA)\operatorname{tr}(XB'))$$
$$= c_0 \operatorname{tr}(\Sigma B\Phi A) + \operatorname{tr}(MA)\operatorname{tr}(MB')$$
$$= c_0 \operatorname{tr}(\Sigma B\Phi A) + \operatorname{tr}(MA)\operatorname{tr}(M'B). \blacksquare$$

THEOREM 3.2.9. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ *with finite third order moment. Let* $c_0 = -2\psi'(0)$. *Then, for any constant matrices* A *and* B, *if the expressions on the left-hand sides are defined, we have*

i) $\mathcal{E}(XAXBX) = c_0(MA\Sigma B'\Phi + \Sigma B'M'A'\Phi + \Sigma A'\Phi BM) + MAMBM,$

ii) $\mathcal{E}(X'AXBX) = c_0(M'A\Sigma B'\Phi + \Phi \operatorname{tr}(\Sigma B'M'A') + \Phi BM \operatorname{tr}(A\Sigma)) + M'AMBM,$

iii) $\mathcal{E}(X'AX'BX) = c_0(M'A\Phi \operatorname{tr}(\Sigma B') + \Phi \operatorname{tr}(AM'B\Sigma) + \Phi A'\Sigma BM) + M'AM'BM,$

iv) $\mathcal{E}(X'AXBX') = c_0(M'A\Sigma \operatorname{tr}(B\Phi) + \Phi B'M'A'\Sigma + \Phi BM' \operatorname{tr}(A\Sigma)) + M'AMBM',$

v) $\mathcal{E}(XAX'BX') = c_0(MA\Phi B'\Sigma + \Sigma \operatorname{tr}(AM'B\Phi) + \Sigma BM' \operatorname{tr}(A\Phi)) + MAM'BM',$

vi) $\mathcal{E}(X'AX'BX') = c_0(M'A\Phi B'\Sigma + \Phi B'MA'\Sigma + \Phi A'\Sigma BM') + M'AM'BM',$

vii) $\mathcal{E}(XAX'BX) = c_0(MA\Phi \operatorname{tr}(B\Sigma) + \Sigma B'MA'\Phi + \Sigma BM \operatorname{tr}(A\Phi)) + MAM'BM,$

viii) $\mathcal{E}(XAXBX') = c_0(MA\Sigma \operatorname{tr}(\Phi B') + \Sigma \operatorname{tr}(AMB\Phi) + \Sigma A'\Phi BM') + MAMBM'.$

PROOF.

i) $$\mathcal{E}(XAXBX)_{ij} = \mathcal{E}\left(\sum_{k=1}^{n}\sum_{\ell=1}^{p}\sum_{r=1}^{n}\sum_{q=1}^{p} x_{ik}a_{k\ell}x_{\ell r}b_{rq}x_{qj}\right)$$

$$= \sum_{i,j,k,\ell} [c_0(\sigma_{i\ell}\phi_{kr}m_{qj} + \sigma_{iq}\phi_{kj}m_{\ell r} + \sigma_{\ell q}\phi_{rj}m_{ik})$$
$$+ m_{ik}m_{\ell r}m_{qj}]a_{k\ell}b_{rq}$$
$$= c_0\left[\sum_{k,\ell,r,q}\sigma_{i\ell}a_{k\ell}\phi_{kr}b_{rq}m_{qj} + \sum_{k,\ell,r,q}\sigma_{iq}b_{rq}m_{\ell r}a_{k\ell}\phi_{kj}\right.$$
$$\left.+ \sum_{k,\ell,r,q} m_{ik}a_{k\ell}\sigma_{\ell q}b_{rq}\phi_{rj}\right]$$
$$+ \sum_{k,\ell,r,q} m_{ik}a_{k\ell}m_{\ell r}b_{rq}m_{qj}$$
$$= (c_0(\Sigma A'\Phi BM + \Sigma B'M'A'\Phi + MA\Sigma B'\Phi)$$
$$+ MAMABM)_{ij}.$$

ii) $$\mathcal{E}(X'AXBX)_{ij} = \mathcal{E}\left(\sum_{k=1}^{p}\sum_{\ell=1}^{p}\sum_{r=1}^{n}\sum_{q=1}^{p} x_{ki}a_{k\ell}x_{\ell r}b_{rq}x_{qj}\right)$$

$$= \sum_{k,\ell,r,q} [c_0(\sigma_{k\ell}\phi_{ir}m_{qj} + \sigma_{kq}\phi_{ij}m_{\ell r} + \sigma_{\ell q}\phi_{rj}m_{ki})$$
$$+ m_{ki}m_{\ell r}m_{qj}]a_{k\ell}b_{rq}$$
$$= c_0\left[\left(\sum_{k,\ell} a_{k\ell}\sigma_{\ell k}\right)\left(\sum_{r,q}\phi_{ir}b_{rq}m_{qj}\right)\right.$$
$$+ \phi_{ij}\sum_{k,\ell,r,q}\sigma_{kq}b_{rq}m_{\ell r}a_{k\ell}$$
$$\left.+ \sum_{k,\ell,r,q} m_{ki}a_{k\ell}\sigma_{\ell q}b_{rq}\phi_{rj}\right] + \sum_{k,\ell,r,q} m_{ki}a_{k\ell}m_{\ell r}b_{rq}m_{qj}$$
$$= (c_0(\mathrm{tr}(A\Sigma)\Phi BM + \Phi\ \mathrm{tr}(\Sigma B'M'A') + M'A\Sigma B'\Phi)$$
$$+ M'AMBM)_{ij}.$$

iii) $$\mathcal{E}(X'AX'BX) = \mathcal{E}(X'B'XA'X)'$$
$$= (c_0(\mathrm{tr}(B'\Sigma)\Phi A'M + \Phi\ \mathrm{tr}(\Sigma AM'B) + M'B'\Sigma A\Phi) + M'B'MA'M)'$$
$$= c_0(M'A\Phi\ \mathrm{tr}(\Sigma B') + \Phi\ \mathrm{tr}(AM'B\Sigma) + \Phi A'\Sigma BM) + M'AM'BM.$$

iv) $$\mathcal{E}(X'AXBX')_{ij} = \mathcal{E}\left(\sum_{k=1}^{p}\sum_{\ell=1}^{p}\sum_{r=1}^{n}\sum_{q=1}^{n} x_{ki}a_{k\ell}x_{\ell r}b_{rq}x_{jq}\right)$$
$$= \sum_{k,\ell,r,q} [c_0(\sigma_{k\ell}\phi_{ir}m_{jq} + \sigma_{kj}\phi_{iq}m_{\ell r} + \sigma_{\ell j}\phi_{rq}m_{ki}) + m_{ki}m_{\ell r}m_{jq}]a_{k\ell}b_{rq}$$
$$= c_0\left[\left(\sum_{k,\ell} a_{k\ell}\sigma_{\ell k}\right)\left(\sum_{r,q}\phi_{ir}b_{rq}m_{jq}\right) + \sum_{k,\ell,r,q}\phi_{iq}b_{rq}m_{\ell r}a_{k\ell}\sigma_{kj}) + \left(\sum_{\ell,k} m_{ki}a_{k\ell}\sigma_{\ell j}\right)\left(\sum_{r,q} b_{rq}\phi_{qr}\right)\right]$$
$$+ \sum_{k,\ell,r,q} m_{ki}a_{k\ell}m_{\ell r}b_{rq}m_{jq}$$
$$= (c_0[\Phi BM'\ \mathrm{tr}(A\Sigma) + \Phi B'M'A'\Sigma + M'A\Sigma\ \mathrm{tr}(B\Phi)] + M'AMBM')_{ij}.$$

v) $X' \sim E_{n,p}(M',\Phi \otimes \Sigma,\psi)$.
$$\mathcal{E}(XAX'BX') = c_0(MA\Phi B'\Sigma + \Sigma\ \mathrm{tr}(\Phi B'MA') + \Sigma BM'\ \mathrm{tr}(A\Phi)) + MAM'BM'$$
$$= c_0(MA\Phi B'\Sigma + \Sigma\ \mathrm{tr}(AM'B\Phi) + \Sigma BM'\ \mathrm{tr}(A\Phi)) + MAM'BM'.$$

vi) $X' \sim E_{n,p}(M',\Phi \otimes \Sigma,\psi)$.
$$\mathcal{E}(X'AX'BX') = c_0(M'A\Phi B'\Sigma + \Phi B'MA'\Sigma + \Phi A'\Sigma BM') + M'AM'BM'.$$

vii) $$\mathcal{E}(XAX'BX) = (\mathcal{E}(X'B'XA'X'))'$$
$$= (c_0(M'B'\Sigma\ \mathrm{tr}(A'\Phi) + \Phi AM'B\Sigma + \Phi A'M'\ \mathrm{tr}(B'\Sigma)) + M'B'MA'M')'$$

$$= c_0(\Sigma BM\ \mathrm{tr}(A\Phi) + \Sigma B'MA'\Phi + MA\Phi\ \mathrm{tr}(B\Sigma)) + MAM'BM.$$

viii) $\mathcal{E}(XAXBX') = (\mathcal{E}(XB'X'A'X'))'$

$$= (c_0(MB'\Phi A\Sigma + \Sigma\ \mathrm{tr}(B'M'A'\Phi) + \Sigma A'M'\ \mathrm{tr}(B'\Phi)) + MB'M'A'M')'$$

$$= c_0(\Sigma A'\Phi BM' + \Sigma\ \mathrm{tr}(AMB\Phi) + MA\Sigma\ \mathrm{tr}(\Phi B')) + MAMBM'. \blacksquare$$

THEOREM 3.2.10. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ *with finite third order moment. Let* $c_0 = -2\psi'(0)$. *Then, for any constant matrices* A *and* B, *if the expressions on the left-hand sides are defined, we have*

i) $\mathcal{E}(X\ \mathrm{tr}(X'AXB)) = c_0(M\ \mathrm{tr}(A'\Sigma)\ \mathrm{tr}(B'\Phi) + \Sigma A'MB'\Phi + \Sigma AMB\Phi) + M\ \mathrm{tr}(M'AMB),$

ii) $\mathcal{E}(XBX\ \mathrm{tr}(AX)) = c_0(MB\Sigma A'\Phi + \Sigma A'\Phi BM + \Sigma B'\Phi\ \mathrm{tr}(AM)) + MBM\ \mathrm{tr}(AM),$

iii) $\mathcal{E}(X'BX\ \mathrm{tr}(AX)) = c_0(M'B\Sigma A'\Phi + \Phi A\Sigma BM + \Phi\ \mathrm{tr}(AM)\ \mathrm{tr}(B\Sigma)) + M'BM\ \mathrm{tr}(AM).$

PROOF.

i) $$\mathcal{E}(X\ \mathrm{tr}(X'AXB))_{ij} = \mathcal{E}\left(x_{ij}\sum_{k=1}^{n}\sum_{\ell=1}^{p}\sum_{q=1}^{n}\sum_{r=1}^{p} x_{\ell k}a_{\ell r}x_{rq}b_{qk}\right)$$

$$= \sum_{k,\ell,r,q} [c_0(\sigma_{i\ell}\phi_{jk}m_{rq} + \sigma_{ir}\phi_{jq}m_{\ell k} + \sigma_{\ell r}\phi_{kq}m_{ij}) + m_{ij}m_{\ell k}m_{rq}]a_{\ell r}b_{qr}$$

$$= c_0\left[\sum_{k,\ell,r,q} \sigma_{i\ell}a_{\ell r}m_{rq}b_{qk}\phi_{jk} + \sum_{k,\ell,r,q} \sigma_{ir}a_{\ell r}m_{\ell k}b_{qk}\phi_{jq} + \left(\sum_{k,q} b_{qk}\phi_{qk}\right)\left(\sum_{r,\ell} a_{\ell r}\sigma_{\ell r}\right)\right]$$

$$+ m_{ij} \sum_{k,\ell,r,q} m_{\ell k} a_{\ell r} m_{rq} b_{qk}$$

$$= (c_0(\Sigma AMB\Phi + \Sigma A'MB'\Phi + M \operatorname{tr}(A'\Sigma) \operatorname{tr}(B'\Phi)) + M \operatorname{tr}(M'AMB))_{ij}.$$

ii) $$\mathcal{E}(XBX \operatorname{tr}(AX))_{ij} = \mathcal{E}\left(\left(\sum_{k=1}^{n}\sum_{\ell=1}^{p} x_{ik} b_{k\ell} x_{\ell j}\right)\left(\sum_{r=1}^{n}\sum_{q=1}^{p} a_{rq} x_{qr}\right)\right)$$

$$= \sum_{k,\ell,r,q} [c_0(\sigma_{i\ell}\phi_{kj} m_{qr} + \sigma_{iq}\phi_{kr} m_{\ell j} + \sigma_{\ell q}\phi_{jr} m_{ik}) + m_{iik} m_{\ell j} m_{qr}] b_{k\ell} a_{rq}$$

$$= c_0\left[\left(\sum_{k,\ell} \sigma_{i\ell} b_{k\ell}\phi_{kj}\right)\left(\sum_{r,q} a_{rq} m_{qr}\right) + \sum_{k,\ell,r,q} \sigma_{iq} a_{rq}\phi_{kr} b_{k\ell} m_{\ell j} + \sum_{k,\ell,r,q} m_{ik} b_{k\ell}\sigma_{\ell q} a_{rq}\phi_{jr}\right] + \left(\sum_{k,\ell} m_{ik} b_{k\ell} m_{\ell j}\right) \cdot \left(\sum_{r,q} a_{rq} m_{qr}\right)$$

$$= (c_0(\Sigma B'\Phi \operatorname{tr}(AM) + \Sigma A'\Phi BM + MB\Sigma A'\Phi) + MBM \operatorname{tr}(AM))_{ij}.$$

iii) $$\mathcal{E}(X'BX \operatorname{tr}(AX))_{ij} = \mathcal{E}\left(\left(\sum_{k=1}^{p}\sum_{\ell=1}^{p} x_{ki} b_{k\ell} x_{\ell j}\right)\left(\sum_{r=1}^{n}\sum_{q=1}^{p} a_{rq} x_{qr}\right)\right)$$

$$= \sum_{k,\ell,r,q} [c_0(\sigma_{k\ell}\phi_{ij} m_{qr} + \sigma_{kq}\phi_{ir} m_{\ell j} + \sigma_{\ell q}\phi_{jr} m_{ki}) + m_{ki} m_{\ell j} m_{qr}] b_{k\ell} a_{rq}$$

$$= c_0\left[\phi_{ij}\left(\sum_{k,\ell} b_{k\ell}\sigma_{\ell k}\right)\left(\sum_{r,q} a_{rq} m_{qr}\right) + \sum_{k,\ell,r,q} \phi_{ir} a_{rq}\sigma_{kq} b_{k\ell} m_{\ell j}\right.$$

$$+ \sum_{k,\ell,r,q} m_{ki} b_{k\ell} \sigma_{\ell q} a_{rq} \phi_{jr} \Bigg] + \left(\sum_{k,\ell} m_{ki} b_{k\ell} m_{\ell j} \right)$$

$$\cdot \left(\sum_{r,q} a_{rq} m_{qr} \right)$$

$$= (c_0(\Phi \operatorname{tr}(B\Sigma) \operatorname{tr}(AM) + \Phi A \Sigma B M + M'B\Sigma A'\Phi)$$

$$+ M'BM \operatorname{tr}(AM))_{ij}. \blacksquare$$

THEOREM 3.2.11. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ *with finite fourth order moments. Let* $c_0 = -2\psi'(0)$ *and* $k_0 = 4\psi''(0)$. *Then, for any constant matrices* A, B, *and* C, *if the expressions on the left-hand sides are defined, we have*

i) $$\mathcal{E}(XAXBXCX) = k_0(\Sigma C'\Phi BA'\Phi + \Sigma A'\Phi B\Sigma C'\Phi + \Sigma B'\Phi \operatorname{tr}(A\Sigma C'\Phi)) + c_0(MAMB\Sigma C'\Phi + MA\Sigma C'M'B'\Phi + \Sigma C'M'B'M'A'\Phi + MA\Sigma B'\Phi CM + \Sigma B'M'A'\Phi CM + \Sigma A'\Phi BMCM) + MAMBMCM,$$

ii) $$\mathcal{E}(X'AXBXCX) = k_0(\Phi \operatorname{tr}(\Sigma C'\Phi B\Sigma A') + \Phi B\Sigma C'\Phi \operatorname{tr}(A\Sigma) + \Phi C\Sigma A'\Sigma B'\Phi) + c_0(M'AMB\Sigma C'\Phi + M'A\Sigma C'M'B'\Phi + \Phi \operatorname{tr}(AMBMC\Sigma) + M'A\Sigma B'\Phi CM + \Phi CM \operatorname{tr}(AMB\Sigma) + \Phi BMCM \operatorname{tr}(A\Sigma)) + M'AMBMCM,$$

iii) $$\mathcal{E}(X'AX'BXCX) = k_0(\Phi \operatorname{tr}(\Sigma C'\Phi A') \operatorname{tr}(B\Sigma) + \Phi A'\Sigma B\Sigma C'\Phi + \Phi C\Sigma B\Sigma A\Phi) + c_0(M'AM'B\Sigma C'\Phi + M'A\Phi \operatorname{tr}(MC\Sigma B) + \Phi \operatorname{tr}(AM'BMC\Sigma) + M'A\Phi CM \operatorname{tr}(B\Sigma) + \Phi CM \operatorname{tr}(AM'B\Sigma) + \Phi A'\Sigma BMCM) + M'AM'BMCM,$$

iv) $$\mathcal{E}(X'AXBX'CX) = k_0(\Phi \operatorname{tr}(\Sigma C'\Sigma A') \operatorname{tr}(B\Phi) + \Phi B\Phi \operatorname{tr}(A\Sigma) \operatorname{tr}(C\Sigma) + \Phi B'\Phi \operatorname{tr}(A\Sigma C'\Sigma)) + c_0(M'AMB\Phi \operatorname{tr}(\Sigma C) + M'A\Sigma C'MB'\Phi + \Phi \operatorname{tr}(AMBM'C\Sigma) + M'A\Sigma CM \operatorname{tr}(B\Phi) + \Phi B'M'A'\Sigma CM + \Phi BM'CM \operatorname{tr}(A\Sigma)) + M'AMBM'CM,$$

v) $$\mathcal{E}(XAX'BXCX) = k_0(\Sigma B\Sigma C'\Phi \operatorname{tr}(\Phi A) + \Sigma B'\Sigma C'\Phi A'\Phi + \Sigma C'\Phi A' \operatorname{tr}(\Sigma B)) + c_0(MAM'B\Sigma C'\Phi + MA\Phi \operatorname{tr}(BMC\Sigma) + MA\Phi CM \operatorname{tr}(\Sigma B)$$

$$+ \Sigma C'M'B'MA'\Phi + \Sigma B'MA'\Phi CM + \Sigma BMCM\ \mathrm{tr}(\Phi A)) + MAM'BMCM,$$

vi) $$\mathcal{E}(X'AXBXCX') = k_0(\Phi B\Sigma\ \mathrm{tr}(\Sigma A)\ \mathrm{tr}(\Phi C) + \Phi C\Phi B\Sigma A\Sigma + \Phi C'\Phi B\Sigma A'\Sigma) + c_0(M'AMB\Sigma\ \mathrm{tr}(\Phi C) + M'A\Sigma\ \mathrm{tr}(BMC\Phi) + M'A\Sigma B'\Phi CM' + \Phi C'M'B'M'A'\Sigma + \Phi CM'\ \mathrm{tr}(AMB\Sigma) + \Phi BMCM'\ \mathrm{tr}(\Sigma A)) + M'AMBMCM'.$$

PROOF.

i) $$\mathcal{E}(XAXBXCX)_{ij} = \mathcal{E}\left(\sum_{q=1}^{p}\sum_{r=1}^{n}\sum_{\ell=1}^{p}\sum_{k=1}^{n}\sum_{t=1}^{p}\sum_{s=1}^{n} x_{is}a_{st}x_{tk}b_{k\ell}x_{\ell r}c_{rq}x_{qj}\right)$$

$$\begin{aligned}
= \sum_{\substack{q,r,\ell\\k,t,s}} \{&k_0[\sigma_{it}\phi_{sk}\sigma_{\ell q}\phi_{rj} + \sigma_{i\ell}\phi_{sr}\sigma_{tq}\phi_{kj} + \sigma_{iq}\phi_{sj}\sigma_{t\ell}\phi_{kr}]\\
&+ c_0[m_{is}m_{tk}\sigma_{\ell q}\phi_{rj} + m_{is}m_{\ell r}\sigma_{tq}\phi_{kj}\\
&+ m_{is}m_{qj}\sigma_{t\ell}\phi_{kr} + m_{tk}m_{\ell r}\sigma_{iq}\phi_{sj} + m_{tk}m_{qj}\sigma_{i\ell}\phi_{sr}\\
&+ m_{\ell r}m_{qj}\sigma_{it}\phi_{sk}] + m_{is}m_{tk}m_{\ell r}m_{qj}\}a_{st}b_{k\ell}c_{rq}
\end{aligned}$$

$$\begin{aligned}
= \sum_{\substack{q,r,\ell\\k,t,s}} \{&k_0[\sigma_{it}a_{st}\phi_{sk}b_{k\ell}\sigma_{\ell q}c_{rq}\phi_{rj}\\
&+ (\sigma_{i\ell}b_{k\ell}\phi_{kj})(\phi_{sr}c_{rq}\sigma_{tq}a_{st})\\
&+ \sigma_{iq}c_{rq}\phi_{kr}b_{k\ell}\sigma_{t\ell}a_{st}\phi_{sj}]\\
&+ c_0[m_{is}a_{st}m_{tk}b_{k\ell}\sigma_{\ell q}c_{rq}\phi_{rj}\\
&+ m_{is}a_{st}\sigma_{tq}c_{rq}m_{\ell r}b_{k\ell}\phi_{kj}\\
&+ m_{is}a_{st}\sigma_{t\ell}b_{k\ell}\phi_{kr}c_{rq}m_{qj}\\
&+ \sigma_{iq}c_{rq}m_{\ell r}b_{k\ell}m_{tk}a_{st}\phi_{sj}\\
&+ \sigma_{i\ell}b_{k\ell}m_{tk}a_{st}\phi_{sr}c_{rq}m_{qj}\\
&+ \sigma_{it}a_{st}\phi_{sk}b_{k\ell}m_{\ell r}c_{rq}m_{qj}]\\
&+ m_{is}a_{st}m_{tk}b_{k\ell}m_{\ell r}c_{rq}m_{qj}\}
\end{aligned}$$

$$\begin{aligned}
= (&k_0[\Sigma A'\Phi B\Sigma C'\Phi + (\Sigma B'\Phi)\ \mathrm{tr}(\Phi C\Sigma A')\\
&+ \Sigma C'\Phi B\Sigma A'\Phi] + c_0[MAMB\Sigma C'\Phi\\
&+ MA\Sigma C'M'B'\Phi + MA\Sigma B'\Phi CM\\
&+ \Sigma C'M'B'M'A'\Phi + \Sigma B'M'A'\Phi CM\\
&+ \Sigma A'\Phi BMCM] + MAMBMCM)_{ij}.
\end{aligned}$$

ii) $$\mathcal{E}(X'AXBXCX)_{ij} = \mathcal{E}\left(\sum_{q=1}^{p}\sum_{r=1}^{n}\sum_{\ell=1}^{p}\sum_{k=1}^{n}\sum_{t=1}^{p}\sum_{s=1}^{p} x_{si}a_{st}x_{tk}b_{k\ell}x_{\ell r}c_{rq}x_{qj}\right)$$

$$\begin{aligned}
&= \sum_{\substack{q,r,\ell\\k,t,s}} \{k_0[\sigma_{st}\phi_{ik}\sigma_{\ell q}\phi_{rj} + \sigma_{s\ell}\phi_{ir}\sigma_{tq}\phi_{kj} \\
&\quad + \sigma_{sq}\phi_{ij}\sigma_{t\ell}\phi_{kr}] + c_0[m_{si}m_{tk}\sigma_{\ell q}\phi_{rj} \\
&\quad + m_{si}m_{\ell r}\sigma_{tq}\phi_{kj} + m_{si}m_{qj}\sigma_{t\ell}\phi_{kr} \\
&\quad + m_{tk}m_{\ell r}\sigma_{sq}\phi_{ij} + m_{tk}m_{qj}\sigma_{s\ell}\phi_{ir} \\
&\quad + m_{\ell r}m_{qj}\sigma_{st}\phi_{ik}] + m_{si}m_{tk}m_{\ell r}m_{qj}\}\, a_{st}b_{k\ell}c_{rq}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{\substack{q,r,\ell\\k,t,s}} k_0[(\phi_{ik}b_{k\ell}\sigma_{\ell q}c_{rq}\phi_{rj})(\sigma_{st}a_{st}) \\
&\quad + \phi_{ir}c_{rq}\sigma_{tq}a_{st}\sigma_{s\ell}b_{k\ell}\phi_{kj} \\
&\quad + (\phi_{ij})(\sigma_{sq}a_{st}\sigma_{t\ell}b_{k\ell}\phi_{kr}c_{rq})] \\
&\quad + c_0[(m_{si}a_{st}m_{tk}b_{k\ell}\sigma_{\ell q}c_{rq}\phi_{rj}) \\
&\quad + m_{si}a_{st}\sigma_{tq}c_{rq}m_{\ell r}b_{k\ell}\phi_{kj} \\
&\quad + m_{si}a_{st}\sigma_{t\ell}b_{k\ell}\phi_{kr}c_{rq}m_{qj} \\
&\quad + (\phi_{ij})(\sigma_{sq}c_{rq}m_{\ell r}b_{k\ell}m_{tk}a_{st}) \\
&\quad + (\phi_{ir}c_{rq}m_{qj})(\sigma_{s\ell}b_{k\ell}m_{tk}a_{st}) \\
&\quad + (\phi_{ik}b_{k\ell}m_{\ell r}c_{rq}m_{qj})(\sigma_{st}a_{st})] \\
&\quad + m_{si}a_{st}m_{tk}b_{k\ell}m_{\ell r}c_{rq}m_{qj} \\
&= (k_0[\Phi B\Sigma C'\Phi\ \mathrm{tr}(\Sigma A) + \Phi C\Sigma A'\Sigma B'\Phi \\
&\quad + \Phi\ \mathrm{tr}(\Sigma A\Sigma B'\Phi C)] + c_0[M'AMB\Sigma C'\Phi \\
&\quad + M'A\Sigma C'M'B'\Phi + M'A\Sigma B'\Phi CM \\
&\quad + \Phi\ \mathrm{tr}(\Sigma C'M'B'M'A') + \Phi CM\ \mathrm{tr}(\Sigma B'M'A') \\
&\quad + \Phi BMCM\ \mathrm{tr}(\Sigma A)] + M'AMBMCM)_{ij}.
\end{aligned}$$

iii) $$\mathcal{E}(X'AX'BXCX)_{ij} = \mathcal{E}\left(\sum_{q=1}^{p}\sum_{r=1}^{n}\sum_{\ell=1}^{p}\sum_{k=1}^{p}\sum_{t=1}^{n}\sum_{s=1}^{p} x_{si}a_{st}x_{kt}b_{k\ell}x_{\ell r}c_{rq}x_{qj}\right)$$

$$\begin{aligned}
&= \sum_{\substack{q,r,\ell\\k,t,s}} \{k_0[\sigma_{sk}\phi_{it}\sigma_{\ell q}\phi_{rj} + \sigma_{s\ell}\phi_{ir}\sigma_{kq}\phi_{tj} \\
&\quad + \sigma_{sq}\phi_{ij}\sigma_{k\ell}\phi_{tr}] + c_0[m_{si}m_{kt}\sigma_{\ell q}\phi_{rj}
\end{aligned}$$

$$+ m_{si}m_{\ell r}\sigma_{kq}\phi_{tj} + m_{si}m_{qj}\sigma_{k\ell}\phi_{tr}$$
$$+ m_{kt}m_{\ell r}\sigma_{sq}\phi_{ij} + m_{kt}m_{qj}\sigma_{s\ell}\phi_{ir}$$
$$+ m_{\ell r}m_{qj}\sigma_{sk}\phi_{it}] + m_{si}m_{kt}m_{\ell r}m_{qj}\}a_{st}b_{k\ell}c_{rq}$$

$$= \sum_{\substack{q,r,\ell \\ k,t,s}} k_0[(\phi_{it}a_{st}\sigma_{sk}b_{k\ell}\sigma_{\ell q}c_{rq}\phi_{rj}$$
$$+ \phi_{ir}c_{rq}\sigma_{kq}b_{k\ell}\sigma_{s\ell}a_{st}\phi_{tj}$$
$$+ (\phi_{ij})(\sigma_{sq}c_{rq}\phi_{tr}a_{st})(\sigma_{k\ell}b_{k\ell})]$$
$$+ c_0[m_{si}a_{st}m_{kt}b_{k\ell}\sigma_{\ell q}c_{rq}\phi_{rj}$$
$$+ (m_{si}a_{st}\phi_{tj})(\sigma_{kq}c_{rq}m_{\ell r}b_{k\ell})$$
$$+ (m_{si}a_{st}\phi_{tr}c_{rq}m_{qj})(\sigma_{k\ell}b_{k\ell})$$
$$+ (\phi_{ij})(\sigma_{sq}c_{rq}m_{\ell r}b_{k\ell}m_{kt}a_{st})$$
$$+ (\phi_{ir}c_{rq}m_{qj})(\sigma_{s\ell}b_{k\ell}m_{kt}a_{st})$$
$$+ \phi_{it}a_{st}\sigma_{sk}b_{k\ell}m_{\ell r}c_{rq}m_{qj}$$
$$+ m_{si}a_{st}m_{kt}b_{k\ell}m_{\ell r}c_{rq}m_{qj}$$
$$= (k_0[\Phi A'\Sigma B\Sigma C'\Phi + \Phi C\Sigma BA\Phi$$
$$+ \Phi\ \mathrm{tr}(\Sigma C'\Phi A')\ \mathrm{tr}(\Sigma B)] + c_0[M'AM'B\Sigma C'\Phi$$
$$+ M'A\Phi\ \mathrm{tr}(\Sigma C'M'B') + M'A\Phi CM\ \mathrm{tr}(\Sigma B)$$
$$+ \Phi\ \mathrm{tr}(\Sigma C'M'B'MA') + \Phi CM\ \mathrm{tr}(\Sigma B'MA')$$
$$+ \Phi A'\Sigma BMCM] + M'AM'BMCM)_{ij}.$$

iv) $$\mathcal{E}(X'AXBX'CX)_{ij} = \mathcal{E}\left(\sum_{q=1}^{p}\sum_{r=1}^{p}\sum_{\ell=1}^{n}\sum_{k=1}^{n}\sum_{t=1}^{p}\sum_{s=1}^{p} x_{si}a_{st}x_{tk}b_{k\ell}x_{r\ell}c_{rq}x_{qj}\right)$$

$$= \sum_{\substack{q,r,\ell \\ k,t,s}} \{k_0[\sigma_{st}\phi_{ik}\sigma_{rq}\phi_{\ell j} + \sigma_{sr}\phi_{i\ell}\sigma_{tq}\phi_{kj}$$
$$+ \sigma_{sq}\phi_{ij}\sigma_{tr}\phi_{k\ell}] + c_0[m_{si}m_{tk}\sigma_{rq}\phi_{\ell j}$$
$$+ m_{si}m_{r\ell}\sigma_{tq}\phi_{kj} + m_{si}m_{qj}\sigma_{tr}\phi_{k\ell}$$
$$+ m_{tk}m_{r\ell}\sigma_{sq}\phi_{ij} + m_{tk}m_{qj}\sigma_{sr}\phi_{i\ell}$$
$$+ m_{r\ell}m_{qj}\sigma_{st}\phi_{ik}] + m_{si}m_{tk}m_{r\ell}m_{qj}\}a_{st}b_{k\ell}c_{rq}$$

$$= \sum_{\substack{q,r,\ell \\ k,t,s}} k_0[(\phi_{ik}b_{k\ell}\phi_{\ell j})(\sigma_{st}a_{st})(\sigma_{rq}c_{rq})$$
$$+ (\phi_{i\ell}b_{k\ell}\phi_{kj})(\sigma_{sr}c_{rq}\sigma_{tq}a_{st})$$

$$
\begin{aligned}
&+ (\phi_{ij})(\sigma_{sq}c_{rq}\sigma_{tr}a_{st})(\phi_{k\ell}b_{k\ell})] \\
&+ c_0[(m_{si}a_{st}m_{tk}b_{k\ell}\phi_{\ell j})(\sigma_{rq}c_{rq}) \\
&+ m_{si}a_{st}\sigma_{tq}c_{rq}m_{r\ell}b_{k\ell}\phi_{kj} \\
&+ (m_{si}a_{st}\sigma_{tr}c_{rq}m_{qj})(\phi_{k\ell}b_{k\ell}) \\
&+ (\phi_{ij})(\sigma_{sq}c_{rq}m_{r\ell}b_{k\ell}m_{tk}a_{st}) \\
&+ \phi_{i\ell}b_{k\ell}m_{tk}a_{st}\sigma_{sr}c_{rq}m_{qj} \\
&+ (\phi_{ij}b_{k\ell}m_{r\ell}c_{rq}m_{qj})(\sigma_{st}a_{st})] \\
&+ m_{si}a_{st}m_{tk}b_{k\ell}m_{r\ell}c_{rq}m_{qj} \\
= (&k_0[\Phi B\Phi\ \mathrm{tr}(\Sigma A)\ \mathrm{tr}(\Sigma C) + (\Phi B'\Phi)\ \mathrm{tr}(\Sigma C\Sigma A') \\
&+ \Phi\ \mathrm{tr}(\Sigma C'\Sigma A')\ \mathrm{tr}(\Phi B)] + c_0[M'AMB\Phi\ \mathrm{tr}(\Sigma C) \\
&+ M'A\Sigma C'MB'\Phi + M'A\Sigma CM\ \mathrm{tr}(\Phi B) \\
&+ \Phi\ \mathrm{tr}(\Sigma C'MB'M'A') + \Phi B'M'A'\Sigma CM \\
&+ \Phi BM'CM\ \mathrm{tr}(\Sigma A)] + M'AMBM'CM)_{ij}.
\end{aligned}
$$

v) $$\mathcal{E}(XAX'BXCX)_{ij} = \mathcal{E}\left(\sum_{q=1}^{p}\sum_{r=1}^{n}\sum_{\ell=1}^{p}\sum_{k=1}^{p}\sum_{t=1}^{n}\sum_{s=1}^{n} x_{is}a_{st}x_{kt}b_{k\ell}x_{\ell r}c_{rq}x_{qj}\right)$$

$$
\begin{aligned}
= &\sum_{\substack{q,r,\ell\\k,t,s}} \{k_0[\sigma_{ik}\phi_{st}\sigma_{\ell q}\phi_{rj} + \sigma_{i\ell}\phi_{sr}\sigma_{kq}\phi_{tj} \\
&+ \sigma_{iq}\phi_{sj}\sigma_{k\ell}\phi_{tr}] + c_0[m_{is}m_{kt}\sigma_{\ell q}\phi_{rj} \\
&+ m_{is}m_{\ell r}\sigma_{kq}\,\phi_{tj} + m_{is}m_{qj}\sigma_{k\ell}\phi_{tr} \\
&+ m_{kt}m_{\ell r}\sigma_{iq}\phi_{sj} + m_{kt}m_{qj}\sigma_{i\ell}\phi_{sr} + m_{\ell r}m_{qj}\sigma_{ik}\phi_{st}] \\
&+ m_{is}m_{kt}m_{\ell r}m_{qj}\}a_{st}b_{k\ell}c_{rq} \\
= &\sum_{\substack{q,r,\ell\\k,t,s}} k_0[(\sigma_{ik}b_{k\ell}\sigma_{\ell q}c_{rq}\phi_{rj})(\phi_{st}a_{st}) \\
&+ \sigma_{i\ell}b_{k\ell}\sigma_{kq}c_{rq}\phi_{sr}a_{st}\phi_{tj} \\
&+ (\sigma_{iq}c_{rq}\phi_{tr}a_{st}\phi_{sj})(\sigma_{k\ell}b_{k\ell})] \\
&+ c_0[m_{is}a_{st}m_{kt}b_{k\ell}\sigma_{\ell q}c_{rq}\phi_{rj} \\
&+ (m_{is}a_{st}\phi_{tj})(\sigma_{kq}c_{rq}m_{\ell r}b_{k\ell}) \\
&+ (m_{is}a_{st}\phi_{tr}c_{rq}m_{qj})(\sigma_{k\ell}b_{k\ell}) \\
&+ \sigma_{iq}c_{rq}m_{\ell r}b_{k\ell}m_{kt}a_{st}\phi_{sj} \\
&+ \sigma_{i\ell}b_{k\ell}m_{kt}a_{st}\phi_{sr}c_{rq}m_{qj} \\
&+ (\sigma_{ik}b_{k\ell}m_{\ell r}c_{rq}m_{qj})(\phi_{st}a_{st})]
\end{aligned}
$$

$$
\begin{aligned}
&+ m_{is}a_{st}m_{kt}b_{k\ell}m_{\ell r}c_{rq}m_{qj}\\
&= (k_0[\Sigma B\Sigma C'\Phi \operatorname{tr}(\Phi A) + \Sigma B'\Sigma C'\Phi A\Phi\\
&+ \Sigma C'\Phi A'\Phi \operatorname{tr}(\Sigma B)]\\
&+ c_0[MAM'B\Sigma C'\Phi + MA\Phi \operatorname{tr}(\Sigma C'M'B')\\
&+ MA\Phi CM \operatorname{tr}(\Sigma B) + \Sigma C'M'B'MA'\Phi\\
&+ \Sigma B'MA'\Phi CM + \Sigma BMCM \operatorname{tr}(\Phi A)]\\
&+ MAM'BMCM)_{ij}.
\end{aligned}
$$

vi) $$\mathcal{E}(X'AXBXCX')_{ij} = \mathcal{E}\left(\sum_{q=1}^{n}\sum_{r=1}^{n}\sum_{\ell=1}^{p}\sum_{k=1}^{n}\sum_{t=1}^{p}\sum_{s=1}^{p} x_{si}a_{st}x_{tk}b_{k\ell}x_{\ell r}c_{rq}x_{jq}\right)$$

$$
\begin{aligned}
&= \sum_{\substack{q,r,\ell\\k,t,s}} \{k_0[\sigma_{st}\phi_{ik}\sigma_{\ell j}\phi_{rq} + \sigma_{s\ell}\phi_{ir}\sigma_{tj}\phi_{kq}\\
&\quad + \sigma_{sj}\phi_{iq}\sigma_{t\ell}\phi_{kr}] + c_0[m_{si}m_{tk}\sigma_{\ell j}\phi_{rq}\\
&\quad + m_{si}m_{\ell r}\sigma_{tj}\phi_{kq} + m_{si}m_{jq}\sigma_{t\ell}\phi_{kr}\\
&\quad + m_{tk}m_{\ell r}\sigma_{sj}\phi_{iq} + m_{tk}m_{jq}\sigma_{s\ell}\phi_{ir}\\
&\quad + m_{\ell r}m_{jq}\sigma_{st}\phi_{ik}] + m_{si}m_{tk}m_{\ell r}m_{jq}\}a_{st}b_{k\ell}c_{rq}\\
&= \sum_{\substack{q,r,\ell\\k,t,s}} k_0[(\phi_{ik}b_{k\ell}\sigma_{\ell j})(\sigma_{st}a_{st})(\phi_{rq}c_{rq})\\
&\quad + \phi_{ir}c_{rq}\phi_{kq}b_{k\ell}\sigma_{s\ell}a_{st}\sigma_{tj}\\
&\quad + \phi_{iq}c_{rq}\phi_{kr}b_{k\ell}\sigma_{t\ell}a_{st}\sigma_{sj}]\\
&\quad + c_0[(m_{si}a_{st}m_{tk}b_{k\ell}\sigma_{\ell j})(\phi_{rq}c_{rq})\\
&\quad + (m_{si}a_{st}\sigma_{tj})(\phi_{kq}c_{rq}m_{\ell r}b_{k\ell})\\
&\quad + m_{si}a_{st}\sigma_{t\ell}b_{k\ell}\phi_{kr}c_{rq}m_{jq}\\
&\quad + \phi_{iq}c_{rq}m_{\ell r}b_{k\ell}m_{tk}a_{st}\sigma_{sj}\\
&\quad + (\phi_{ir}c_{rq}m_{jq})(\sigma_{s\ell}b_{k\ell}m_{tk}a_{st})\\
&\quad + (\phi_{ik}b_{k\ell}m_{\ell r}c_{rq}m_{jq})(\sigma_{st}a_{st})]\\
&\quad + m_{si}a_{st}m_{tk}b_{k\ell}m_{\ell r}c_{rq}m_{jq}\\
&= (k_0[\Phi B\Sigma \operatorname{tr}(\Sigma A) \operatorname{tr}(\Phi C) + \Phi C\Phi B\Sigma A\Sigma\\
&\quad + \Phi C'\Phi B\Sigma A'\Sigma] + c_0[M'AMB\Sigma \operatorname{tr}(\Phi C)\\
&\quad + M'A\Sigma \operatorname{tr}(\Phi C'M'B') + M'A\Sigma B'\Phi CM'\\
&\quad + \Phi C'M'B'M'A'\Sigma + \Phi CM' \operatorname{tr}(\Sigma B'M'A')\\
&\quad + \Phi BMCM' \operatorname{tr}(\Sigma A)] + M'AMBMCM')_{ij}. \blacksquare
\end{aligned}
$$

THEOREM 3.2.12. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ *with finite fourth order moment. Let* $c_0 = -2\psi'(0)$ *and* $k_0 = 4\psi''(0)$. *Then, for any constant matrices* A, B, *and* C, *if the expressions on the left-hand sides are defined, we have*

i) $$\mathcal{E}(XA\ \mathrm{tr}(XBXCX')) = k_0(\Sigma^2B'\Phi C\Phi A + \Sigma^2B'\Phi A\ \mathrm{tr}(C\Phi) + \Sigma B'\Phi C'\Phi A\ \mathrm{tr}(\Sigma)) + c_0(MA\ \mathrm{tr}(MB\Sigma)\ \mathrm{tr}(C\Phi) + MA\ \mathrm{tr}(MC\Phi B)\ \mathrm{tr}(\Sigma) + \Sigma MBMC\Phi A + MA\ \mathrm{tr}(MC'\Phi B\Sigma) + \Sigma B'M'MC'\Phi A + \Sigma MC'M'B\Phi A) + MA\ \mathrm{tr}(MBMCM'),$$

ii) $$\mathcal{E}(XAX\ \mathrm{tr}(BXCX')) = k_0(\Sigma B\Sigma A'\Phi C'\Phi + \Sigma A'\Phi\ \mathrm{tr}(B\Sigma)\ \mathrm{tr}(C\Phi) + \Sigma B'\Sigma A'\Phi C\Phi) + c_0(MAM\ \mathrm{tr}(B\Sigma)\ \mathrm{tr}(C\Phi) + MA\Sigma BMC\Phi + \Sigma BMC\Phi AM + MA\Sigma B'MC'\Phi\ ] + \Sigma B'MC'\Phi AM + \Sigma A'\Phi\ \mathrm{tr}(MC'M'B')) + MAM\ \mathrm{tr}(BMCM'),$$

iii) $$\mathcal{E}(X'AX\ \mathrm{tr}(BXCX')) = k_0(\Phi C'\Phi\ \mathrm{tr}(\Sigma B\Sigma A') + \Phi\ \mathrm{tr}(A\Sigma)\ \mathrm{tr}(B\Sigma)\ \mathrm{tr}(C\Phi) + \Phi C\Phi\ \mathrm{tr}(\Sigma B'\Sigma A')) + c_0(M'AM\ \mathrm{tr}(B\Sigma)\ \mathrm{tr}(C\Phi) + M'A\Sigma BMC\Phi + \Phi C'M'B'\Sigma AM + M'A\Sigma B'MC'\Phi + \Phi CM'B\Sigma AM + \Phi\ \mathrm{tr}(A\Sigma)\ \mathrm{tr}(MCM'B)) + M'AM\ \mathrm{tr}(BMCM'),$$

iv) $$\mathcal{E}(XBXCX\ \mathrm{tr}(AX)) = k_0(\Sigma B'\Phi C\Sigma A'\Phi + \Sigma A'\Phi B\Sigma C'\Phi + \Sigma C'\Phi A\Sigma B'\Phi) + c_0(MB\Sigma C'\Phi\ \mathrm{tr}(AM) + \Sigma C'M'B'\Phi\ \mathrm{tr}(AM) + MBMC\Sigma A'\Phi + \Sigma B'\Phi CM\ \mathrm{tr}(MA) + MB\Sigma A'\Phi CM + \Sigma A'\Phi BMCM) + MBMCM\ \mathrm{tr}(AM),$$

v) $$\mathcal{E}(X'BXCX\ \mathrm{tr}(AX)) = k_0(\Phi C\Sigma A'\Phi\ \mathrm{tr}(B\Sigma) + \Phi A\Sigma B\Sigma C'\Phi + \Phi\ \mathrm{tr}(\Sigma B'\Sigma C'\Phi A)) + c_0(M'B\Sigma C'\Phi\ \mathrm{tr}(MA) + \Phi\ \mathrm{tr}(MA)\ \mathrm{tr}(MC\Sigma B) + M'BMC\Sigma A'\Phi + \Phi CM\ \mathrm{tr}(B\Sigma)\ \mathrm{tr}(MA) + M'B\Sigma A'\Phi CM + \Phi A\Sigma BMCM) + M'BMCM\ \mathrm{tr}(MA),$$

vi) $$\mathcal{E}(X'BX'CX\ \mathrm{tr}(AX)) = k_0(\Phi B'\Sigma C\Sigma A'\Phi + \Phi A\Sigma B\Phi\ \mathrm{tr}(\Sigma C) + \Phi\ \mathrm{tr}(\Sigma C\Sigma B\Phi A)) + c_0(M'B\Phi\ \mathrm{tr}(MA)\ \mathrm{tr}(\Sigma C)$$

$$+ \Phi \operatorname{tr}(MA)\operatorname{tr}(MB'\Sigma C') + M'BM'C\Sigma A'\Phi$$
$$+ \Phi B'\Sigma CM \operatorname{tr}(MA) + M'B\Phi A\Sigma CM$$
$$+ \Phi A\Sigma BM'CM) + M'BM'CM \operatorname{tr}(AM).$$

PROOF.

i) $\mathcal{E}(XA\operatorname{tr}(XBXCX'))_{ij}$

$$= \mathcal{E}\left(\left(\sum_{s=1}^{n} x_{is}a_{sj}\right)\left(\sum_{q=1}^{n}\sum_{r=1}^{n}\sum_{\ell=1}^{p}\sum_{k=1}^{n}\sum_{t=1}^{p} x_{tk}b_{k\ell}x_{\ell r}c_{rq}x_{tq}\right)\right)$$

$$= \mathcal{E}\left(\sum_{\substack{q,r,\ell\\k,t,s}} x_{is}a_{sj}x_{tk}b_{k\ell}x_{\ell r}c_{rq}x_{tq}\right)$$

$$= \sum_{\substack{q,r,\ell\\k,t,s}} \{k_0[\sigma_{it}\phi_{sk}\sigma_{\ell t}\phi_{rq} + \sigma_{i\ell}\phi_{sr}\sigma_{tt}\phi_{kq} + \sigma_{it}\phi_{sq}\sigma_{t\ell}\phi_{kr}]$$
$$+ c_0[m_{is}m_{tk}\sigma_{\ell t}\phi_{rq} + m_{is}m_{\ell r}\sigma_{tt}\phi_{kq}$$
$$+ m_{is}m_{tq}\sigma_{t\ell}\phi_{kr} + m_{tk}m_{\ell r}\sigma_{it}\phi_{sq} + m_{tk}m_{tq}\sigma_{i\ell}\phi_{sr}]$$
$$+ m_{is}m_{tk}m_{\ell r}m_{tq}\}a_{sj}b_{k\ell}c_{rq}$$

$$= \sum_{\substack{q,r,\ell\\k,t,s}} k_0[(\sigma_{it}\sigma_{t\ell}b_{k\ell}\phi_{sk}a_{sj})(\phi_{rq}c_{rq})$$
$$+ (\sigma_{i\ell}b_{k\ell}\phi_{kq}c_{rq}\phi_{sr}a_{sj})\sigma_{tt}$$
$$+ \sigma_{it}\sigma_{t\ell}b_{k\ell}\phi_{kr}c_{rq}\phi_{sq}a_{sj}]$$
$$+ c_0[(m_{is}a_{sj})(m_{tk}b_{k\ell}\sigma_{\ell t})(\phi_{rq}c_{rq})$$
$$+ (m_{is}a_{is})\sigma_{tt}(\phi_{kq}c_{rq}m_{\ell r}b_{k\ell})$$
$$+ (m_{is}a_{sj})(\sigma_{t\ell}b_{k\ell}\phi_{kr}c_{rq}m_{tq})$$
$$+ \sigma_{it}m_{tk}b_{k\ell}m_{\ell r}c_{rq}\phi_{sq}a_{sj} + \sigma_{i\ell}b_{k\ell}m_{tk}m_{tq}c_{rq}\phi_{sr}a_{sj}$$
$$+ \sigma_{it}m_{tq}c_{rq}m_{\ell r}b_{k\ell}\phi_{sk}a_{sj}]$$
$$+ (m_{is}a_{sj})(m_{tk}b_{k\ell}m_{\ell r}c_{rq}m_{tq})$$
$$= (k_0[\Sigma^2B'\Phi A \operatorname{tr}(\Phi C) + \Sigma B'\Phi C'\Phi A \operatorname{tr}(\Sigma)$$
$$+ \Sigma^2B'\Phi C\Phi A)] + c_0[MA \operatorname{tr}(MB\Sigma)\operatorname{tr}(\Phi C)$$
$$+ MA \operatorname{tr}(\Sigma)\operatorname{tr}(\Phi C'M'B') + MA \operatorname{tr}(\Sigma B'\Phi CM')$$
$$+ \Sigma MBMC\Phi A + \Sigma B'M'MC'\Phi A + \Sigma MC'M'B'\Phi A]$$
$$+ MA \operatorname{tr}(MBMCM'))ij.$$

ii) $\mathcal{E}(XAX\operatorname{tr}(BXCX'))_{ij}$

$$= \mathcal{E}\left(\left(\sum_{t=1}^{p}\sum_{s=1}^{n} x_{is}a_{st}x_{tj}\right)\left(\sum_{q=1}^{n}\sum_{r=1}^{n}\sum_{\ell=1}^{p}\sum_{k=1}^{p} b_{k\ell}x_{\ell r}c_{rq}x_{kq}\right)\right)$$

$$= \mathcal{E}\left(\sum_{\substack{q,r,\ell\\k,t,s}} x_{is}a_{st}x_{tj}b_{k\ell}x_{\ell r}c_{rq}x_{kq}\right)$$

$$\begin{aligned}
= \sum_{\substack{q,r,\ell\\k,t,s}} \{&k_0[\sigma_{it}\phi_{sj}\sigma_{\ell k}\phi_{rq} + \sigma_{i\ell}\phi_{sr}\sigma_{tk}\phi_{jq} + \sigma_{ik}\phi_{sq}\sigma_{t\ell}\phi_{jr}]\\
&+ c_0[m_{is}m_{tj}\sigma_{\ell k}\phi_{rq} + m_{is}m_{\ell r}\sigma_{tk}\phi_{jq} + m_{is}m_{kq}\sigma_{t\ell}\phi_{jr}\\
&+ m_{tj}m_{\ell r}\sigma_{is}\phi_{sq} + m_{tj}m_{kq}\sigma_{i\ell}\phi_{sr} + m_{\ell r}m_{kq}\sigma_{it}\phi_{sj}]\\
&+ m_{is}m_{tj}m_{\ell r}m_{kq}\}a_{st}b_{k\ell}c_{rq}
\end{aligned}$$

$$\begin{aligned}
= \sum_{\substack{q,r,\ell\\k,t,s}} &k_0[(\sigma_{it}a_{st}\phi_{sj})(\sigma_{\ell k}b_{k\ell})(\phi_{rq}c_{rq}) + \sigma_{i\ell}b_{k\ell}\sigma_{tk}a_{st}\phi_{sr}c_{rq}\phi_{jq}\\
&+ \sigma_{ik}b_{k\ell}\sigma_{t\ell}a_{st}\phi_{sq}c_{rq}\phi_{jr}] + c_0[(m_{is}a_{st}m_{tj})(\sigma_{\ell k}b_{k\ell})(\phi_{rq}c_{rq})\\
&+ m_{is}a_{st}\sigma_{tk}b_{k\ell}m_{\ell r}c_{rq}\phi_{jq} + m_{is}a_{st}\sigma_{t\ell}b_{k\ell}m_{kq}c_{rq}\phi_{jr}\\
&+ \sigma_{ik}b_{k\ell}m_{\ell r}c_{rq}\phi_{sq}a_{st}m_{tj} + \sigma_{i\ell}b_{k\ell}m_{kq}c_{rq}\phi_{sr}a_{st}m_{tj}\\
&+ (\sigma_{it}a_{st}\phi_{sj})(m_{kq}c_{rq}m_{\ell r}b_{k\ell})] + (m_{is}a_{st}m_{tj})(m_{kq}c_{rq}m_{\ell r}b_{k\ell})
\end{aligned}$$

$$\begin{aligned}
= (&k_0[\Sigma A'\Phi\ \mathrm{tr}(\Sigma B)\ \mathrm{tr}(\Phi C) + \Sigma B'\Sigma A'\Phi C\Phi + \Sigma B\Sigma A'\Phi C'\Phi]\\
&+ c_0[MAM\ \mathrm{tr}(\Sigma B)\ \mathrm{tr}(\Phi C) + MA\Sigma BMC\Phi + MA\Sigma B'MC'\Phi\\
&+ \Sigma BMC\Phi AM + \Sigma B'MC'\Phi AM + \Sigma A'\Phi\ \mathrm{tr}(MC'M'B')]\\
&+ MAM\ \mathrm{tr}(MC'M'B'))_{ij}.
\end{aligned}$$

iii) $\mathcal{E}(X'AX\ \mathrm{tr}(BXCX'))_{ij}$

$$= \mathcal{E}\left(\left(\sum_{t=1}^{p}\sum_{s=1}^{p} x_{si}a_{st}x_{tj}\right)\left(\sum_{q=1}^{n}\sum_{r=1}^{n}\sum_{\ell=1}^{p}\sum_{k=1}^{p} b_{k\ell}x_{\ell r}c_{rq}x_{kq}\right)\right)$$

$$= \mathcal{E}\left(\sum_{\substack{q,r,\ell\\k,t,s}} x_{si}a_{st}x_{tj}b_{k\ell}x_{\ell r}c_{rq}x_{kq}\right)$$

$$= \sum_{\substack{q,r,\ell\\k,t,s}} \{k_0[\sigma_{st}\phi_{ij}\sigma_{\ell k}\phi_{rq} + \sigma_{s\ell}\phi_{ir}\sigma_{tk}\phi_{jq} + \sigma_{sk}\phi_{iq}\sigma_{t\ell}\phi_{jr}]$$

$$+ c_0[m_{si}m_{tj}\sigma_{\ell k}\phi_{rq} + m_{si}m_{\ell r}\sigma_{tk}\phi_{jq} + m_{si}m_{kq}\sigma_{t\ell}\phi_{jr}$$
$$+ m_{tj}m_{\ell r}\sigma_{sk}\phi_{iq} + m_{tj}m_{kq}\sigma_{s\ell}\phi_{ir} + m_{\ell r}m_{kq}\sigma_{st}\phi_{ij}]$$
$$+ m_{si}m_{tj}m_{\ell r}m_{kq}\}a_{st}b_{k\ell}c_{rq}$$

$$= \sum_{\substack{q,r,\ell\\k,t,s}} k_0[(\phi_{ij})(\sigma_{st}a_{st})(\sigma_{\ell k}b_{k\ell})(\phi_{rq}c_{rq}) + (\phi_{ir}c_{rq}\phi_{jq})(\sigma_{s\ell}b_{k\ell}\sigma_{tk}a_{st})$$
$$+ (\phi_{iq}c_{rq}\phi_{jr})(\sigma_{sk}b_{k\ell}\sigma_{t\ell}a_{st})] + c_0[(m_{si}a_{st}m_{tj})(\sigma_{\ell k}b_{k\ell})(\phi_{rq}c_{rq})$$
$$+ m_{si}a_{st}\sigma_{tk}b_{k\ell}m_{\ell r}c_{rq}\phi_{jq} + m_{si}a_{st}\sigma_{t\ell}b_{k\ell}m_{kq}c_{rq}\phi_{jr}$$
$$+ \phi_{iq}c_{rq}m_{\ell r}b_{k\ell}\sigma_{sk}a_{st}m_{tj} + \phi_{ir}c_{rq}m_{kq}b_{k\ell}\sigma_{s\ell}a_{st}m_{tj}$$
$$+ (\phi_{ij})(\sigma_{st}a_{st})(m_{kq}c_{rq}m_{\ell r}b_{k\ell})] + (m_{si}a_{st}m_{tj})(m_{kq}c_{rq}m_{\ell r}b_{k\ell})$$
$$= (k_0[\Phi \operatorname{tr}(\Sigma A) \operatorname{tr}(\Sigma B) \operatorname{tr}(\Phi C) + \Phi C\Phi \operatorname{tr}(\Sigma B'\Sigma A')$$
$$+ \Phi C'\Phi \operatorname{tr}(\Sigma B\Sigma A')] + c_0[M'AM \operatorname{tr}(\Sigma B) \operatorname{tr}(\Phi C)$$
$$+ M'A\Sigma BMC\Phi + M'A\Sigma B'MC'\Phi + \Phi C'M'B'\Sigma AM$$
$$+ \Phi CM'B\Sigma AM + \Phi \operatorname{tr}(\Sigma A) \operatorname{tr}(MC'M'B')]$$
$$+ M'AM \operatorname{tr}(MC'M'B'))_{ij}.$$

iv) $\mathcal{E}(XBXCX \operatorname{tr}(AX))_{ij}$

$$= \mathcal{E}\left(\left(\sum_{t=1}^{p}\sum_{s=1}^{n} x_{ts}a_{st}\right)\left(\sum_{q=1}^{p}\sum_{r=1}^{n}\sum_{\ell=1}^{p}\sum_{k=1}^{n} x_{ik}b_{k\ell}x_{\ell r}c_{rq}x_{qj}\right)\right)$$
$$= \mathcal{E}\left(\sum_{\substack{q,r,\ell\\k,t,s}} x_{ts}a_{st}x_{ik}b_{k\ell}c_{rq}x_{qj}\right)$$
$$= \sum_{\substack{q,r,\ell\\k,t,s}} \{k_0[\sigma_{ti}\phi_{sk}\sigma_{\ell q}\phi_{rj} + \sigma_{t\ell}\phi_{sr}\sigma_{iq}\phi_{kj} + \sigma_{tq}\phi_{sj}\sigma_{i\ell}\phi_{kr}]$$
$$+ c_0[m_{ts}m_{ik}\sigma_{\ell q}\phi_{rj} + m_{ts}m_{\ell r}\sigma_{iq}\phi_{kj} + m_{ts}m_{qj}\sigma_{i\ell}\phi_{kr}$$
$$+ m_{ik}m_{\ell r}\sigma_{tq}\phi_{sj} + m_{ik}m_{qj}\sigma_{t\ell}\phi_{sr} + m_{\ell r}m_{qj}\sigma_{ti}\phi_{sk}]$$
$$+ m_{ts}m_{ik}m_{\ell r}m_{qj}\}a_{st}b_{k\ell}c_{rq}$$

$$= \sum_{\substack{q,r,\ell \\ k,t,s}} k_0[\sigma_{ti}a_{st}\phi_{sk}b_{k\ell}\sigma_{\ell q}c_{rq} + \sigma_{iq}c_{rq}\phi_{sr}a_{st}\sigma_{t\ell}b_{k\ell}\phi_{kj}$$
$$+ \sigma_{i\ell}b_{k\ell}\phi_{kr}c_{rq}\sigma_{tq}a_{st}\phi_{sj}] + c_0[(m_{ik}b_{k\ell}\sigma_{\ell q}c_{rq}\phi_{rj})(m_{ts}a_{st})$$
$$+ (\sigma_{iq}c_{rq}m_{\ell r}b_{k\ell}\phi_{kj})(m_{ts}a_{st}) + (\sigma_{i\ell}b_{k\ell}\phi_{kr}c_{rq}m_{qj})(m_{ts}a_{st})$$
$$+ m_{ik}b_{k\ell}m_{\ell r}c_{rq}\sigma_{tq}a_{st}\phi_{sj} + m_{ik}b_{k\ell}\sigma_{t\ell}a_{st}\phi_{sr}c_{rq}m_{qj}$$
$$+ \sigma_{ti}a_{st}\phi_{sk}b_{k\ell}m_{\ell r}c_{rq}m_{qj}] + (m_{ik}b_{k\ell}m_{\ell r}c_{rq}m_{qj})(m_{ts}a_{st})$$
$$= (k_0[\Sigma A'\Phi B\Sigma C'\Phi + \Sigma C'\Phi A\Sigma B'\Phi + \Sigma B'\Phi C\Sigma A'\Phi]$$
$$+ c_0[MB\Sigma C'\Phi \text{ tr}(MA) + \Sigma C'M'B'\Phi \text{ tr}(MA)$$
$$+ \Sigma B'\Phi CM \text{ tr}(MA) + MBMC\Sigma A'\Phi + MB\Sigma A'\Phi CM$$
$$+ \Sigma A'\Phi MCM] + MBMCM \text{ tr}(MA))_{ij}.$$

v) $\mathcal{E}(X'BXCX\text{tr}(AX))_{ij}$

$$= \mathcal{E}\left(\left(\sum_{t=1}^{p}\sum_{s=1}^{n} x_{ts}a_{st}\right)\left(\sum_{q=1}^{p}\sum_{r=1}^{n}\sum_{\ell=1}^{p}\sum_{k=1}^{p} x_{ki}b_{k\ell}x_{\ell r}c_{rq}x_{qj}\right)\right)$$
$$= \mathcal{E}\left(\sum_{\substack{q,r,\ell \\ k,t,s}} x_{ts}a_{st}x_{ki}b_{k\ell}x_{\ell r}c_{rq}x_{qj}\right)$$

$$= \sum_{\substack{q,r,\ell \\ k,t,s}} \{k_0[\sigma_{tk}\phi_{si}\sigma_{\ell q}\phi_{rj} + \sigma_{t\ell}\phi_{sr}\sigma_{kq}\phi_{ij} + \sigma_{tq}\phi_{sj}\sigma_{k\ell}\phi_{ir}]$$
$$+ c_0[m_{ts}m_{ki}\sigma_{\ell q}\phi_{rj} + m_{ts}m_{\ell r}\sigma_{kq}\phi_{ij} + m_{ts}m_{qj}\sigma_{k\ell}\phi_{ir}$$
$$+ m_{ki}m_{\ell r}\sigma_{tq}\phi_{sj} + m_{ki}m_{qj}\sigma_{t\ell}\phi_{sr} + m_{\ell r}m_{qj}\sigma_{tk}\phi_{si}]$$
$$+ m_{ts}m_{ki}m_{\ell r}m_{qj}\}a_{st}b_{k\ell}c_{rq}$$

$$= \sum_{\substack{q,r,\ell \\ k,t,s}} k_0[\phi_{si}a_{st}\sigma_{tk}b_{k\ell}\sigma_{\ell q}c_{rq}\phi_{rj} + (\phi_{ij})(\phi_{sr}c_{rq}\sigma_{kq}b_{k\ell}\sigma_{t\ell}a_{st})$$
$$+ (\phi_{ir}c_{rq}\sigma_{tq}a_{st}\phi_{sj})(\sigma_{k\ell}b_{k\ell})] + c_0[(m_{ki}b_{k\ell}\sigma_{\ell q}c_{rq}\phi_{rj}(m_{ts}a_{st})$$
$$+ (\phi_{ij})(m_{ts}a_{st})(\sigma_{kq}c_{rq}m_{\ell r}b_{k\ell}) + (\phi_{ir}c_{rq}m_{qj})(m_{ts}a_{st})(\sigma_{k\ell}b_{k\ell})$$
$$+ m_{ki}b_{k\ell}m_{\ell r}c_{rq}\sigma_{tq}a_{st}\phi_{sj} + m_{ki}b_{k\ell}\sigma_{t\ell}a_{st}\phi_{sr}c_{rq}m_{qj}$$
$$+ \phi_{si}a_{st}\sigma_{tk}b_{k\ell}m_{\ell r}c_{rq}m_{qj}] + (m_{ki}b_{k\ell}m_{\ell r}c_{rq}m_{qj})(m_{ts}a_{st})$$
$$= (k_0[\Phi A\Sigma B\Sigma C'\Phi + \Phi \text{ tr}(\Phi C\Sigma B\Sigma A') + \Phi C\Sigma A'\Phi \text{ tr}(\Sigma B)]$$
$$+ c_0[M'B\Sigma C'\Phi \text{ tr}(MA) + \Phi \text{ tr}(MA) \text{ tr}(\Sigma C'M'B')$$

$$
\begin{aligned}
&+ \Phi CM \operatorname{tr}(MA) \operatorname{tr}(\Sigma B') + M'BMC\Sigma A'\Phi + M'B\Sigma A'\Phi CM \\
&+ \Phi A\Sigma BMCM] + M'BMCM \operatorname{tr}(MA))_{ij}.
\end{aligned}
$$

vi) $\mathcal{E}(X'BX'CX \operatorname{tr}(AX))_{ij}$

$$
\begin{aligned}
&= \mathcal{E}\left(\left(\sum_{t=1}^{p}\sum_{s=1}^{n} x_{ts}a_{st}\right)\left(\sum_{q=1}^{p}\sum_{r=1}^{p}\sum_{\ell=1}^{n}\sum_{k=1}^{p} x_{ki}b_{k\ell}x_{r\ell}c_{rq}x_{qj}\right)\right) \\
&= \mathcal{E}\left(\sum_{\substack{q,r,\ell\\k,t,s}} x_{ts}a_{st}x_{ki}b_{k\ell}x_{r\ell}c_{rq}x_{qj}\right) \\
&= \sum_{\substack{q,r,\ell\\k,t,s}} \{k_0[\sigma_{tk}\phi_{si}\sigma_{rq}\phi_{\ell j} + \sigma_{tr}\phi_{s\ell}\sigma_{kq}\phi_{ij} + \sigma_{tq}\phi_{sj}\sigma_{kr}\phi_{i\ell}] \\
&\quad + c_0[m_{ts}m_{ki}\sigma_{rq}\phi_{\ell j} + m_{ts}m_{r\ell}\sigma_{kq}\phi_{ij} \\
&\quad + m_{ts}m_{qj}\sigma_{kr}\phi_{i\ell} + m_{ki}m_{r\ell}\sigma_{tq}\phi_{sj} + m_{ki}m_{qj}\sigma_{tr}\phi_{s\ell} \\
&\quad + m_{r\ell}m_{qj}\sigma_{tk}\phi_{si}] + m_{ts}m_{ki}m_{r\ell}m_{qj}\}a_{st}b_{k\ell}c_{rq} \\
&= \sum_{\substack{q,r,\ell\\k,t,s}} k_0[(\phi_{si}a_{st}\sigma_{tk}b_{k\ell}\phi_{\ell j})(\sigma_{rq}c_{rq}) \\
&\quad + (\phi_{ij})(\phi_{s\ell}b_{k\ell}\sigma_{kq}c_{rq}\sigma_{tr}a_{st}) + \phi_{i\ell}b_{k\ell}\sigma_{kr}c_{rq}\sigma_{tq}a_{st}\phi_{sj}] \\
&\quad + c_0[(m_{ki}b_{k\ell}\phi_{\ell j})(m_{ts}a_{st})(\sigma_{rq}c_{rq}) + (\phi_{ij})(m_{ts}a_{st})(\sigma_{kq}c_{rq}m_{r\ell}b_{k\ell}) \\
&\quad + (\phi_{i\ell}b_{k\ell}\sigma_{kr}c_{rq}m_{qj})(m_{ts}a_{st}) + m_{ki}b_{k\ell}m_{r\ell}c_{rq}\sigma_{tq}a_{st}\phi_{sj} \\
&\quad + m_{ki}b_{k\ell}\phi_{s\ell}a_{st}\sigma_{tr}c_{rq}m_{qj} + \phi_{si}a_{st}\sigma_{tk}b_{k\ell}m_{r\ell}c_{rq}m_{qj}] \\
&\quad + (m_{ki}b_{k\ell}m_{r\ell}c_{rq}m_{qj})(m_{ts}a_{st}) \\
&= (k_0[\Phi A\Sigma B\Phi \operatorname{tr}(\Sigma C) + \Phi \operatorname{tr}(\Phi B'\Sigma C'\Sigma A') + \Phi B'\Sigma C\Sigma A'\Phi] \\
&\quad + c_0[M'B\Phi \operatorname{tr}(MA) \operatorname{tr}(\Sigma C) + \Phi \operatorname{tr}(MA) \operatorname{tr}(\Sigma C'MB') \\
&\quad + \Phi B'\Sigma CM \operatorname{tr}(MA) + M'BM'C\Sigma A'\Phi + M'B\Phi A\Sigma CM \\
&\quad + \Phi A\Sigma BM'CM] + M'BM'CM \operatorname{tr}(MA))_{ij}. \blacksquare
\end{aligned}
$$

REMARK 3.2.2. The expected values of many of the expressions in Theorems 3.2.8-3.2.13 were computed by Nel (1977) for the case where X has matrix variate normal distribution. If $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$, then $-2\psi'(0) = 1$ and $4\psi''(0) = 1$. Therefore taking $c_0 = k_0 = 1$, our results give the expected values for the normal case, and so Nel's results can be obtained as special cases of the formulae presented here.

Next, we give some applications of Theorems 3.2.8-3.2.13.

THEOREM 3.2.13. *Let* $X \sim E_{p,n}(O, \Sigma \otimes I_n, \psi)$ *with finite fourth order moments. Let* $Y = X\left(I_n - \frac{e_n e_n'}{n}\right)X'$. *Let* $c_0 = -2\psi'(0)$ *and* $k_0 = 4\psi''(0)$. *Then,*

i) $$\mathcal{E}(Y) = c_0(n-1)\Sigma, \tag{3.36}$$

ii) $$\begin{aligned}\mathrm{Cov}(y_{ij}, y_{k\ell}) &= k_0(n-1)(\sigma_{i\ell}\sigma_{jk} + \sigma_{ik}\sigma_{j\ell}) \\ &\quad + (n-1)^2(k_0 - c_0^2)\sigma_{ij}\sigma_{k\ell},\end{aligned} \tag{3.37}$$

iii) $$\mathrm{Var}(y_{ij}) = (nk_0 - (n-1)c_0^2)(n-1)\sigma_{ij}^2 + k_0(n-1)\sigma_{ii}\sigma_{jj}. \tag{3.38}$$

PROOF. Let $A = I_n - \frac{e_n e_n'}{n}$.

i) Using Theorem 3.2.8, we get

$$\mathcal{E}(XAX') = c_0\Sigma\,\mathrm{tr}\left(\left(I_n - \frac{e_n e_n'}{n}\right)I_n\right) + O = c_0(n-1)\Sigma.$$

ii) Let $\mathbf{q}^m$ be a p-dimensional column vector such that

$$q_i^m = \begin{cases} 1 & \text{if } i = m \\ 0 & \text{if } i \neq m \end{cases}, \quad m = 1,2,\ldots,p;\ i = 1,2,\ldots,p.$$

Then, $y_{ij} = \mathbf{q}^{i\prime}Y\mathbf{q}^j = \mathbf{q}^{i\prime}XAX'\mathbf{q}^j$ and $y_{k\ell} = \mathbf{q}^{k\prime}Y\mathbf{q}^\ell = \mathbf{q}^{k\prime}XAX'\mathbf{q}^\ell$. Now from (3.36), it follows that

$$\mathcal{E}(y_{ij}) = c_0(n-1)\sigma_{ij} \text{ and } \mathcal{E}(y_{k\ell}) = c_0(n-1)\sigma_{k\ell}.$$

Since, $X \sim E_{p,n}(O, \Sigma \otimes I_n, \psi)$, we have $X' \sim E_{n,p}(O, I_n \otimes \Sigma, \psi)$. Using Theorem 3.2.12, we get

$$
\begin{aligned}
\mathcal{E}(y_{ij}y_{k\ell}) &= \mathcal{E}(q^{i\prime}XAX'q^jq^{k\prime}XAX'q^{\ell}) \\
&= q^{i\prime}\mathcal{E}(XAX'q^jq^{k\prime}XAX')q^{\ell} \\
&= q^{i\prime}(k_0(\Sigma\ \mathrm{tr}(I_nAI_nA)\ \mathrm{tr}(q^jq^{k\prime}\Sigma) \\
&\quad + \Sigma q^jq^{k\prime}\Sigma\ \mathrm{tr}(AI_n)\ \mathrm{tr}(AI_n) + \Sigma(q^jq^{k\prime})'\Sigma\ \mathrm{tr}(AI_nAI_n))q^{\ell} \\
&= k_0((n-1)q^{i\prime}\Sigma q^{\ell}\ \mathrm{tr}(q^{k\prime}\Sigma q^j) + q^{i\prime}\Sigma q^jq^{k\prime}\Sigma q^{\ell}(n-1)^2 \\
&\quad + q^{i\prime}\Sigma q^kq^{j\prime}\Sigma q^{\ell}(n-1)) \\
&= k_0((n-1)\sigma_{i\ell}\sigma_{kj} + (n-1)^2\sigma_{ij}\sigma_{k\ell} + (n-1)\sigma_{ik}\sigma_{j\ell}).
\end{aligned}
$$

Then,

$$
\begin{aligned}
\mathrm{Cov}(y_{ij}, y_{k\ell}) &= \mathcal{E}(y_{ij}y_{k\ell}) - \mathcal{E}(y_{ij})\mathcal{E}(y_{k\ell}) \\
&= k_0((n-1)\sigma_{i\ell}\sigma_{kj} + (n-1)^2\sigma_{ij}\sigma_{k\ell} + (n-1)\sigma_{ik}\sigma_{j\ell}) \\
&\quad - c_0^2(n-1)^2\sigma_{ij}\sigma_{k\ell} \\
&= k_0(n-1)(\sigma_{i\ell}\sigma_{jk} + \sigma_{ik}\sigma_{j\ell}) + (n-1)^2\ (k_0 - c_0^2)\sigma_{ij}\sigma_{k\ell}
\end{aligned}
$$

which proves (3.37).

iii) Take $k = i$ and $\ell = j$ in (3.37). Then,

$$
\begin{aligned}
\mathrm{Var}(y_{ij}) &= k_0(n-1)(\sigma_{ij}^2 + \sigma_{ii}\sigma_{jj}) + (n-1)^2\ (k_0 - c_0^2)\sigma_{ij}^2 \\
&= (n-1)(nk_0 - (n-1)c_0^2)\sigma_{ij}^2 + k_0(n-1)\sigma_{ii}\sigma_{jj}. \blacksquare
\end{aligned}
$$

EXAMPLE 3.2.1. Let $X \sim Mt_{p,n}(m, O, \Sigma \otimes \Phi)$, $m > 4$, and $Y = X\left(I_n - \frac{e_ne_n'}{n}\right)X'$. We want to find $\mathcal{E}(Y)$, $\mathrm{Cov}(y_{ij}, y_{k\ell})$ and $\mathrm{Var}(y_{ij})$. In order to compute them, we need to know $c_0 = -2\psi'(0)$ and $k_0 = 4\psi''(0)$. Let $u = \frac{x_{11}}{\sqrt{\sigma_{11}}}$. Then, $u \sim Mt_{1,1}(m,0,1)$; that is u has a one-dimensional Student's t-distribution with m degrees of freedom. (This will be shown in Chapter 4.) Hence, using Theorem 3.2.6 we get $\mathcal{E}(u^2) = c_0$ and $\mathcal{E}(u^4) = 3k_0$. It is known that

$$\mathcal{E}(u^r) = \frac{m^{\frac{r}{2}}\,\beta\left(\frac{r+1}{2}, \frac{m-r}{2}\right)}{\beta\left(\frac{1}{2}, \frac{m}{2}\right)} \quad \text{for } m > r \text{ and } r \text{ even,}$$

(see Mood, Graybill, and Boes, 1974, p. 543). In particular, we have

$$\begin{aligned}
\mathcal{E}(u^2) &= \frac{m\,\beta\left(\frac{2+1}{2}, \frac{m-2}{2}\right)}{\beta\left(\frac{1}{2}, \frac{m}{2}\right)} \\
&= m\,\frac{\frac{1}{2}\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{m-2}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\left(\frac{m-2}{2}\right)\Gamma\left(\frac{m-2}{2}\right)} \\
&= \frac{m}{m-2},
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{E}(u^4) &= \frac{m^2\,\beta\left(\frac{4+1}{2}, \frac{m-4}{2}\right)}{\beta\left(\frac{1}{2}, \frac{m}{2}\right)} \\
&= m^2\,\frac{\frac{3}{2}\frac{1}{2}\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{m-4}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\left(\frac{m-2}{2}\right)\left(\frac{m-4}{2}\right)\Gamma\left(\frac{m-4}{2}\right)} \\
&= \frac{3m^2}{(m-2)(m-4)}.
\end{aligned}$$

Hence, $c_0 = \frac{m}{m-2}$ and $k_0 = \frac{m^2}{(m-2)(m-4)}$. Now, using Theorem 3.2.14, we get

$$\begin{aligned}
&\mathcal{E}(Y) = (n-1)\frac{m}{m-2}\Sigma, \\
&\mathrm{Cov}(y_{ij}, y_{k\ell}) = (n-1)\left[\frac{m^2}{(m-2)(m-4)}(\sigma_{i\ell}\sigma_{jk} + \sigma_{ik}\sigma_{j\ell}) \right. \\
&\qquad\qquad \left. + (n-1)\left(\frac{m^2}{(m-2)(m-4)} - \frac{m^2}{(m-2)^2}\right)\sigma_{ij}\sigma_{k\ell}\right]
\end{aligned}$$

$$
\begin{aligned}
&= (n-1)\frac{m^2}{(m-2)(m-4)} \\
&\quad \cdot \left[\sigma_{i\ell}\sigma_{jk} + \sigma_{ik}\sigma_{j\ell} + (n-1)\left(1 - \frac{m-4}{m-2}\right)\sigma_{ij}\sigma_{k\ell}\right] \\
&= (n-1)\frac{m^2}{(m-2)(m-4)} \\
&\quad \cdot \left[\sigma_{i\ell}\sigma_{jk} + \sigma_{ik}\sigma_{j\ell} + \frac{2(n-1)}{m-2}\sigma_{ij}\sigma_{k\ell}\right],
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{Var}(y_{ij}) &= (n-1)\frac{m^2}{(m-2)(m-4)}\left[\sigma_{ij}^2 + \sigma_{ii}\sigma_{jj} + \frac{2(n-1)}{m-2}\sigma_{ij}^2\right] \\
&= (n-1)\frac{m^2}{(m-4)(m-2)}\left[\sigma_{ii}\sigma_{jj} + \frac{2n+m-4}{m-2}\sigma_{ij}^2\right].
\end{aligned}
$$

CHAPTER 4

MIXTURES OF NORMAL DISTRIBUTIONS

4.1. MIXTURE BY DISTRIBUTION FUNCTION

Muirhead (1982) gave a definition of scale mixture of vector variate normal distributions. Using Corollary 2.7.4.1, the scale mixture of matrix variate normal distributions can be defined as follows (Gupta and Varga, 1992c).

DEFINITION 4.1.1. Let M: $p \times n$, Σ: $p \times p$, and Φ: $n \times n$ be constant matrices such that $\Sigma > O$ and $\Phi > O$. Assume $G(z)$ is a distribution function on $(0,\infty)$. Let $X \in \mathbb{R}^{p\times n}$ and define

$$g(X) = \frac{1}{(2\pi)^{\frac{pn}{2}} |\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} \int_0^\infty z^{-\frac{pn}{2}} \operatorname{etr}\left(\frac{-1}{2z}((X - M)'\Sigma^{-1}(X - M)\Phi^{-1})\right) dG(z). \tag{4.1}$$

Then the m.e.c. distribution whose p.d.f. is $g(X)$ is called a scale mixture of matrix variate normal distributions.

REMARK 4.1.1. In this chapter we will denote the p.d.f. of $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$ by $f_{N_{p,n}(M,\Sigma \otimes \Phi)}(X)$. With this notation, (4.1) can be written as

$$g(X) = \int_0^\infty f_{N_{p,n}(M,z\Sigma \otimes \Phi)}(X)\, dG(z).$$

REMARK 4.1.2. Let X be a $p \times n$ random matrix. Then, $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$, $\Sigma > O$, $\Phi > O$ has the p.d.f. defined by (4.1) if and only if its characteristic function is

$$\phi_X(T) = \operatorname{etr}(iT'M) \int_0^\infty \operatorname{etr}\left(-\frac{z(T'\Sigma T\Phi)}{2}\right) dG(z), \tag{4.2}$$

that is

$$\psi(v) = \int_0^\infty \exp\left(-\frac{zv}{2}\right) dG(z).$$

This statement follows from a more general result proved later in Theorem 4.1.5.

REMARK 4.1.3. From (4.2), it follows that $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$, $\Sigma > O$, $\Phi > O$ has the p.d.f. defined by (4.1) if and only if $X \approx z^{\frac{1}{2}} Y$ where $Y \sim N_{p,n}(M, \Sigma \otimes \Phi)$, z has the distribution $G(z)$, and z and Y are independent.

The relationship between the characteristic function of a scale mixture of normal distributions and its stochastic representation is pointed out in the next theorem, due to Cambanis, Huang, and Simons (1981).

THEOREM 4.1.1. *Let* $X \sim E_{p,n}(O, I_p \otimes I_n, \psi)$ *have the stochastic representation* $X \approx rU$. *Let* $G(z)$ *be a distribution function on* $(0, \infty)$. *Then,*

$$\psi(v) = \int_0^\infty \exp\left(-\frac{zv}{2}\right) dG(z) \tag{4.3}$$

if and only if r *is absolutely continuous with p.d.f.*

$$\ell(r) = \frac{1}{2^{\frac{pn}{2}-1}\Gamma\left(\frac{pn}{2}\right)} r^{pn-1} \int_0^{\infty} \frac{\exp\left(-\frac{r^2}{2z}\right)}{z^{\frac{pn}{2}}} dG(z). \tag{4.4}$$

PROOF. First, assume (4.3) holds. Then, with the notation of Remark 4.1.3, we can write $X \approx z^{\frac{1}{2}}Y$. Hence, $rU \approx z^{\frac{1}{2}}Y$, and $\mathrm{tr}(r^2U'U) \approx \mathrm{tr}(zY'Y)$. Consequently,

$$r^2 \approx z\,\mathrm{tr}(Y'Y). \tag{4.5}$$

Since $Y \sim N_{p,n}(O, I_p \otimes I_n)$, and hence $\mathrm{tr}(Y'Y) \sim \chi^2_{pn}$. Now, χ^2_{pn} has the p.d.f.

$$f_1(w) = \frac{1}{2^{\frac{pn}{2}}\Gamma\left(\frac{pn}{2}\right)} w^{\frac{pn}{2}-1} e^{-\frac{w}{2}}, \quad w > 0.$$

Denoting r^2 by s, from (4.5) we obtain the p.d.f. of s as

$$\begin{aligned} f_2(s) &= \int_0^{\infty} \frac{1}{z} f_1\left(\frac{s}{z}\right) dG(z) \\ &= \int_0^{\infty} \frac{1}{z} \frac{1}{2^{\frac{pn}{2}}\Gamma\left(\frac{pn}{2}\right)} \left(\frac{s}{z}\right)^{\frac{pn}{2}-1} e^{-\frac{s}{2z}} dG(z) \\ &= \frac{1}{2^{\frac{pn}{2}}\Gamma\left(\frac{pn}{2}\right)} s^{\frac{pn}{2}-1} \int_0^{\infty} \frac{e^{-\frac{s}{2z}}}{z^{\frac{pn}{2}}} dG(z). \end{aligned}$$

Since $r^2 = s$, we have $J(s \to r) = 2r$ and so the p.d.f. of r is

$$\ell(r) = \frac{1}{2^{\frac{pn}{2}-1}\Gamma\left(\frac{pn}{2}\right)} r^{pn-1} \int_0^{\infty} \frac{\exp\left(-\frac{r^2}{2z}\right)}{z^{\frac{pn}{2}}} dG(z)$$

which is (4.4).

The other direction of the theorem follows from the fact that r and ψ determine each other. ■

The question arises when the p.d.f. of a m.e.c. distribution can be expressed as a scale mixture of matrix variate normal distributions. With the help of the next theorem, we can answer this question. This theorem was first derived by Schoenberg (1938) and the proof given here is due to Fang, Kotz and Ng (1990) who made use of a derivation of Kingman (1972).

THEOREM 4.1.2. *Let* ψ: $[0,\infty) \to \mathbb{R}$ *be a real function. Then,* $\psi(t't)$, $t \in \mathbb{R}^k$ *is a characteristic function for every* $k \geq 1$ *if and only if*

$$\psi(u) = \int_0^{\infty} \exp\left(-\frac{uz}{2}\right) dG(z) \tag{4.6}$$

where $G(z)$ *is a distribution function on* $[0,\infty)$.

PROOF. First, assume (4.6) holds. Let $k \geq 1$ be an integer and x be a k-dimensional random vector with p.d.f.

$$g(x) = \frac{1}{(2\pi)^{\frac{k}{2}}} \int_0^{\infty} z^{-\frac{k}{2}} \exp\left(-\frac{x'x}{2z}\right) dG(z).$$

Then, the characteristic function of x is

$$\phi_x(t) = \int_{\mathbb{R}^k} \exp(it'x)\, g(x)\, dx$$

$$= \int_{\mathbb{R}^k} \exp(it'x) \frac{1}{(2\pi)^{\frac{k}{2}}} \int_0^{\infty} z^{-\frac{k}{2}} \exp\left(-\frac{x'x}{2z}\right) dG(z)\, dx$$

$$= \int_0^{\infty} \int_{\mathbb{R}^k} \exp(it'x) \frac{1}{(2\pi z)^{\frac{k}{2}}} \exp\left(-\frac{x'x}{2z}\right) dx\, dG(z)$$

$$= \int_0^{\infty} \exp\left(-\frac{t'tz}{2}\right) dG(z),$$

where we used the fact that $\int_{\mathbb{R}^k} \exp(it'x) \frac{1}{(2\pi z)^{\frac{k}{2}}} \exp\left(-\frac{x'x}{2z}\right) dx$ is the characteristic function of $N_k(0,zI_k)$ and hence is equal to $\exp\left(-\frac{t'tz}{2}\right)$. Hence, $\psi(t't)$ is the characteristic function of x, where $t \in \mathbb{R}^k$.

Conversely, assume $\psi(t't)$ is a characteristic function for every $k \geq 1$. Then, we can choose an infinite sequence of random variables $(x_1,x_2,...)$, such that for every $k \geq 1$ integer, the characteristic function of $(x_1,x_2,...,x_k)$ is $\psi(t't)$, where $t \in \mathbb{R}^k$. Let $\{\pi(1),\pi(2),...,\pi(k)\}$ be a permutation of the numbers $\{1,2,...,k\}$. Then the characteristic function of $(x_{\pi(1)},x_{\pi(2)},...,x_{\pi(k)})$ is also $\psi(t't)$. Hence, $(x_1,x_2,...,x_k)$ and $(x_{\pi(1)},x_{\pi(2)},...,x_{\pi(k)})$ are identically distributed. Thus the infinite sequence $(x_1,x_2,...)$ is exchangeable.

Let $(\Omega,\mathcal{A},P)$ be the probability space on which $x_1,x_2,...,x_n,...$ are defined. From De Finetti's theorem (see Billingsley, 1979, p. 425), we know that there exists a sub σ-field of $\mathcal{A}$ say $\mathcal{F}$ such that conditional upon $\mathcal{F}$, x_i's are identically distributed and conditionally independent. The conditional independence means that for every integer n

$$P(x_i \in M_i,\ i = 1,2,...,n \mid \mathcal{F}) = \prod_{i=1}^{n} P(x_i \in M_i \mid \mathcal{F}) \qquad (4.7)$$

where $M_i \in \mathcal{B}(\mathbb{R})$, $i = 1,2,\ldots,n$. Moreover, $\mathcal{F}$ has the property that for every permutation of $\{1,2,\ldots,n\}$, say $\{\pi(1),\pi(2),\ldots,\pi(n)\}$,

$$P(x_i \in M_i,\ i = 1,2,\ldots,n \mid \mathcal{F}) = P(x_{\pi(i)} \in M_i,\ i = 1,2,\ldots,n \mid \mathcal{F}). \tag{4.8}$$

Let $(x_i \mid \mathcal{F})_\omega$ be a regular conditional distribution of x_i given $\mathcal{F}$ and $\mathcal{E}(x_i \mid \mathcal{F})_\omega$ be a conditional expectation of x_i given $\mathcal{F}$, $i = 1,2,\ldots$. Here, $\omega \in \Omega$ (see Billingsley, 1979, p. 390). Then, for $g: \mathbb{R} \to \mathbb{R}$ integrable, we have

$$\mathcal{E}(g(x_i) \mid \mathcal{F})_\omega = \int_{\mathbb{R}} g(x_i)\, d(x_i \mid \mathcal{F})_\omega \tag{4.9}$$

(see Billingsley, 1979, p. 399). Now, it follows from (4.7) and (4.8) that for fixed $\omega \in \Omega$,

$$\begin{aligned} P(x_1 \in M \mid \mathcal{F})_\omega &= P(x_1 \in M,\ x_i \in \mathbb{R},\ i = 2,\ldots,k \mid \mathcal{F})_\omega \\ &= P(x_k \in M,\ x_1 \in \mathbb{R},\ x_i \in \mathbb{R},\ i = 2,\ldots,k-1 \mid \mathcal{F})_\omega \\ &= P(x_k \in M \mid \mathcal{F})_\omega. \end{aligned}$$

Hence, for fixed $\omega \in \Omega$, and any positive integer k,

$$(x_1 \mid \mathcal{F})_\omega = (x_k \mid \mathcal{F})_\omega \text{ almost everywhere.} \tag{4.10}$$

Define

$$\phi(t)_\omega = \int_{\mathbb{R}} e^{itx_1}\, d(x_1 \mid \mathcal{F})_\omega\,, \tag{4.11}$$

where $t \in \mathbb{R}$, and $\omega \in \Omega$. Then, from (4.10) we get

$$\phi(t)_\omega = \int_{\mathbb{R}} e^{itx_k}\, d(x_k \mid \mathcal{F})_\omega. \tag{4.12}$$

For fixed $t \in \mathbb{R}$, $\phi(t)_\omega$ is a $\mathcal{F}$-measurable random variable. On the other hand, for fixed ω, $\phi(t)_\omega$ is a continuous function of t since it is the characteristic function of the distribution defined by $(x_1 \mid \mathcal{F})_\omega$. Since $\phi(t)_\omega$ is a characteristic function, we have $|\phi(t)_\omega| \le 1$ and $\phi(0)_\omega = 1$. We also have

$$\phi(-t)_\omega = \mathcal{E}(e^{i(-t)x_1} \mid \mathcal{F})_\omega = \overline{\mathcal{E}(e^{itx_1} \mid \mathcal{F})_\omega} = \overline{\phi(t)}_\omega.$$

From (4.9) and (4.12), we see that for any positive integer k,

$$\phi(t)_\omega = \mathcal{E}(e^{itx_k} \mid \mathcal{F})_\omega. \tag{4.13}$$

Using (4.7) and (4.13), we get

$$\mathcal{E}\left(\exp\left(i \sum_{j=1}^{n} t_j x_j\right) \Big| \mathcal{F}\right)_\omega = \prod_{j=1}^{n} \mathcal{E}\left(e^{it_j x_j} \mid \mathcal{F}\right)_\omega$$

$$= \prod_{j=1}^{n} \phi(t_j)_\omega. \tag{4.14}$$

Therefore,

$$\mathcal{E}\left(\exp\left(i \sum_{j=1}^{n} t_j x_j\right)\right) = \mathcal{E}\left(\prod_{j=1}^{n} \phi(t_j)\right).$$

The left-hand side of the last expression is the characteristic function of $(x_1, x_2, \ldots, x_n)$. Hence, we get

$$\mathcal{E}\left(\prod_{j=1}^{n}\phi(t_j)\right)=\psi\left(\sum_{j=1}^{n}t_j^2\right). \tag{4.15}$$

Let u and v be real numbers and define $w = (u^2 + v^2)^{\frac{1}{2}}$. Then, we can write

$$\begin{aligned}\mathcal{E}(\,|\,\phi(w)-\phi(u)\phi(v)\,|^2) &= \mathcal{E}((\phi(w)-\phi(u)\phi(v))(\overline{\phi(w)-\phi(u)\phi(v)}))\\ &= \mathcal{E}((\phi(w)-\phi(u)\phi(v))(\phi(-w)-\phi(-u)\phi(-v)))\\ &= \mathcal{E}(\phi(w)\phi(-w)) + \mathcal{E}(\phi(u)\phi(-u)\phi(v)\phi(-v))\\ &\quad - \mathcal{E}(\phi(w)\phi(-u)\phi(-v)) - \mathcal{E}(\phi(-w)\phi(u)\phi(v)).\end{aligned}$$

Using (4.15), we see that all four terms in the last expression equal $\psi(2w^2)$, hence,

$$\mathcal{E}(\,|\,\phi(w)-\phi(u)\phi(v)\,|^2) = 0.$$

That means, $\phi(u)\phi(v) = \phi(w)$ with probability one, or equivalently $\phi(u)_\omega\,\phi(v)_\omega = \phi(w)_\omega$ for $\omega \in C(u,v)$, where $C(u,v) \subset \Omega$ and $P(C(u,v)) = 1$.

Now, we have

$$\phi(u)_\omega\,\phi(0)_\omega = \phi(\,|\,u\,|\,) \text{ for } \omega \in C(u,0)$$

and

$$\phi(-u)_\omega\,\phi(0)_\omega = \phi(\,|\,u\,|\,) \text{ for } \omega \in C(-u,0).$$

But $\phi(0)_\omega = 1$ and so $\phi(u)_\omega = \phi(-u)_\omega$ for $\omega \in C(u,0) \cap C(-u,0)$. However we have already shown that $\phi(-u)_\omega = \overline{\phi(u)_\omega}$. Hence $\phi(u)_\omega = \overline{\phi(u)_\omega}$ and therefore $\phi(u)_\omega$ is real for $\omega \in C(u,0) \cap C(-u,0)$. Similarly, $\phi(v)_\omega$ is real for $\omega \in C(0,v) \cap c(0,-v)$.

Define

$$C = \bigcap_{u,v \text{ rational numbers}} \{C(u,v) \cap C(u,0) \cap C(-u,0) \cap C(0,v) \cap C(0,-v)\}.$$

Then, $P(C) = 1$, $\phi(u)_\omega \, \phi(v)_\omega = \phi(\sqrt{u^2 + v^2}\,)_\omega$ and $\phi(u)_\omega$ is real for u, v rational and $\omega \in C$. However we have already shown that $\phi(t)_\omega$ is continuous in t for fixed $\omega \in \Omega$. Hence $\phi(t)_\omega$ is real for all $t \in \mathbb{R}$ and $\omega \in C$. Moreover, with the notation $\xi(t)_\omega = \phi\left(t^{\frac{1}{2}}\right)_\omega$, we have

$$\xi(t_1)_\omega \, \xi(t_2)_\omega = \xi(t_1 + t_2)_\omega, \quad t_1 \geq 0, \ t_2 \geq 0 \tag{4.16}$$

for $t_1^{\frac{1}{2}}, t_2^{\frac{1}{2}}$ rational. Since $\xi(t)_\omega$ is continuous in t, we conclude that (4.16) holds for all nonnegative numbers t_1 and t_2. Now using Corollary 1.4.1.1 we find that the solution of (4.16) is

$$\xi(t)_\omega = e^{-k(\omega)t}, \quad t \geq 0$$

where $k(\omega)$ is a positive number depending on ω (see Feller, 1957, p. 413). So, we can write

$$\phi(t)_\omega = e^{-\frac{z(\omega)}{2} t^2}$$

where $z(\omega)$ depends on ω. This is also true if $t < 0$ since $\phi(-t)_\omega = \phi(t)_\omega$. Now, $z(\omega)$ defines a random variable z with probability one. Therefore, we can write

$$\phi(t) = e^{-\frac{z}{2} t^2} \tag{4.17}$$

and hence $z = -2 \log \phi(1)$. Since, $\phi(1)$ is $\mathcal{F}$-measurable, so is z, and we have

$$\mathcal{E}(y \mid z) = \mathcal{E}(\mathcal{E}(y \mid \mathcal{F}) \mid z) \tag{4.18}$$

for any random variable y. Now take $y = e^{itx_1}$, then using (4.13), (4.17) and (4.18) we get

$$\begin{aligned}
\mathcal{E}(e^{itx_1} \mid z) &= \mathcal{E}(\mathcal{E}(e^{itx_1} \mid \mathcal{F}) \mid z) \\
&= \mathcal{E}(\phi(t) \mid z) \\
&= \mathcal{E}\left(e^{-\frac{z}{2}t^2} \mid z\right) \\
&= e^{-\frac{z}{2}t^2}.
\end{aligned}$$

Hence, the characteristic function of x_1 is

$$\begin{aligned}
\psi(t^2) &= \mathcal{E}(e^{itx_1}) \\
&= \mathcal{E}(\mathcal{E}(e^{itx_1} \mid z)) \\
&= \mathcal{E}(e^{-\frac{z}{2}t^2}) \\
&= \int_0^\infty \exp\left(-\frac{zt^2}{2}\right) dG(z),
\end{aligned}$$

where $G(z)$ denotes the distribution function of z. Thus,

$$\psi(u) = \int_0^\infty \exp\left(-\frac{zu}{2}\right) dG(z)$$

which proves (4.6). ■

We also need the following lemma.

LEMMA 4.1.1. *Let* $X \sim E_{p,n}(O,\Sigma \otimes \Phi,\psi)$. *Let* Y *be a* $q \times m$ *submatrix of* X *such that* $qm < pn$ *and* $P(Y = O) = 0$. *Then* Y *is absolutely continuous.*

PROOF. From Theorem 3.1.6, it follows that Y is absolutely continuous. ∎

Now, we can prove the following theorem.

THEOREM 4.1.3. *Let* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ *such that* $P(X = M) = 0$. *Then, the distribution of* X *is absolutely continuous and the p.d.f. of* X *can be written as a scale mixture of matrix variate normal distributions if and only if for every integer* $k > pn$ *there exists* $Y \sim E_{q,m}(M_1,\Sigma_1 \otimes \Phi_1,\psi_1)$, *such that* $qm \geq k$, $\Sigma_1 > O$, $\Phi_1 > O$ *and* Y *has a submatrix* Y_0 *with* $Y_0 \approx X$.

PROOF. First, assume that for every integer $k > pn$ there exists $Y \sim E_{q,m}(M_1,\Sigma_1 \otimes \Phi_1,\psi_1)$ such that $qm \geq k$ and Y has a submatrix Y_0 with $Y_0 \approx X$. Then, for fixed k, from Remark 2.3.1, it follows that $\psi_1 = \psi$. Moreover, from Lemma 4.1.1, it follows that the distribution of Y_0 and consequently that of X is absolutely continuous. Let $\mathbf{w} = \operatorname{vec}\left(\Sigma_1^{-\frac{1}{2}}(Y - M_1)\Phi_1^{-\frac{1}{2}}\right)'$, then $\mathbf{w} \sim E_{qm}(\mathbf{o},I_{qm},\psi)$. Let $\mathbf{v}$ be a k-dimensional subvector of $\mathbf{w}$. Then, $\mathbf{v} \sim E_k(\mathbf{o},I_k,\psi)$ and the characteristic function of $\mathbf{v}$ is $\phi_{\mathbf{v}}(\mathbf{t}) = \psi(\mathbf{t}'\mathbf{t})$, where $\mathbf{t} \in \mathbb{R}^k$. Using Theorem 4.1.2, we get $\psi(u) = \int_0^\infty \exp\left(-\frac{uz}{2}\right) dG(z)$. Therefore, the p.d.f. of X is

$$g(X) = \int_0^\infty f_{N_{p,n}(M,z\Sigma \otimes \Phi)}(X)\, dG(z).$$

Next, assume that X can be written as a scale mixture of matrix variate normal distributions; that is, the p.d.f. of X is

$g(X) = \int_0^\infty f_{N_{p,n}(M,z\Sigma \otimes \Phi)}(X)\, dG(z)$. Then, we have

$$\psi(u) = \int_0^\infty \exp\left(-\frac{uz}{2}\right) dG(z).$$

It follows from Theorem 4.1.2, that $\psi(\mathbf{t}'\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^k$ is a characteristic function for every $k \geq 1$. Choose $k > pn$, and let $q \geq p$, $m \geq n$, such that $qm \geq k$. Define a qm-dimensional random vector $\mathbf{w}$ such that $\mathbf{w} \sim E_{qm}(\mathbf{o}, I_{qm}, \psi)$. Let $\mathbf{w} = \text{vec}(S')$ where S is $q \times m$ matrix, then $S \sim E_{q,m}(O, I_q \otimes I_m, \psi)$. Further define

$$\Sigma_1 = \begin{pmatrix} \Sigma & O \\ O & I_{q-p} \end{pmatrix}, \quad \Phi_1 = \begin{pmatrix} \Phi & O \\ O & I_{m-n} \end{pmatrix} \text{ and } M_1 = \begin{pmatrix} M & O \\ O & O \end{pmatrix}$$

where M_1 is $q \times m$. Let $Y = \Sigma_1^{\frac{1}{2}} S \Phi^{\frac{1}{2}} + M_1$. Then,

$$Y \sim E_{q,m}\left(\begin{pmatrix} M & O \\ O & O \end{pmatrix}, \begin{pmatrix} \Sigma & O \\ O & I_{q-p} \end{pmatrix} \otimes \begin{pmatrix} \Phi & O \\ O & I_{m-n} \end{pmatrix}, \psi\right).$$

Partition Y into $Y = \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{22} & Y_{22} \end{pmatrix}$ where Y_{11} is $p \times n$. Then, $Y_{11} \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ and hence $X \approx Y_{11}$. ■

EXAMPLE 4.1.1. Here, we list some m.e.c. distributions together with the distribution function $G(z)$ which generates the p.d.f. through (4.1).

a) Matrix variate normal distribution:

$$G(1) = 1.$$

b) ε-contaminated matrix variate normal distribution:

$$G(1) = 1 - \varepsilon, \ G(\sigma^2) = \varepsilon.$$

c) Matrix variate Cauchy distribution:

$$G(z) = \frac{1}{4\pi} \int_0^z t^{-\frac{3}{2}} e^{-\frac{1}{2t}} dt.$$

d) Matrix variate t-distribution with m degrees of freedom:

$$G(z) = \frac{\left(\frac{m}{2}\right)^{\frac{m}{2}}}{\Gamma\left(\frac{m}{2}\right)} \int_0^z t^{-(1+\frac{m}{2})} e^{-\frac{m}{2t}} dt.$$

Here, (a) and (b) are obvious, and (c) and (d) will be shown in Section 4.2.

Next, we give an example which shows that the p.d.f. of an absolutely continuous elliptically contoured distribution is not always expressible as the scale mixture of normal distributions.

EXAMPLE 4.1.2. Let x be a one-dimensional random variable with p.d.f. $f(x) = \frac{\sqrt{2}}{\pi} \frac{1}{1+x^4}$. Assume $f(x)$ has a scale mixture representation. Then, from Theorem 4.1.3, there exists a $q \times m$ dimensional elliptical distribution $Y \sim E_{q,m}(M_1, \Sigma_1 \otimes \Phi_1, \psi_1)$, such that $qm > 5$ and one element of Y is identically distributed as x. Therefore, there exists a 5-dimensional random vector $\mathbf{w}$ such that $\mathbf{w} \sim E_5(\mathbf{m}_2, \Sigma_2, \psi)$ and $w_1 \approx x$.

Now, $f(x) = h(x^2)$ where $h(z) = \frac{\sqrt{2}}{\pi} \frac{1}{1+z^2}$. Let the p.d.f. of $\mathbf{w}$ be $f_1(\mathbf{w}) = h_1((\mathbf{w} - \mathbf{m}_2)' \Sigma_2^{-1} (\mathbf{w} - \mathbf{m}_2))$. It follows, from Theorem 3.1.4, that $h_1(z) = \frac{1}{\pi^2} \frac{\partial^2 h(z)}{\partial z^2}$. We have

$$\frac{\partial h(z)}{\partial z} = \frac{\sqrt{2}}{\pi} \frac{\partial}{\partial z} \frac{1}{1+z^2}$$

$$= -\frac{2\sqrt{2}}{\pi} \frac{z}{(1+z^2)^2},$$

and

$$\frac{\partial^2 h(z)}{\partial z} = -\frac{2\sqrt{2}}{\pi} \frac{1-3z^2}{(1+z^2)^3}.$$

Consequently we get $h_1(z) = \frac{2\sqrt{2}}{\pi^3} \frac{3z^2-1}{(1+z^2)^3}$. However, $h_1(z) < 0$ for $0 < z < \frac{1}{\sqrt{3}}$ and hence $h(x^2)$ cannot be a p.d.f. This is a contradiction. Therefore, $f(x)$ cannot be written as a scale mixture of normal distributions.

Next, we prove some important theorems about scale mixture representations.

THEOREM 4.1.4. *Let* λ: $\mathbb{R}^{p\times n} \to \mathbb{R}^{q\times m}$ *be a Borel-measurable matrix variate function. Assume that if* $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$ *then the p.d.f. of* $W = \lambda(X)$ *is* $\ell^{\lambda}_{N_{p,n}(M,\Sigma\otimes\Phi)}(W)$. *Then, if* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ *with p.d.f.*

$g(X) = \int_0^\infty f_{N_{p,n}(M,z\Sigma\otimes\Phi)}(X)dG(z)$, *the p.d.f. of* $W = \lambda(X)$ *is*

$$\ell(W) = \int_0^\infty \ell^{\lambda}_{N_{p,n}(M,z\Sigma\otimes\Phi)}(W)\, dG(z).$$

PROOF. Let $A \subset \mathbb{R}^{q\times m}$. Then,

$$\int_A \ell(W)dW = \int_A \int_0^\infty \ell^{\lambda}_{N_{p,n}(M,z\Sigma\otimes\Phi)}(W)\, dG(z)dW$$

$$= \int_0^\infty \int_A \ell^{\lambda}_{N_{p,n}(M,z\Sigma\otimes\Phi)}(W)dW\, dG(z)$$

$$= \int_0^\infty P(\lambda(X) \in A \mid X \sim N_{p,n}(M, z\Sigma \otimes \Phi))\, dG(z)$$

$$= \int_0^\infty \int_{\mathbb{R}^{p\times n}} \chi_A(\lambda(X)) f_{N_{p,n}(M,z\Sigma\otimes\Phi)}(X) dX\, dG(z)$$

$$= \int_{\mathbb{R}^{p\times n}} \chi_A(\lambda(X)) \int_0^\infty f_{N_{p,n}(M,z\Sigma\otimes\Phi)}(X)\, dG(z)\, dX$$

$$= \int_{\mathbb{R}^{p\times n}} \chi_A(\lambda(X))\, g(X)\, dX$$

$$= P(\lambda(x) \in A \mid X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)). \blacksquare$$

COROLLARY 4.1.4.1. *Let* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ *with p.d.f.* $g(X) = \int_0^\infty f_{N_{p,n}(M,z\Sigma\otimes\Phi)}(X)\, dG(z)$. *Let* C: $q \times m$, A: $q \times p$, *and* B: $n \times m$ *be constant matrices, such that* $\text{rk}(A) = q$ *and* $\text{rk}(B) = m$. *Then, from Theorem 4.1.4 and Theorem 2.1.2, it follows that the p.d.f. of* $AXB + C$ is

$$g^*(X) = \int_0^\infty f_{N_{p,n}(AMB+C,z(A\Sigma A')\otimes(B'\Phi B))}(X)\, dG(z).$$

Furthermore, if X, M, *and* Σ *are partitioned into*

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix},\quad M = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix},\quad \text{and } \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

where X_1: $q \times n$, M_1: $q \times n$, *and* Σ_{11}: $q \times q$, $1 \le q < p$, *then the p.d.f. of* X_1 *is*

$$g_1(X) = \int_0^\infty f_{N_{p,n}(M_1,z\Sigma_{11}\otimes\Phi)}(X)\, dG(z).$$

COROLLARY 4.1.4.2. *Let* $X \sim E_{p,n}(\mu e_n', \Sigma \otimes I_n, \psi)$ *with p.d.f.* $g(X)$, *where* $g(X) = \int_0^\infty f_{N_{p,n}(\mu e_n', z\Sigma\otimes\Phi)}(X)\, dG(z)$. *Then,*

a) *the p.d.f. of* $\mathbf{y}_1 = \frac{X\mathbf{e}_n}{n}$ *is*

$$g_1(\mathbf{y}_1) = \int_0^\infty f_{N_p(\mu, z\,\Sigma/n)}(\mathbf{y}_1)\, dG(z),$$

b) *the p.d.f. of* $Y_2 = X\left(I_n - \frac{\mathbf{e}_n\mathbf{e}_n'}{n}\right)X'$ *is*

$$g_2(Y_2) = \int_0^\infty f_{W_p(z\Sigma, n-1)}(Y_2)\, dG(z),$$

and

c) *the p.d.f. of* $Y_3 = XX'$, *for* $\mu = \mathbf{o}$, *is*

$$g_3(Y_3) = \int_0^\infty f_{W_p(z\Sigma, n)}(Y_3)\, dG(z).$$

THEOREM 4.1.5. *Let* λ: $\mathbb{R}^{p\times n} \to \mathbb{R}^{q\times m}$ *be a Borel-measurable matrix variate function. Assume that if* $X \sim N_{p,n}(M, \Sigma \otimes \Phi)$, *then* $\mathcal{E}(\lambda(X))$ *exists and it is denoted by* $\mathcal{E}_{N_{p,n}(M,\Sigma\otimes\Phi)}(\lambda(X))$. *Then, if* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ *with p.d.f.* $g(X) = \int_0^\infty f_{N_{p,n}(M, z\Sigma\otimes\Phi)}(X)\, dG(z)$, *such that* $\mathcal{E}(\lambda(X))$ *exists and it is denoted by* $\mathcal{E}_{E_{p,n}(M,\Sigma\otimes\Phi,\psi)}(\lambda(X))$, *we have*

$$\mathcal{E}_{E_{p,n}(M,\Sigma\otimes\Phi,\psi)}(\lambda(X)) = \int_0^\infty \mathcal{E}_{N_{p,n}(M,z\Sigma\otimes\Phi,\psi)}(\lambda(X))\, dG(z).$$

PROOF.

$$\mathcal{E}_{E_{p,n}(M,\Sigma\otimes\Phi,\psi)}(\lambda(X)) = \int_{\mathbb{R}^{p\times n}} \lambda(X)\, g(X)\, dX$$

$$= \int_{\mathbb{R}^{p\times n}} \lambda(X) \int_0^\infty f_{N_{p,n}(M,z\Sigma\otimes\Phi)}(X)\, dG(z) dX$$

$$= \int_0^\infty \int_{\mathbb{R}^{p\times n}} \lambda(X) f_{N_{p,n}(M,z\Sigma\otimes\Phi)}(X)\, dX\, dG(z)$$

$$= \int_0^\infty \mathcal{E}_{N_{p,n}(M,z\Sigma\otimes\Phi)}(\lambda(X))\, dG(z). \blacksquare$$

COROLLARY 4.1.5.1. *With the notations of Theorem 4.1.5, if* Cov(X) *exists, then*

$$\mathrm{Cov}(X) = \left(\int_0^\infty z\, dG(z)\right) \Sigma \otimes \Phi.$$

Next, we give a theorem which shows the relationship between the characteristic function of a scale mixture of normal distributions and the characteristic function of a conditional distribution. This theorem is due to Cambanis, Huang, and Simons (1981).

THEOREM 4.1.6. *Let* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ *with p.d.f.* $g(X) = \int_0^\infty f_{N_{p,n}(M,z\Sigma\otimes\Phi)}(X)\, dG(z)$. *Let* $X, M,$ *and* Σ *be partitioned as* $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, $M = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}$, $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, *where* X_1 *and* M_1 *are* $q \times n$ *and* Σ_{11} *is* $q \times q$. *Then, the conditional p.d.f. of* $X_1 | X_2$ *can be written as*

$$g_1(X_1 \mid X_2) = \int_0^\infty f_{N_{q,n}(M_1+\Sigma_{12}\Sigma_{22}^{-1}(X_2-M_2), z\Sigma_{11\cdot 2}\otimes\Phi)}(X)\, dG_{q(X_2)}(z), \tag{4.19}$$

where $q(X_2) = \mathrm{tr}((X_2 - M_2)' \Sigma_{22}^{-1}(X_2 - M_2)\Phi^{-1})$ *and*

$$G_{a^2}(z) = \frac{\int_0^z v^{-\frac{(p-q)n}{2}} \exp\left(-\frac{a^2}{2v}\right) dG(v)}{\int_0^\infty v^{-\frac{(p-q)n}{2}} \exp\left(-\frac{a^2}{2v}\right) dG(v)} \quad \textit{if } a > 0,\ z \geq 0, \tag{4.20}$$

and $G_0(z) = 1$ *if* $z \geq 0$.

PROOF. Let $X \approx r\Sigma^{-\frac{1}{2}} U \Phi^{-\frac{1}{2}} + M$ be the stochastic representation of X. Then, (4.4) gives the p.d.f. of r. It follows, from Theorem 2.6.4, that the stochastic representation of $X_1 \mid X_2$ has the form

$$X_1 \mid X_2 \approx r_{q(X_2)}\, \Sigma_{11\cdot 2}^{-1/2}\, U_1\, \Phi^{-\frac{1}{2}} + (M_1 + (X_2 - M_2)\Sigma_{22}^{-1}\Sigma_{21}).$$

Here, $\mathrm{vec}(U_1')$ is uniformly distributed on S_{qn}. It follows, from (2.31) and (4.4), that

$$P(r_{a^2} \leq c) = \frac{\int_a^{\sqrt{c^2+a^2}} (r^2-a^2)^{\frac{qn}{2}-1} r^{-(pn-2)} r^{pn-1} \int_0^\infty \frac{\exp\left(-\frac{r^2}{2s}\right)}{s^{\frac{pn}{2}}} dG(s)dr}{\int_a^\infty (r^2-a^2)^{\frac{pn}{2}-1} r^{-(pn-2)} r^{pn-1} \int_0^\infty \frac{\exp\left(-\frac{r^2}{2s}\right)}{s^{\frac{pn}{2}}} dG(s)dr}.$$

Let $y^2 = r^2 - a^2$, then $J(r \to y) = \frac{r}{y}$, and we have

$$P(r_{a^2} \le c) = \frac{\int_0^c y^{qn-1} \int_0^\infty s^{-\frac{pn}{2}} \exp\left(-\frac{y^2+a^2}{2s}\right) dG(s)dy}{\int_0^\infty y^{qn-1} \int_0^\infty s^{-\frac{pn}{2}} \exp\left(-\frac{y^2+a^2}{2s}\right) dG(s)dy}$$

$$= \frac{\int_0^c y^{qn-1} \int_0^\infty s^{-\frac{pn}{2}} \exp\left(-\frac{y^2}{2s}\right) \exp\left(-\frac{a^2}{2s}\right) dG(s)dy}{\int_0^\infty s^{-\frac{pn}{2}} \exp\left(-\frac{a^2}{2s}\right) \int_0^\infty y^{qn-1} \exp\left(-\frac{y^2}{2s}\right) dy\, dG(s)}. \tag{4.21}$$

In order to compute $\int_0^\infty y^{qn-1} \exp\left(-\frac{y^2}{2s}\right) dy$ we substitute $t = \frac{y^2}{2s}$. Then, $J(y \to t) = \frac{s}{y}$, and hence

$$\int_0^\infty y^{qn-1} \exp\left(-\frac{y^2}{2s}\right) dy = (2s)^{\frac{qn}{2}-1} s \int_0^\infty t^{\frac{qn}{2}-1} \exp(-t)dt$$

$$= 2^{\frac{qn}{2}-1} s^{\frac{qn}{2}} \Gamma\left(\frac{qn}{2}\right).$$

Substituting this into (4.2.1), we get

$$P(r_{a^2} \le c) = \frac{\int_0^c y^{qn-1} \int_0^\infty s^{-\frac{pn}{2}} \exp\left(-\frac{y^2}{2s}\right) \exp\left(-\frac{a^2}{2s}\right) dG(s)dy}{\int_0^\infty 2^{\frac{qn}{2}-1} \Gamma\left(\frac{qn}{2}\right) s^{-\frac{(p-q)n}{2}} \exp\left(-\frac{a^2}{2s}\right) dG(s)}$$

$$= \int_0^c \frac{1}{2^{\frac{qn}{2}-1}\Gamma\left(\frac{qn}{2}\right)} y^{qn-1} \int_0^\infty \frac{\exp\left(-\frac{y^2}{2s}\right)}{s^{\frac{qn}{2}}} \frac{s^{-\frac{(p-q)n}{2}} \exp\left(-\frac{a^2}{2s}\right)}{\int_0^\infty s^{-\frac{(p-q)n}{2}} \exp\left(-\frac{a^2}{2s}\right) dG(s)} dG(s)\, dy. \tag{4.22}$$

Define the distribution function

$$G_{a^2}(z) = \frac{\int_0^z s^{-\frac{(p-q)n}{2}} \exp\left(-\frac{a^2}{2s}\right) dG(s)}{\int_0^\infty s^{-\frac{(p-q)n}{2}} \exp\left(-\frac{a^2}{2s}\right) dG(s)}.$$

Then, from (4.22), we get

$$P(r_{a^2} \le c) = \int_0^c \frac{1}{2^{\frac{qn}{2}-1}\Gamma\left(\frac{qn}{2}\right)} y^{qn-1} \int_0^\infty \frac{\exp\left(-\frac{y^2}{2z}\right)}{z^{\frac{qn}{2}}} d(G_{a^2}(z))dy.$$

Hence the p.d.f. of r_{a^2} is

$$\ell^*(r_{a^2}) = \frac{1}{2^{\frac{qn}{2}-1}\Gamma\left(\frac{qn}{2}\right)} r_a^{qn-1} \int_0^\infty \frac{\exp\left(-\frac{r_a^2}{2z}\right)}{z^{\frac{qn}{2}}} dG(z).$$

Now, from Theorem 4.1.1, this means that the p.d.f. of $X_1 \mid X_2$ has the form (4.19) with $G_{a^2}(z)$ defined by (4.20). ■

4.2. MIXTURE BY WEIGHTING FUNCTION

Chu (1973) showed another way to obtain the p.d.f. of a m.e.c. distribution from the density functions of matrix variate normal distributions. For this purpose, he used Laplace transform. We recall here that if $f(t)$ is a real function defined on the set of nonnegative real numbers, then its Laplace transform, $\mathcal{L}[f(t)]$ is defined by

$$g(s) = \mathcal{L}[f(t)]$$

$$= \int_0^\infty e^{-st} f(t)dt.$$

Moreover, the inverse Laplace transform of a function $g(s)$ (see Abramowitz and Stegun, 1965, p. 1020) is defined by

$$f(t) = \mathcal{L}^{-1}[g(s)]$$

$$= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{st} g(s)ds,$$

where c is an appropriately chosen real number. It is known that $\mathcal{L}^{-1}[g(s)]$ exists if $g(s)$ is differentiable for sufficiently large s and $g(s) = o(s^{-k})$ as $s \to \infty$, $k > 1$.

The following theorem was proved by Chu (1973) for the vector variate case.

THEOREM 4.2.1. *Let* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ *with p.d.f.* $g(X)$ *where* $g(X) = |\Sigma|^{-\frac{n}{2}} |\Phi|^{-\frac{p}{2}} h(\mathrm{tr}((X - M)' \Sigma^{-1}(X - M)\Phi^{-1}))$. *If* $h(t)$, $t \in [0,\infty)$ *has the inverse Laplace transform, then we have*

$$g(X) = \int_0^\infty f_{N_{p,n}(M, z^{-1}\Sigma \otimes \Phi)}(X)\, w(z)dz \tag{4.23}$$

where

$$w(z) = (2\pi)^{\frac{pn}{2}} z^{-\frac{pn}{2}} \mathcal{L}^{-1}[h(2t)]. \tag{4.24}$$

PROOF. From (4.24), we get

$$h(2t) = \mathcal{L}[(2\pi)^{-\frac{pn}{2}} z^{\frac{pn}{2}} w(z)]$$

$$= \int_0^\infty e^{-tz} (2\pi)^{-\frac{pn}{2}} z^{\frac{pn}{2}} w(z)\, dz.$$

Hence,

$$g(X) = |\Sigma|^{-\frac{n}{2}} |\Phi|^{-\frac{p}{2}} \int_0^\infty \operatorname{etr}\left(-\frac{1}{2} \operatorname{tr}((X-M)'\Sigma^{-1}(X-M)\Phi^{-1})z\right)(2\pi)^{-\frac{pn}{2}} z^{\frac{pn}{2}} w(z)dz$$

$$= \int_0^\infty \frac{1}{(2\pi)^{\frac{pn}{2}} |z^{-1}\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} \operatorname{etr}\left(-\frac{1}{2} \operatorname{tr}((X-M)'\Sigma^{-1}(X-M)\Phi^{-1})z\right) w(z)dz$$

$$= \int_0^\infty f_{N_{p,n}(M, z^{-1}\Sigma \otimes \Phi)}(X)\, w(z)\, dz. \blacksquare$$

REMARK 4.2.1. Let $g(\mathbf{x}) = \int_0^\infty f_{N_p(\mathbf{0}, z^{-1}\Sigma)}(\mathbf{x})\, w(z)\, dz$. Assume, $w(z)$ is a function (not a functional), and define $u(z) = \frac{1}{z^2} w\left(\frac{1}{z}\right)$. Then,

$$g(X) = \int_0^\infty f_{N_{p,n}(M,z\Sigma\otimes\Phi)} u(z)\, dz. \tag{4.25}$$

Indeed, let $t = \frac{1}{z}$. Then, $J(z \to t) = \frac{1}{t^2}$ and so (4.23) can be rewritten as

$$\begin{aligned} g(X) &= \int_0^\infty f_{N_{p,n}(M,t\Sigma\otimes\Phi)}(X)\, w\left(\frac{1}{t}\right)\frac{1}{t^2}\, dt \\ &= \int_0^\infty f_{N_{p,n}(M,t\Sigma\otimes\Phi)}(X)\, u(t)dt. \end{aligned}$$

REMARK 4.2.2. Even if $w(z)$ is a functional the representation (4.25) may exist as parts (a) and (b) of the next example show.

EXAMPLE 4.2.1. Here, we list some m.e.c. distributions together with the functions $w(z)$ and $u(z)$ which generate the p.d.f. through (4.23) and (4.25).

a) Matrix variate normal distribution;

$w(z) = \delta(z - 1)$, and
$u(z) = \delta(z - 1)$.

b) ε-contaminated matrix variate normal distribution;

$w(z) = (1 - \varepsilon)\, \delta(z - 1) + \varepsilon\delta(z - \sigma^{-2})$, and
$u(z) = (1 - \varepsilon)\, \delta(z - 1) + \varepsilon\delta(z - \sigma^2)$.

c) Matrix variate Cauchy distribution:

$w(z) = \dfrac{1}{4\pi\sqrt{ze^z}}$, and

$$u(z) = \frac{1}{4\pi z\sqrt{ze^{\frac{1}{2}}}}\,.$$

d) Matrix variate t-distribution with m degrees of freedom;

$$w(z) = \frac{\left(\frac{mz}{2}\right)^{\frac{m}{2}} e^{-\frac{mz}{2}}}{z\Gamma\left(\frac{m}{2}\right)}\,, \text{ and}$$

$$u(z) = \frac{\left(\frac{m}{2z}\right)^{\frac{m}{2}} e^{-\frac{m}{2z}}}{z\Gamma\left(\frac{m}{2}\right)}\,.$$

e) The one-dimensional distribution with p.d.f.

$$g(x) = \frac{\sqrt{2}}{\pi\sigma}\frac{1}{1+\left(\frac{x}{\sigma}\right)^4};$$

$$w(z) = \frac{1}{\sqrt{\pi z}}\sin\frac{z}{2}\,, \text{ and}$$

$$u(z) = \frac{1}{\sqrt{\pi z^3}}\sin\frac{1}{2z}\,.$$

The functions $w(z)$ in parts (a), (c), (d), and (e) are given in Chu (1973) and $u(z)$ can be easily computed from $w(z)$.

REMARK 4.2.3. It may be noted that $u(z)$ is not always nonnegative as part (e) of Example 4.2.1 shows. However, if it is nonnegative then defining $G(z) = \int_0^z u(s)\,ds$, (4.25) yields $g(X) = \int_0^\infty f_{N_{p,n}(M,z\Sigma\otimes\Phi)}(X)\,dG(z)$ which is the expression given in Remark 4.1.1. We have to see that $\int_0^\infty u(s)ds = 1$ but this will follow from the next theorem if we take $v = 0$ in (4.26). Therefore, using Example 4.2.1 we obtain the results in parts (c) and (d) of Example 4.1.1.

Now, we can state theorems similar to those in Section 4.1.

THEOREM 4.2.2. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ *be absolutely continuous. Then,* X *has the p.d.f. defined by (4.23) if and only if the characteristic function of* X *is*

$$\phi_X(T) = \operatorname{etr}(iT'M) \int_0^\infty \operatorname{etr}\left(-\frac{(T'\Sigma T\Phi)}{2z}\right) w(z)dz,$$

that is

$$\psi(v) = \int_0^\infty \exp\left(-\frac{v}{2z}\right) w(z)dz.$$

Also, X *has the p.d.f. defined by (4.25) if and only if*

$$\phi_X(T) = \operatorname{etr}(iT'M) \int_0^\infty \operatorname{etr}\left(-\frac{z(T'\Sigma T\Phi)}{2}\right) u(z)dz,$$

that is

$$\psi(v) = \int_0^\infty \exp\left(-\frac{zv}{2}\right) u(z)dz. \tag{4.26}$$

THEOREM 4.2.3. *Let* λ: $\mathbb{R}^{p \times n} \to \mathbb{R}^{q \times m}$ *be a Borel-measurable matrix variate function. Assume that if* $X \sim N_{p,n}(M, \Sigma \otimes \Phi)$, *then the p.d.f. of* $W = \lambda(X)$ *is* $\ell^{\lambda}_{N_{p,n}(M,\Sigma\otimes\Phi)}(W)$. *Then, if* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ *with p.d.f.*

$g(X) = \int_0^\infty f_{N_{p,n}(M,z^{-1}\Sigma\otimes\Phi)}(X)\, w(z)\, dz,$ *the p.d.f. of* $W = \lambda(X)$ *is*

$$\ell(W) = \int_0^\infty \ell^{\lambda}_{N_{p,n}(M,z^{-1}\Sigma\otimes\Phi)}(W)\, w(z)\, dz. \tag{4.27}$$

If $g(X) = \int_0^\infty f_{N_{p,n}(M,z\Sigma\otimes\Phi)}(X)\, u(z)\, dz,$ *then the p.d.f. of* $W = \lambda(X)$ *is*

$$\ell(W) = \int_0^\infty \ell^{\lambda}_{N_{p,n}(M,z\Sigma\otimes\Phi)}(W)\, u(z)\, dz. \tag{4.28}$$

PROOF. In the proof of Theorem 4.1.4 if $dG(z)$ is replaced by $u(z)dz$, we obtain (4.27). In the same proof if we replace $dG(z)$ by $w(z)dz$ and $N_{p,n}(M,z\Sigma \otimes \Phi)$ by $N_{p,n}(M,z^{-1}\Sigma \otimes \Phi)$ we obtain (4.28). ■

COROLLARY 4.2.3.1. *Let* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ *with p.d.f.* $g(X) = \int_0^\infty f_{N_{p,n}(M,z^{-1}\Sigma\otimes\Phi)}(X)\, w(z)\, dz.$ *Let* C: $q \times m$, A: $q \times p$, *and* B: $n \times m$ *be constant matrices such that* $\text{rank}(A) = q$ *and* $\text{rank}(B) = m$. *Then, from Theorem 4.2.3 and Theorem 2.1.2 it follows that the p.d.f. of* $AXB + C$ *is*

$$g^*(X) = \int_0^\infty f_{N_{p,n}(AMB+C,z^{-1}(A\Sigma A')\otimes(B'\Phi B))}(X)\, w(z)\, dz.$$

If

$$g(X) = \int_0^\infty f_{N_{p,n}(M,z\Sigma\otimes\Phi)}(X)\, u(z)\, dz \tag{4.29}$$

then

$$g^*(X) = \int_0^\infty f_{N_{p,n}(AMB+C,z(A\Sigma A')\otimes(B'\Phi B))}(X)\, u(z)\, dz.$$

If X, M *and* Σ *are partitioned as* $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, $M = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}$, *and* $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, *where* X_1 is $q \times n$, M_1 is $q \times n$, *and* Σ_{11} is $q \times q$, $1 \le q < p$, *then the p.d.f. of* X_1 *is*

$$g_1(X_1) = \int_0^\infty f_{N_{p,n}(M_1,z^{-1}\Sigma_{11}\otimes\Phi)}(X_1)\, w(z)\, dz$$

and if (4.29) holds, then

$$g_1(X_1) = \int_0^\infty f_{N_{p,n}(M_1,z\Sigma_{11}\otimes\Phi)}(X_1)\, u(z)\, dz.$$

COROLLARY 4.2.3.2. *Let* $X \sim E_{p,n}(\mu e_n', \Sigma \otimes I_n, \psi)$ *with p.d.f.* $g(X)$ *where* $g(X) = \int_0^\infty f_{N_{p,n}(\mu e_n',z^{-1}\Sigma\otimes\Phi)}(X)\, w(z)\, dz$ *and* $\mu \in \mathbb{R}^p$. *Then,*

a) *the p.d.f. of* $\mathbf{y}_1 = \dfrac{Xe_n}{n}$ *is*

$$g_1(\mathbf{y}_1) = \int_0^\infty f_{N_p(\mu,z^{-1}\,\Sigma/n)}(\mathbf{y}_1)\, w(z)\, dz.$$

b) *the p.d.f. of* $Y_2 = X\left(I_n - \dfrac{e_n e_n'}{n}\right)X'$, *for* $p \le n-1$, *is*

$$g_2(Y_2) = \int_0^\infty f_{W_p(z^{-1}\Sigma,n-1)}(Y_2)\, w(z)\, dz.$$

and

c) *the p.d.f. of* $Y_3 = XX'$, *for* $p \le n$ *and* $\mu = \mathbf{o}$, *is*

$$g_3(Y_3) = \int_0^\infty f_{W_p(z^{-1}\Sigma,n)}(Y_3)\, w(z)\, dz.$$

If $g(x) = \int_0^\infty f_{N_{p,n}(\mu e_n', z\Sigma\otimes\Phi)}(x)\, u(z)\, dz$, *then*

$$g_1(y_1) = \int_0^\infty f_{N_{p,n}(\mu, z\,\Sigma/n)}(y_1)\, u(z)\, dz,$$

$$g_2(Y_2) = \int_0^\infty f_{W_p(z\Sigma,n-1)}(Y_2)\, u(z)\, dz,$$

and

$$g_3(Y_3) = \int_0^\infty f_{W_p(z\Sigma,n)}(Y_3)\, u(z)\, dz.$$

REMARK 4.2.4. It follows, from Example 4.2.1 and Corollary 4.2.3.1, that any submatrix of a random matrix with Cauchy distribution also has Cauchy distribution. Also any submatrix of a random matrix having t-distribution with m degrees of freedom has t-distribution with m degrees of freedom.

EXAMPLE 4.2.2. Let $X \sim E_{p,n}(\mu e_n', \Sigma \otimes I_n, \psi)$ have matrix variate t-distribution with m degrees of freedom. Then applying Corollary 4.2.3.2 with $w(z) = \dfrac{m\left(\frac{mz}{2}\right)^{\frac{m}{2}-1} e^{-\frac{mz}{2}}}{2\Gamma\left(\frac{m}{2}\right)}$, we see that

a) the p.d.f. of $y_1 = \frac{Xe_n}{n}$ is

$$g_1(y_1) = \frac{m^{\frac{m}{2}} n^{\frac{p}{2}} \Gamma\left(\frac{m+p}{2}\right)}{\pi^{\frac{p}{2}} \Gamma\left(\frac{m}{2}\right) |\Sigma|^{\frac{1}{2}}} (m + n(y_1 - \mu)' \Sigma^{-1}(y_1 - \mu))^{-\frac{m+p}{2}},$$

b) the p.d.f. of $Y_2 = X\left(I_n - \frac{e_n e_n'}{n}\right)X'$ is

$$g_2(Y_2) = \frac{m^{\frac{m}{2}} \Gamma\left(\frac{m+p(n-1)}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma_p\left(\frac{m-1}{2}\right)|\Sigma|^{\frac{n-1}{2}}} (m + \mathrm{tr}(\Sigma^{-1}Y_2))^{-\frac{m+p(n-1)}{2}} |Y_2|^{\frac{n-p-2}{2}},$$

c) the p.d.f. of $Y_3 = XX'$, if $\mu = \mathbf{o}$, is

$$g_3(Y_3) = \frac{m^{\frac{m}{2}} \Gamma\left(\frac{m+pn}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma_p\left(\frac{n}{2}\right)|\Sigma|^{\frac{n}{2}}} (m + \mathrm{tr}(\Sigma^{-1}Y_3))^{-\frac{m+pn}{2}} |Y_3|^{\frac{n-p-1}{2}}.$$

Here, $\Gamma_p(t) = \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^{p} \Gamma\left(t - \frac{i-1}{2}\right)$.

The mixture representation of the p.d.f. of a m.e.c. distribution makes it possible to derive monotone likelihood ratio (MLR) properties. To do this, we also need the following lemma (Karlin, 1956) given in Eaton (1972), p. B.2.

LEMMA 4.2.1. *Assume* $p(x,r)$ *and* $q(r,\theta)$ *are functions which have monotone likelihood ratios. Then* $g(x,\theta) = \int p(x,r)\, q(r,\theta)\, dr$ *also has monotone likelihood ratio.*

THEOREM 4.2.4. *Let* $x \sim E_n(\mu e_n, \sigma^2 I_n, \psi)$ *with p.d.f.* $g(x,\sigma)$ *where* $\mu \in \mathbb{R}$ *and* $g(x,\sigma) = \int_0^\infty f_{N_n(\mu e_n, z\sigma^2 I_n)}(x)\, u(z)\, dz$. *Assume*

$$u(c_1 r_1)\, u(c_2 r_2) \le u(c_1 r_2)\, u(c_2 r_1), \tag{4.30}$$

for $0 < c_1 < c_2$, $0 < r_1 < r_2$.

a) *If* $n > p$ *and* $g_1(y,\sigma)$ *denotes the p.d.f. of* $y = x'\left(I_n - \frac{e_n e_n'}{n}\right)x$, *then*

$$g_1(y_1,\sigma_1)\, g_1(y_2,\sigma_2) \ge g_1(y_1,\sigma_2)\, g_1(y_2,\sigma_1),$$

for $0 < \sigma_1 < \sigma_2$, $0 < y_1 < y_2$; *that is,* $g_1(y,\sigma)$ *has MLR.*

b) *If* $n \ge p$, $\mu = 0$, *and* $g_2(v,\sigma)$ *denotes the p.d.f. of* $v = x'x$, *then*

$$g_2(v_2,\sigma_1)\, g_2(v_2,\sigma_2) \ge g_2(v_2,\sigma_2)\, g_2(v_2,\sigma_1),$$

for $0 < \sigma_1 < \sigma_2$, $0 < v_1 < v_2$; *that is,* $g_2(v,\sigma)$ *has MLR.*

PROOF. a) We know, from Corollary 4.2.3.2, that

$$g_1(y,\sigma) = \int_0^\infty f_{W_1(z\sigma^2, n-1)}(y)\, u(z)\, dz.$$

Let $r = z\sigma^2$. Then, $z = \frac{r}{\sigma^2}$ and $J(z \to r) = \frac{1}{\sigma^2}$. Thus,

$$g_1(y,\sigma) = \int_0^\infty f_{W_1(r,n-1)}(y)\, u\left(\frac{r}{\sigma^2}\right)\frac{1}{\sigma^2}\, dr.$$

Let $p(y,r) = f_{W_1(r,n-1)}(y) = \dfrac{1}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)} \dfrac{y^{\frac{n-1}{2}-1} e^{-\frac{y}{2r}}}{r^{\frac{n-1}{2}}}$ and $q(r,\sigma) = u\left(\frac{r}{\sigma^2}\right)\frac{1}{\sigma^2}$.

It is easy to see that $p(y_1,r_1)\, p(y_2,r_2) \ge p(y_1,r_2)\, p(y_2,r_1)$, if $0 < r_1 < r_2$,

$0 < y_1 < y_2$. Thus $p(y,r)$ has MLR. It follows, from (4.30), that $q(r,\sigma)$ also has MLR. Using Lemma 4.2.1, we obtain the desired result.

b) It follows, from Corollary 4.2.3.2, that

$$g_2(y,\sigma) = \int_0^\infty f_{W_1(z\sigma^2,n)}(y)\, u(z)\, dz.$$

Then, proceeding in a similar way as in the proof of part (a), we find that $g_2(v,\sigma)$ has MLR. ■

The following theorem was given by Chu (1973) for the vector variate case.

THEOREM 4.2.5. *Let* λ: $\mathbb{R}^{p\times n} \to \mathbb{R}^{q\times m}$ *be a Borel-measurable matrix variate function. Assume that if* $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$ *then* $\mathcal{E}(\lambda(X))$ *exists and it is denoted by* $\mathcal{E}_{N_{p,n}(M,\Sigma\otimes\Phi)}(\lambda(X))$. *Then, if* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ *with p.d.f.*

$g(X) = \int_0^\infty f_{N_{p,n}(M,z^{-1}\Sigma\otimes\Phi)}(X)\, w(z)\, dz$ *such that* $\mathcal{E}(\lambda(X))$ *exists and it is denoted by* $\mathcal{E}_{E_{p,n}(M,\Sigma\otimes\Phi,\psi)}(\lambda(X))$, *we have*

$$\mathcal{E}_{E_{p,n}(M,\Sigma\otimes\Phi,\psi)}(\lambda(X)) = \int_0^\infty \mathcal{E}_{N_{p,n}(M,z^{-1}\Sigma\otimes\Phi)}(\lambda(X))\, w(z)dz. \tag{4.31}$$

If $g(X) = \int_0^\infty f_{N_{p,n}(M,z\Sigma\otimes\Phi)}(X)\, u(z)\, dz$ *and* $\mathcal{E}(\lambda(X))$ *exists, then*

$$\mathcal{E}_{E_{p,n}(M,\Sigma\otimes\Phi,\psi)}(\lambda(X)) = \int_0^\infty \mathcal{E}_{N_{p,n}(M,z\Sigma\otimes\Phi)}(\lambda(X))\, u(z)dz. \tag{4.32}$$

PROOF. In the proof of Theorem 4.1.5 if $dG(z)$ is replaced by $u(z)\, dz$, we obtain (4.31). In the same proof if we replace $dG(z)$ by $w(z)\, dz$ and $N_{p,n}(M,z\Sigma \otimes \Phi)$ by $N_{p,n}(M,z^{-1}\Sigma \otimes \Phi)$ we obtain (4.32). ■

COROLLARY 4.2.5.1. *With the notation of Theorem 4.2.4, if* Cov(X) *exists, then* $\mathrm{Cov}(X) = \left(\int_0^\infty \frac{w(s)}{s}\, ds \right) \Sigma \otimes \Phi$ *and also* $\mathrm{Cov}(X) = \left(\int_0^\infty u(s)\, s\, ds \right) \Sigma \otimes \Phi$.

CHAPTER 5

QUADRATIC FORMS AND OTHER FUNCTIONS OF ELLIPTICALLY CONTOURED MATRICES

5.1. EXTENSION OF COCHRAN'S THEOREM TO MULTIVARIATE ELLIPTICALLY CONTOURED DISTRIBUTIONS

Anderson and Fang (1987) studied how results, similar to Cochran's theorem can be derived for m.e.c. distributions. This section presents their results. Results from Anderson and Fang (1982b) are also used.

We will need the following lemma.

LEMMA 5.1.1. *Let* X: $p \times n$ *be a random matrix with p.d.f.* $f(XX')$. *Let* $A = XX'$, *then the p.d.f. of* A *is*

$$\frac{\pi^{\frac{pn}{2}}}{\Gamma_p\left(\frac{n}{2}\right)} |A|^{\frac{n-p-1}{2}} f(A), \quad A > O,$$

where $\Gamma_p(t) = \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^{p} \Gamma\left(t - \frac{i-1}{2}\right)$.

PROOF. See Anderson (1984), p. 533. ■

The next lemma generalizes the result of Lemma 5.1.1.

LEMMA 5.1.2. *Let* X *be a random* $p \times n$ *matrix, and write* $X = (X_1, X_2, \ldots, X_m)$ *where* X_i is $p \times n_i$, $i = 1, \ldots, m$. *Assume* X *has the p.d.f.* $p(X) = f(X_1X_1', X_2X_2', \ldots, X_mX_m')$. *Further let* $W_i = X_iX_i'$, $i = 1, \ldots, m$. *Then, the p.d.f. of* $(W_1, W_2, \ldots, W_m)$ *is*

$$\frac{\pi^{\frac{pn}{2}}}{\prod_{i=1}^{m} \Gamma_p\left(\frac{n_i}{2}\right)} \prod_{i=1}^{n} |W_i|^{\frac{n_i-p-1}{2}} f(W_1,W_2,\ldots,W_m), \quad W_i > O, \quad i = 1,\ldots,m. \tag{5.1}$$

PROOF. We prove, by induction, that the p.d.f. of $(W_1,\ldots,W_k,X_{k+1},\ldots,X_m)$ is

$$\frac{\pi^{\frac{p}{2}\sum_{i=1}^{k} n_i}}{\prod_{i=1}^{k} \Gamma_p\left(\frac{n_i}{2}\right)} \prod_{i=1}^{k} |W_i|^{\frac{n_i-p-1}{2}} p(X_1,\ldots,X_k,X_{k+1},\ldots,X_m)\,|_{X_1X_1'=W_1,\ldots,X_kX_k'=W_k},$$

$$k = 1,\ldots,m. \tag{5.2}$$

In the proof, $p(Y)$ will denote the p.d.f. of any random matrix Y.

If $k = 1$, we can write

$$p(W_1,X_2,\ldots,X_m) = p(W_1 \mid X_2,\ldots,X_m)\, p(X_2,\ldots,X_m)$$

$$= \frac{\pi^{\frac{pn_1}{2}}}{\Gamma_p\left(\frac{n_1}{2}\right)} |W_1|^{\frac{n_1-p-1}{2}} p(X_1 \mid X_2,\ldots,X_m)\,|_{X_1X_1'=W_1} p(X_2,\ldots,X_m)$$

$$= \frac{\pi^{\frac{pn_1}{2}}}{\Gamma_p\left(\frac{n_1}{2}\right)} |W_1|^{\frac{n_1-p-1}{2}} p(X_1,X_2,\ldots,X_m)\,|_{X_1X_1'=W_1},$$

where we used Lemma 5.1.1.

Now, assume the statement is true for $k = \ell < m$. Then, for $k = \ell + 1$, we get

$$
\begin{aligned}
&p(W_1,\ldots,W_{\ell+1},X_{\ell+2},\ldots,X_m)\\
&= p(W_{\ell+1} \mid W_1,\ldots,W_\ell,X_{\ell+2},\ldots,X_m)\, p(W_1,\ldots,W_\ell,X_{\ell+2},\ldots,X_m)\\
&= \frac{\pi^{\frac{pn_{\ell+1}}{2}}}{\Gamma_p\left(\frac{n_{\ell+1}}{2}\right)} |W_{\ell+1}|^{\frac{n_{\ell+1}-p-1}{2}}.
\end{aligned}
$$

$$
p(X_{\ell+1} \mid W_1,\ldots,W_\ell,X_{\ell+2},\ldots,X_m)\big|_{X_{\ell+1}X'_{\ell+1}=W_{\ell+1}}\; p(W_1,\ldots,W_\ell,X_{\ell+2},\ldots,X_m)
$$

$$
= \frac{\pi^{\frac{pn_{\ell+1}}{2}}}{\Gamma_p\left(\frac{n_{\ell+1}}{2}\right)} |W_{\ell+1}|^{\frac{n_{\ell+1}-p-1}{2}}
$$

$$
\cdot\, p(W_1,\ldots,W_\ell,X_{\ell+1},X_{\ell+2},\ldots,X_m)\big|_{X_{\ell+1}X'_{\ell+1}=W_{\ell+1}}
$$

$$
= \frac{\pi^{\frac{pn_{\ell+1}}{2}}}{\Gamma_p\left(\frac{n_{\ell+1}}{2}\right)} |W_{\ell+1}|^{\frac{n_{\ell+1}-p-1}{2}} \frac{\pi^{\frac{p}{2}\sum_{i=1}^{\ell} n_i}}{\prod_{i=1}^{\ell}\Gamma_p\left(\frac{n_i}{2}\right)} \prod_{i=1}^{\ell} |W_i|^{\frac{n_i-p-1}{2}}
$$

$$
\cdot\, p(X_1,\ldots,X_\ell,X_{\ell+1},X_{\ell+2},\ldots,X_m)\big|_{X_1X'_1=W_1,\ldots,X_{\ell+1}X'_{\ell+1}=W_{\ell+1}}
$$

$$
= \frac{\pi^{\frac{p}{2}\sum_{i=1}^{\ell+1} n_i}}{\prod_{i=1}^{\ell+1}\Gamma_p\left(\frac{n_i}{2}\right)} \prod_{i=1}^{\ell+1} |W_i|^{\frac{n_i-p-1}{2}}
$$

$$
\cdot\, p(X_1,\ldots,X_{\ell+1},X_{\ell+2},\ldots,X_m)\big|_{X_1X'_1=W_1,\ldots,X_{\ell+1}X'_{\ell+1}=W_{\ell+1}},
$$

where we used Lemma 5.1.1 and the induction hypothesis. Taking $k = m$ in (5.2) we obtain (5.1). ■

DEFINITION 5.1.1. Let $X \sim E_{p,n}(O,\Sigma \otimes I_n,\psi)$, $\Sigma > O$. Partition X as $X = (X_1,X_2,\ldots,X_m)$, where X_i is $p \times n_i$, $i = 1,\ldots,m$. Then, $G_{p,m}\left(\Sigma, \frac{n_1}{2}, \frac{n_2}{2},\ldots, \frac{n_m}{2},\psi\right)$ denotes the distribution of $(X_1X_1',X_2X_2',\ldots,X_mX_m')$.

REMARK 5.1.1. If in Definition 5.1.1, $\Sigma = I_p$, we also use the notation $G_{p,m}\left(\frac{n_1}{2}, \frac{n_2}{2},\ldots, \frac{n_m}{2},\psi\right)$; that is, I_p can be dropped from the notation.

REMARK 5.1.2. Definition 5.1.1 generalizes the Wishart distribution. In fact, if $m = 1$ and $\psi(z) = \exp\left(-\frac{z}{2}\right)$, then $G_{p,1}\left(\Sigma, \frac{n}{2}, \psi\right)$ is the same as $W_p(\Sigma,n)$.

THEOREM 5.1.1. *Let* $X \sim E_{p,n}(O,\Sigma \otimes I_n,\psi)$, $\Sigma > O$ *and* rAU *be the stochastic representation of* X. *Partition* X *as* $X = (X_1,X_2,\ldots,X_m)$, *where* X_i *is* $p \times n_i$, $i = 1,\ldots,m$. *Then,*

$$(X_1X_1',X_2X_2',\ldots,X_mX_m') \approx r^2A(z_1V_1,z_2V_2,\ldots,z_mV_m)A', \tag{5.3}$$

where $(z_1,z_2,\ldots,z_{m-1}) \approx D\left(\frac{n_1}{2}, \frac{n_2}{2},\ldots, \frac{n_{m-1}}{2}; \frac{n_m}{2}\right)$, $\sum_{i=1}^{m} z_i = 1$, $V_i = U_iU_i'$ *with* $\text{vec}(U_i')$ *uniformly distributed on* S_{pn_i}, *and* r, $V_1,V_2,\ldots,V_m$, $(z_1,z_2,\ldots,z_m)$ *are independent.*

PROOF. From Theorem 2.5.6, it follows that $(X_1,X_2,\ldots,X_m) \approx rA(\sqrt{z_1}\,U_1,\sqrt{z_2}\,U_2,\ldots,\sqrt{z_m}\,U_m)$, from which (5.3) follows immediately. ■

THEOREM 5.1.2. *Let* $(W_1,W_2,\ldots,W_m) \sim G_{p,m}\left(\Sigma, \frac{n_1}{2}, \frac{n_2}{2},\ldots, \frac{n_m}{2}, \psi\right)$, *where* W_i *is* $p \times p$, $i = 1,\ldots,m$. *Then, for* $1 \leq \ell \leq m$,

$$(W_1,W_2,\ldots,W_\ell) \sim G_{p,\ell}\left(\Sigma,\frac{n_1}{2},\frac{n_2}{2},\ldots,\frac{n_\ell}{2},\psi\right).$$

PROOF. Define $X \sim E_{p,n}(O,\Sigma \otimes I_n,\psi)$ and partition X as $X = (X_1,X_2,\ldots,X_m)$, where X_i is $p \times n_i$ dimensional, $i = 1,\ldots,n$. Then, by Definition 5.1.1 we have

$$(X_1X_1',X_2X_2',\ldots,X_mX_m') \approx (W_1,W_2,\ldots,W_\ell).$$

Hence, $(X_1X_1',X_2X_2',\ldots,X_\ell X_\ell') \approx (W_1,W_2,\ldots,W_\ell)$. Let $Y = (X_1,X_2,\ldots,X_\ell)$. Then, $Y \sim E_{p,n^*}(O,\Sigma \otimes I_{n^*},\psi)$ with $n^* = \sum_{i=1}^{\ell} n_i$. Therefore, $(X_1X_1',X_2X_2',\ldots,X_\ell X_\ell') \sim G_{p,\ell}\left(\Sigma,\frac{n_1}{2},\frac{n_2}{2},\ldots,\frac{n_\ell}{2},\psi\right)$, which completes the proof.■

THEOREM 5.1.3. *Let* $(W_1,W_2,\ldots,W_m) \sim G_{p,m}\left(\Sigma,\frac{n_1}{2},\frac{n_2}{2},\ldots,\frac{n_m}{2},\psi\right)$, *where* $m > 1$, W_i *is* $p \times p$, $i = 1,\ldots,m$. *Then,* $(W_1 + W_2,W_3,\ldots,W_m) \sim G_{p,m-1}\left(\Sigma,\frac{n_1+n_2}{2},\frac{n_3}{2},\ldots,\frac{n_m}{2},\psi\right)$.

PROOF. Let X be defined as in the proof of Theorem 5.1.2. Define $X_0 = (X_1,X_2)$ and $Y = (X_0,X_3,\ldots,X_m)$. Then,

$$\begin{aligned}(X_0X_0',X_3X_3',\ldots,X_mX_m') &= (X_1X_1' + X_2X_2',X_3X_3',\ldots,X_mX_m') \\ &\approx (W_1 + W_2,W_3,\ldots,W_m).\end{aligned}$$

We also have

$$Y = (X_0,X_3,\ldots,X_m) \sim E_{p,m}(O,\Sigma \otimes I_n,\psi).$$

Hence, $(X_0X_0',X_3X_3',\ldots,X_mX_m') \sim G_{p,m-1}\left(\Sigma,\frac{n_1+n_2}{2},\frac{n_3}{2},\ldots,\frac{n_m}{2},\psi\right)$ which completes the proof. ■

THEOREM 5.1.4. *Let* $X \sim E_{p,n}(O, \Sigma \otimes I_n, \psi)$ *with* $\Sigma > O$ *and* $P(X = O) = 0$ and stochastic representation $X \approx r\Sigma^{\frac{1}{2}}U$. *Partition* X *as* $X = (X_1, X_2, \ldots, X_m)$, *where* X_i *is* $p \times n_i$, $i = 1,\ldots,m$, $1 < m \le n$, $p \le n_i$, $i = 1,2,\ldots,m-1$, $1 \le n_m$. *Let* $W_i = X_i X_i'$, $i = 1,\ldots,m-1$, *then the p.d.f. of* $(W_1, W_2, \ldots, W_{m-1})$ *is given by*

$$p(W_1,W_2,\ldots,W_{m-1}) = \frac{\Gamma\left(\frac{pn}{2}\right) |\Sigma|^{-\frac{n-n_m}{2}}}{\Gamma\left(\frac{pn_m}{2}\right) \prod_{i=1}^{m-1} \Gamma_p\left(\frac{pn_i}{2}\right)} \prod_{i=1}^{m-1} |W_i|^{\frac{n_i-p-1}{2}}$$

$$\int\limits_{\left(\mathrm{tr}\left(\Sigma^{-1}\sum_{i=1}^{m-1} W_i\right)\right)^{\frac{1}{2}}}^{\infty} r^{2-pn}\left(r^2 - \mathrm{tr}\left(\Sigma^{-1}\sum_{i=1}^{m-1} W_i\right)\right)^{\frac{pn_m}{2}-1} dF(r), \qquad (5.4)$$

where $W_i > O$, $i = 1,\ldots,m$, *and* $F(r)$ *is the distribution function of* r.

PROOF. Let $Y = \Sigma^{-\frac{1}{2}}X$, $Y_i = \Sigma^{-\frac{1}{2}}X_i$, $i = 1,2,\ldots,m$, and $V_i = Y_i Y_i'$, $i = 1,2,\ldots,m-1$. Then $Y \sim E_{p,n}(O, I_p \otimes I_n, \psi)$. From Theorem 3.1.6, it follows that the density of $Y = (Y_1, Y_2, \ldots, Y_{m-1})$ is

$$f(Y_1,Y_2,\ldots,Y_{m-1}) = \frac{\Gamma\left(\frac{np}{2}\right)}{\pi^{\frac{p(n-n_m)}{2}} \Gamma\left(\frac{pn_m}{2}\right)} \int\limits_{\left(\mathrm{tr}\left(\sum_{i=1}^{m-1} Y_i Y_i'\right)\right)^{\frac{1}{2}}}^{\infty} r^{2-pn}$$

$$\left(r^2 - \mathrm{tr}\left(\sum_{i=1}^{m-1} Y_i Y_i'\right)\right)^{\frac{pn_m}{2}-1} dF(r).$$

Then, Lemma 5.1.2 gives the p.d.f. of $(V_1, V_2, \ldots, V_{m-1})$ as

$$p(V_1,V_2,\dots,V_{m-1}) = \frac{\Gamma\left(\frac{pn}{2}\right)}{\Gamma\left(\frac{pn_m}{2}\right)\prod_{i=1}^{m-1}\Gamma_p\left(\frac{pn_i}{2}\right)} \prod_{i=1}^{m-1} |V_i|^{\frac{n_i-p-1}{2}}$$

$$\int_{\left(\mathrm{tr}\left(\sum_{i=1}^{m-1} V_i\right)\right)^{\frac{1}{2}}}^{\infty} r^{2-pn}\left(r^2 - \mathrm{tr}\left(\sum_{i=1}^{m-1} V_i\right)\right)^{\frac{pn_m}{2}-1} dF(r). \tag{5.5}$$

Since, $(V_1,V_2,\dots,V_{m-1}) = \Sigma^{-\frac{1}{2}}(W_1,W_2,\dots,W_{m-1})\Sigma^{-\frac{1}{2}}$ and $J((V_1,V_2,\dots,V_{m-1}) \to (W_1,W_2,\dots,W_{m-1})) = |\Sigma|^{-\frac{(p+1)(m-1)}{2}}$, from (5.5) we get

$$p(W_1,W_2,\dots,W_{m-1}) = \frac{\Gamma\left(\frac{pn}{2}\right)}{\Gamma\left(\frac{pn_m}{2}\right)\prod_{i=1}^{m-1}\Gamma_p\left(\frac{pn_i}{2}\right)} \prod_{i=1}^{m-1} |\Sigma|^{-\frac{n_i-p-1}{2}} \prod_{i=1}^{m-1} |W_i|^{\frac{n_i-p-1}{2}}$$

$$\int_{\left(\mathrm{tr}\left(\Sigma^{-1}\sum_{i=1}^{m-1} W_i\right)\right)^{\frac{1}{2}}}^{\infty} \left(r^2 - \mathrm{tr}\left(\Sigma^{-1}\sum_{i=1}^{m-1} W_i\right)\right)^{\frac{pn_m}{2}-1} |\Sigma|^{-\frac{(p+1)(m-1)}{2}} dF(r)$$

and since $\left(\prod_{i=1}^{m-1} |\Sigma|^{-\frac{n_i-p-1}{2}}\right) |\Sigma|^{-\frac{(p+1)(m-1)}{2}} = |\Sigma|^{-\frac{1}{2}\sum_{i=1}^{m-1} n_i}$, we obtain (5.4). ■

If X is absolutely continuous, we obtain the following result.

THEOREM 5.1.5. *Let* $X \sim E_{p,n}(O,\Sigma \otimes I_n,\psi)$ *have the p.d.f.*

$$f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}}} h(\mathrm{tr}(X'\Sigma^{-1}X)).$$

Partition X *as* $X = (X_1, X_2, \dots, X_m)$, *where* X_i *is* $p \times n_i$, $i = 1,\dots,m$, $1 \le m \le p$. *Let* $W_i = X_iX_i'$, $i = 1,\dots,m$, *then the p.d.f. of* $(W_1, W_2, \dots, W_m)$ *is*

$$p(W_1, W_2, \dots, W_m) = \frac{\pi^{\frac{pn}{2}} |\Sigma|^{-\frac{n}{2}}}{\prod_{i=1}^{m} \Gamma_p\left(\frac{n_i}{2}\right)} \prod_{i=1}^{m} |W_i|^{\frac{n_i-p-1}{2}} h\left(\mathrm{tr}\, \Sigma^{-1} \sum_{i=1}^{m} W_i\right),$$

$$W_i > O, \quad i = 1,\dots,m. \tag{5.6}$$

PROOF. Since the p.d.f. of $X = (X_1, X_2, \dots, X_m)$ is $|\Sigma|^{-\frac{n}{2}} h\left(\mathrm{tr}\, \Sigma^{-1} \sum_{i=1}^{m} X_iX_i'\right)$, from Lemma 5.1.2, we obtain (5.6). ∎

COROLLARY 5.1.5.1. *Let* $X \sim E_{p,n}(O, \Sigma \otimes I_n, \psi)$ *with the p.d.f.*

$$f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}}} h(\mathrm{tr}(X'\Sigma^{-1}X)).$$

Then, the p.d.f. of $A = XX'$ *is given by*

$$p(A) = \frac{\pi^{\frac{pn}{2}} |\Sigma|^{-\frac{n}{2}}}{\Gamma_p\left(\frac{n}{2}\right)} |A|^{\frac{n-p-1}{2}} h(\mathrm{tr}(\Sigma^{-1}A)), \quad A > O.$$

LEMMA 5.1.3. *Let* x *and* z *be independent, and* y *and* z *be independent one-dimensional random variables, such that* $P(x > 0) = P(y > 0) = P(z > 0) = 1$ *and* $xz \approx yz$. *Assume one of the following conditions holds:*

i) $\phi_{\log z}(t) \neq 0$ *almost everywhere*

ii) $P(x < 1) = 1$.

Then, $x \approx y$.

PROOF. Note that $xz \approx yz$ is equivalent to $\log x + \log z \approx \log y + \log z$, which is again equivalent to

$$\phi_{\log x}(t)\, \phi_{\log z}(t) = \phi_{\log y}(t)\, \phi_{\log z}(t), \tag{5.7}$$

where $\phi_x(t)$ denotes the characteristic function of x.

i) Since $\phi_{\log z}(t) \neq 0$ almost everywhere, and the characteristic functions are continuous, we get

$$\phi_{\log x}(t) = \phi_{\log y}(t).$$

Hence, $x \approx y$.

ii) Since $\phi_{\log z}(0) = 1$ and the characteristic functions are continuous, there exists $\delta > 0$, such that $\phi_{\log z}(t) \neq 0$ for $t \in (-\delta,\delta)$.

Then, from (5.7) we get

$$\phi_{\log x}(t) = \phi_{\log y}(t) \text{ for } t \in (-\delta,\delta). \tag{5.8}$$

Since $P(x < 1) = 1$, we have $P(\log x < 0) = 1$. However, $P(\log x < 0) = 1$, together with (5.8), implies that $\log x \approx \log y$ (see Marcinkiewicz, 1938). Thus, we get $x \approx y$. ■

THEOREM 5.1.6. *Let* $(W_1,W_2,\dots,W_m) \sim G_{p,m}\left(\Sigma, \frac{n_1}{2}, \frac{n_2}{2},\dots, \frac{n_m}{2}, \psi\right)$, *where* n_i *is positive integer and* W_i *is* $p \times p$, $i = 1,2,\dots,m$ *and* $\Sigma > O$. *Let* $\mathbf{v}$ *be a p-dimensional constant nonzero vector. Then,*

$$(\mathbf{v}'W_1\mathbf{v},\mathbf{v}'W_2\mathbf{v},\dots,\mathbf{v}'W_m\mathbf{v}) \sim G_{1,m}\left(\frac{n_1}{2}, \frac{n_2}{2},\dots, \frac{n_m}{2}, \psi^*\right), \tag{5.9}$$

where $\psi^*(z) = \psi(\mathbf{v}'\Sigma\mathbf{v}z)$.

PROOF. Let $(X_1,X_2,\dots,X_m) \sim E_{p,n}(O,\Sigma \otimes I_n,\psi)$, with $n = \sum_{i=1}^{m} n_i$ and X_i is $p \times n_i$, $i = 1,2,\dots,m$. Then, by definition,

$(X_1X_1', X_2X_2', \dots, X_mX_m') \approx (W_1, W_2, \dots, W_m)$. Hence,

$$(\mathbf{v}'X_1X_1'\mathbf{v}, \mathbf{v}'X_2X_2'\mathbf{v}, \dots, \mathbf{v}'X_mX_m'\mathbf{v}) \approx (\mathbf{v}'W_1\mathbf{v}, \mathbf{v}'W_2\mathbf{v}, \dots, \mathbf{v}'W_m\mathbf{v}). \quad (5.10)$$

Define $\mathbf{y}_i' = \mathbf{v}'X_i$, $i = 1,2,\dots,m$. Then, $\mathbf{y}_i$ is $n_i \times 1$ and $(\mathbf{y}_1', \mathbf{y}_2', \dots, \mathbf{y}_m') = \mathbf{v}'(X_1, X_2, \dots, X_m)$, and hence,

$$(\mathbf{y}_1', \mathbf{y}_2', \dots, \mathbf{y}_m') \sim E_{1,n}(O, \mathbf{v}'\Sigma\mathbf{v} \otimes I_n, \psi) = E_{1,n}(O, I_n, \psi^*)$$

with $\psi^*(z) = \psi(\mathbf{v}'\Sigma\mathbf{v}z)$. Therefore,

$$(\mathbf{y}_1'\mathbf{y}_1, \mathbf{y}_2'\mathbf{y}_2, \dots, \mathbf{y}_m'\mathbf{y}_m) \sim G_{1,m}\left(\frac{n_1}{2}, \frac{n_2}{2}, \dots, \frac{n_m}{2}, \psi^*\right). \quad (5.11)$$

Since, $\mathbf{y}_i'\mathbf{y}_i = \mathbf{v}'X_iX_i'\mathbf{v}$, $i = 1,2,\dots,m$, (5.10) and (5.11) give (5.9). ■

Now, we derive some result which can be regarded as the generalizations of Cochran's theorem for normal variables to the m.e.c. distribution.

THEOREM 5.1.7. *Let* $X \sim E_{p,n}(O, \Sigma \otimes I_n, \psi)$, $\Sigma > O$ *and assume there exists a p-dimensional constant vector* $\mathbf{v}$ *such that* $P(\mathbf{v}'\Sigma^{-\frac{1}{2}}X = \mathbf{o}) = 0$. *Let* A *be an* $n \times n$ *symmetric matrix and* $k \le n$ *a positive integer. Then,*

$$XAX' \sim G_{p,1}\left(\Sigma, \frac{k}{2}, \psi\right) \quad (5.12)$$

if and only if $A^2 = A$ *and* $\mathrm{rk}(A) = k$.

PROOF. It is enough to consider the case $\Sigma = I_p$ because otherwise we can define $Y = \Sigma^{-\frac{1}{2}}X$, and $XAX' \sim G_{p,1}\left(\Sigma, \frac{k}{2}, \psi\right)$ is equivalent to $YAY' \sim G_{p,1}\left(\frac{k}{2}, \psi\right)$.

First, assume $A^2 = A$ and $\mathrm{rk}(A) = k$. Then, using Theorem 1.3.12, we can write

$$A = G\begin{pmatrix} I_k & O \\ O & O \end{pmatrix}G',$$

where $G \in O(n)$ and O's denote zero matrices of appropriate dimensions.

Define $n \times k$ matrix $C = \begin{pmatrix} I_k \\ O \end{pmatrix}$ and let $Y = XGC$. Then, $Y \sim E_{p,k}(O, I_p \otimes I_k, \psi)$ and

$$XAX' = YY'. \tag{5.13}$$

From Definition 5.1.1, we have

$$YY' \sim G_{p,1}\left(\frac{k}{2}, \psi\right). \tag{5.14}$$

From (5.13) and (5.14), we obtain (5.12).

On the other hand, assume (5.12) holds, and define $\mathbf{y} = X'\mathbf{v}$. Then, $\mathbf{y} \sim E_n(\mathbf{o}, I_n, \psi^*)$, where $\psi^*(t) = \psi(\mathbf{v}'\mathbf{v}t)$. Moreover, $P(\mathbf{y} = \mathbf{o}) = 0$. From Theorem 5.1.2, we get

$$\mathbf{y}'A\mathbf{y} \sim G_{1,1}\left(\frac{k}{2}, \psi^*\right). \tag{5.15}$$

Let $\mathbf{y} \approx r\mathbf{u}$ be the stochastic representation of $\mathbf{y}$. Then,

$$\mathbf{y}'A\mathbf{y} \approx r^2\mathbf{u}'A\mathbf{u}. \tag{5.16}$$

Let $\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}$, where $\mathbf{y}_1$ is k-dimensional. Then, from Corollary 2.5.6.1, we get

$$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} \approx \begin{pmatrix} r\sqrt{w}\,\mathbf{u}_1 \\ r\sqrt{1-w}\,\mathbf{u}_2 \end{pmatrix}, \tag{5.17}$$

where $\mathbf{r}$, $\mathbf{w}$, $\mathbf{u}_1$ and $\mathbf{u}_2$ are independent, $\mathbf{u}_1$ is uniformly distributed on S_k, $\mathbf{u}_2$ is uniformly distributed on S_{n-k}, and $w \sim B\left(\frac{k}{2}, \frac{n-k}{2}\right)$. Since $\mathbf{y}_1 \sim E_k(\mathbf{o}, I_k, \psi^*)$, we get

$$\mathbf{y}_1'\mathbf{y}_1 \sim G_{1,1}\left(\frac{k}{2}, \psi^*\right). \tag{5.18}$$

From (5.17), we obtain

$$\mathbf{y}_1'\mathbf{y}_1 \approx r^2 w \, \mathbf{u}_1'\mathbf{u}_1 = r^2 w. \tag{5.19}$$

From (5.15), (5.16), (5.18), and (5.19), we get

$$r^2 \mathbf{u}'A\mathbf{u} \approx r^2 w. \tag{5.20}$$

Since $P(0 < w < 1) = 1$ and $P(r^2 > 0) = 1$, we have $P(r^2 w > 0) = 1$. Therefore, (5.20) implies $P(\mathbf{u}'A\mathbf{u} > 0) = 1$. Using (ii) of Lemma 5.1.3, from (5.20) we obtain $\mathbf{u}'A\mathbf{u} \approx w$. Thus,

$$\mathbf{u}'A\mathbf{u} \sim B\left(\frac{k}{2}, \frac{n-k}{2}\right). \tag{5.21}$$

Define $\mathbf{z} \sim N_n(0, I_n)$. Then, from Theorem 2.5.4 it follows that $\frac{\mathbf{z}}{\|\mathbf{z}\|} \approx \mathbf{u}$, and from (5.21) we get

$$\frac{\mathbf{z}'A\mathbf{z}}{\|\mathbf{z}\|^2} \sim B\left(\frac{k}{2}, \frac{n-k}{2}\right). \tag{5.22}$$

Now, $A = GDG'$, where $G \in O(n)$ and $D = \begin{pmatrix} I_k & O \\ O & O \end{pmatrix}$, where O's denote zero matrices of appropriate dimensions. Let $\mathbf{t} = G'\mathbf{z}$. Then, $\mathbf{t} \sim N_n(\mathbf{0}, I_n)$ and $\frac{\mathbf{z}'A\mathbf{z}}{\|\mathbf{z}\|^2} = \frac{\mathbf{t}'D\mathbf{t}}{\|\mathbf{t}\|^2} \sim B\left(\frac{k}{2}, \frac{n-k}{2}\right)$. However $\mathbf{z} \approx \mathbf{t}$ and so we get

$$\frac{z'Dz}{\|z\|^2} \approx B\left(\frac{k}{2}, \frac{n-k}{2}\right). \tag{5.23}$$

From (5.22) and (5.23), we get

$$\frac{z'Az}{\|z\|^2} \approx \frac{z'Dz}{\|z\|^2}. \tag{5.24}$$

Now, $z'Az = \|z\|^2 \frac{z'Az}{\|z\|^2}$ with $\|z\|^2$ and $\frac{z'Az}{\|z\|^2}$ being independent. Moreover, $z'Dz = \|z\|^2 \frac{z'Dz}{\|z\|^2}$, with $\|z\|^2$ and $\frac{z'Dz}{\|z\|^2}$ being independent. Therefore, from (5.23), we get $z'Az \sim z'Dz$. Since, $z'Dz \sim \chi^2_k$, we get

$$z'Az \sim \chi^2_k. \tag{5.25}$$

Now, (5.25) implies that $A^2 = A$, and $rk(A) = k$. ■

Theorem 5.1.7 can be generalized in the following way.

THEOREM 5.1.8. *Let* $X \sim E_{p,n}(O, \Sigma \otimes I_n, \psi)$, $\Sigma > O$ *and assume there exists a p-dimensional constant vector* $\mathbf{v}$, *such that* $P(\mathbf{v}'\Sigma^{-\frac{1}{2}} X = \mathbf{o}) = 0$. *Let* $A_1, A_2, \ldots, A_m$ be $n \times n$ *symmetric matrices and* $k_1, k_2, \ldots, k_m$ *positive integers with* $\sum_{i=1}^{m} k_i \leq n$. *Then,*

$$(XA_1X', XA_2X', \ldots, XA_mX') \sim G_{p,m}\left(\Sigma, \frac{k_1}{2}, \frac{k_2}{2}, \ldots, \frac{k_m}{2}, \psi\right) \tag{5.26}$$

if and only if

$$rk(A_i) = k_i, \ i = 1, \ldots, m, \tag{5.27}$$

$$A_i^2 = A_i, \ \ i = 1,\dots,m \tag{5.28}$$

and

$$A_iA_j = O, \ \ i \neq j, \ \ i,j = 1,\dots,m. \tag{5.29}$$

PROOF. As in Theorem 5.1.7, it is enough to consider the case $\Sigma = I_p$.

First, assume (5.26), (5.27) and (5.28) are satisfied. Then, from Theorem 1.3.14, there exists $G \in O(n)$ such that

$$G'A_1G = \begin{pmatrix} I_{k_1} & O \\ O & O \end{pmatrix}, \ G'A_2G = \begin{pmatrix} O & O & O \\ O & I_{k_2} & O \\ O & O & O \end{pmatrix}$$

$$\dots, \ G'A_mG = \begin{pmatrix} O & O & O \\ O & I_{k_m} & O \\ O & O & O \end{pmatrix}.$$

Let $k = \sum_{i=1}^{m} k_i$ and define the $n \times k$ matrix $C = \begin{pmatrix} I_k \\ O \end{pmatrix}$. Moreover, define the $k \times k$ matrices $C_i = \begin{pmatrix} O & O & O \\ O & I_{k_i} & O \\ O & O & O \end{pmatrix}$, $i = 1,\dots,m$. Then, $G'A_iG = CC_iC'$, $i = 1,\dots,m$.

Define $Y = XGC$. Then, $Y \sim E_{p,k}(O, I_p \otimes I_k, \psi)$ and

$$XA_iX' = YC_iY', \ \ i = 1,\dots,m. \tag{5.30}$$

Partition Y into $Y = (Y_1, Y_2, \dots, Y_m)$ where Y_i is $p \times k_i$, $i = 1,\dots,m$. Then,

$$YC_iY' = Y_iY_i', \ \ i = 1,\dots,m \tag{5.31}$$

and by Definition 4.1, we get

$$(Y_1Y_1',Y_2Y_2',\dots,Y_mY_m') \sim G_{p,m}\left(\frac{k_1}{2},\frac{k_2}{2},\dots,\frac{k_m}{2},\psi\right). \tag{5.32}$$

From (5.30), (5.31), and (5.32) we obtain (5.26).

Next, assume (5.26) holds. Then, it follows from Theorem 5.1.2, that

$$XA_1X' \sim G_{p,1}\left(\frac{k_1}{2},\psi\right), \quad i = 1,\dots,m.$$

Using Theorem 5.1.7, we get $rk(A_i) = k_i$ and $A_i^2 = A_i$. It also follows from Theorem 5.1.2, that

$$(XA_iX',XA_jX') \sim G_{p,2}\left(\frac{k_i}{2},\frac{k_j}{2},\psi\right).$$

Now, using Theorem 5.1.3, we get

$$(X(A_i + A_j)X') \sim G_{p,1}\left(\frac{k_i + k_j}{2},\psi\right), \quad i \neq j.$$

Using Theorem 5.1.7 again, we get $(A_i + A_j)^2 = A_i + A_j$. However, we already know $A_i^2 = A_i$, and $A_j^2 = A_j$. Hence, we get $A_iA_j = O$. ∎

THEOREM 5.1.9. *Let* $X \sim E_{p,n}(O,\Sigma \otimes I_n,\psi)$, $\Sigma > O$ *and assume there exists a p-dimensional constant vector* $\mathbf{v}$, *such that* $P(\mathbf{v}'\Sigma^{-\frac{1}{2}}X = \mathbf{o}) = 0$. *Let* A *and* B *be symmetric idempotent* $n \times n$ *matrices, with* $1 \leq rk(A) < n$, $1 \leq rk(B) < n$, *such that* $AB = O$. *Then,* XAX' *and* XBX' *are independent if and only if* $X \sim N_{p,n}(O,\sigma^2 I_p \otimes I_n)$, *where* $\sigma^2 > 0$.

PROOF. Without loss of generality, we can assume $\Sigma = I_p$.

Let $n_1 = rk(A)$, $n_2 = rk(B)$. Since A ans B are symmetric, we have $BA = AB = O$. Using Theorem 1.3.14, we can find $G \in O(n)$, such that

$$G'AG = \begin{pmatrix} I_{n_1} & O & O \\ O & O & O \\ O & O & O \end{pmatrix}, \text{ and } G'BG = \begin{pmatrix} O & O & O \\ O & I_{n_2} & O \\ O & O & O \end{pmatrix}.$$

Let $n_0 = n_1 + n_2$ and define the $n \times n_0$ matrix $C = \begin{pmatrix} I_{n_0} \\ O \end{pmatrix}$. Moreover, define the $n_0 \times n_0$ matrices $C_1 = \begin{pmatrix} I_{n_1} & O \\ O & O \end{pmatrix}$ and $C_2 = \begin{pmatrix} O & O \\ O & I_{n_2} \end{pmatrix}$. Then, $G'AG = CC_1C'$ and $G'BG = CC_2C'$.

Define $Y = XGC$, then $XAX' = YC_1Y'$ and $XBX' = YC_2Y'$. Partition Y into $Y = (Y_1 \;\; Y_2)$, where Y_1 is $p \times n_1$. Then, $YC_1Y' = Y_1Y_1'$ and $YC_2Y' = Y_2Y_2'$.

First, assume $X \sim N_{p,n}(O, \sigma^2 I_p \otimes I_n)$. Then, $Y \sim N_{p,n_0}(O, \sigma^2 I_p \otimes I_k)$. Thus, the columns of Y are independent and so are Y_1 and Y_2. Hence, Y_1Y_1' and Y_2Y_2' are independent. Therefore, XAX' and XBX' are independent.

On the other hand, assume XAX' and XBX' are independent. Define $\mathbf{y} = X'\mathbf{v}$. Then, $\mathbf{y} \sim E_n(\mathbf{o}, I_n, \psi^*)$, where $\psi^*(t) = \psi(\mathbf{v}'\mathbf{v}t)$. Moreover, $P(\mathbf{y} = \mathbf{o}) = 0$. Since, XAX' and XBX' are independent, so are $\mathbf{y}'A\mathbf{y}$ and $\mathbf{y}'B\mathbf{y}$. Define $\mathbf{w} = G'\mathbf{y}$, then $\mathbf{w} \sim E_n(\mathbf{o}, I_n, \psi^*)$ with $P(\mathbf{w} = \mathbf{o}) = 0$. Let $r\mathbf{u}$ be the stochastic representation of $\mathbf{w}$. Then, $P(r = 0) = 0$. Partition $\mathbf{w}$ into $\mathbf{w} = \begin{pmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \mathbf{w}_3 \end{pmatrix}$ where $\mathbf{w}_1$ is n_1-dimensional, and $\mathbf{w}_2$ is n_2-dimensional.

Let $r \begin{pmatrix} \sqrt{z_1}\, \mathbf{u}_1 \\ \sqrt{z_2}\, \mathbf{u}_2 \\ \sqrt{z_3}\, \mathbf{u}_3 \end{pmatrix}$ be the representation of $\begin{pmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \mathbf{w}_3 \end{pmatrix}$ given by Theorem 2.5.6.

Since $P(r = 0) = 0$, we get $P(\mathbf{w}_1 = \mathbf{o}) = P(\mathbf{w}_2 = \mathbf{o}) = 0$. Now, $\mathbf{y}'A\mathbf{y} = \mathbf{w}_1'\mathbf{w}_1$ and $\mathbf{y}'B\mathbf{y} = \mathbf{w}_2'\mathbf{w}_2$. Define $\mathbf{w}_0 = \begin{pmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{pmatrix}$. Then, $\mathbf{w}_0 \sim E_{r_0}(\mathbf{o}, I_{n^*}, \psi^*)$. Let $r_0\mathbf{u}_0$ be the stochastic representation of $\mathbf{w}_0$. Then,

$$(\mathbf{w}_1'\mathbf{w}_1, \mathbf{w}_2'\mathbf{w}_2) \approx r_0^2(s_1, s_2)$$

where r_0^2 and (s_1,s_2) are independent, $s_1 + s_2 = 1$ and

$$s_1 \sim B\left(\frac{n_1}{2}, \frac{n_2}{2}\right). \tag{5.33}$$

Moreover, $\frac{w_1'w_1}{w_2'w_2} \approx \frac{s_1}{s_2}$ and $w_1'w_1 + w_2'w_2 \approx r_0^2$ are independent. Therefore, $w_1'w_1 \approx \sigma_0^2\chi_{p_1}^2$ and $w_2'w_2 \approx \sigma_0^2\chi_{p_2}^2$ for $\sigma_0 > 0$ (see Lukacs, 1956, p. 208). Since $r_0^2 \approx w_1'w_1 + w_2'w_2$, we have

$$r_0^2 \approx \sigma^2\chi p \text{ with } p = p_1 + p_2. \tag{5.34}$$

We also have $\frac{1 - s_2}{s_2} \approx \frac{w_1'w_1}{w_2'w_2}$. Consequently, $\frac{1}{s_2} \approx \frac{w_1'w_1 + w_2'w_2}{w_2'w_2}$, and $s_2 \approx \frac{w_2'w_2}{w_1'w_1 + w_2'w_2} \sim B\left(\frac{p_2}{2}, \frac{p_1}{2}\right)$. Hence,

$$s_1 \sim B\left(\frac{p_1}{2}, \frac{p_2}{2}\right). \tag{5.35}$$

From (5.33) and (5.35), we get $p_1 = n_1$, $p_2 = n_2$. Thus, $p = n_0$. From (5.34), we get $r_0^2 \approx \sigma_0^2\chi_{n_0}^2$. Since, u_0 is uniformly distributed on S_{n_0}, we get

$w_0 \sim N_{n_0}(o,\sigma_0^2 I_{r_0})$. Consequently, $\psi^*(z) = \exp\left(-\frac{\sigma_0^2 z}{2}\right)$. Therefore,

$\psi(z) = \exp\left(-\frac{\sigma^2 z}{2}\right)$, with $\sigma^2 = \frac{\sigma_0^2}{v'v}$. Hence, $X \sim N_{p,n}(O,\sigma^2 I_p \otimes I_n)$. ■

COROLLARY 5.1.9.1. *Let* $X \sim E_{p,n}(O,I_p \otimes I_n,\psi)$ *and let* x_i *denote the* i^{th} *column of* X, $i = 1,2,\ldots,n$. Define $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $S(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})$.

Assume there exists a p-dimensional constant vector $\mathbf{v}$ *such that* $P(\mathbf{v}'X = \mathbf{o}) = 0$. *Then,* $\bar{\mathbf{x}}$ *and* $S(X)$ *are independent if and only if* $X \sim N_{p,n}(O, \sigma^2 I_p \otimes I_n)$ *with* $\sigma^2 > 0$.

PROOF. If $X \sim N_{p,n}(O, \sigma^2 I_p \otimes I_n)$, then $\mathbf{x}_i$'s constitute a random sample from the distribution $N_p(\mathbf{o}, \sigma^2 I_p)$ and the independence of $\bar{\mathbf{x}}$ and $S(X)$ is a well-known result (see Anderson, 1984, p. 71).

On the other hand, if $\bar{\mathbf{x}}$ and $S(X)$ are independent, then $\bar{\mathbf{x}}\bar{\mathbf{x}}'$ and $S(X)$ are also independent. We have $\bar{\mathbf{x}}\bar{\mathbf{x}}' = \frac{1}{n^2} X \mathbf{e}_n \mathbf{e}_n' X'$ and

$$S(X) = \frac{1}{n} X \left(I_n - \frac{\mathbf{e}_n \mathbf{e}_n'}{n} \right) X'.$$

Let $A = \mathbf{e}_n \mathbf{e}_n'$ and $B = I_n - \frac{\mathbf{e}_n \mathbf{e}_n'}{n}$. Then, A and B satisfy the conditions of Theorem 5.1.9. Therefore $X \sim N_{p,n}(O, \sigma^2 I_p \otimes I_n)$. ∎

5.2. RANK OF QUADRATIC FORMS

The main result of this section uses the following lemma proved by Okamoto (1973).

LEMMA 5.2.1. *Let* X *be a* $p \times n$ *random matrix with absolute continuous distribution. Let* A *be an* $n \times n$ *symmetric matrix with* $\mathrm{rk}(A) = q$. *Then*

i) $P\{\mathrm{rk}(XAX') = \min(p,q)\} = 1$

and

ii) $P\{$*nonzero eigenvalues of* XAX' *are distinct*$\} = 1$.

PROOF. See Okamoto (1973). ∎

Matrix variate elliptically contoured distributions are not necessarily absolutely continuous. However, as the following theorem shows (see Gupta and Varga, 1991b), a result similar to that of Okamoto can be derived for this class of distributions also, if we assume that the distribution is symmetric about the origin and it assumes zero with probability zero.

THEOREM 5.2.1. *Let* $X \sim E_{p,n}(O,\Sigma \otimes \otimes,\psi)$ *with* $P(X = O) = 0$. *Let* A *be an* $n \times n$ *symmetric matrix. Then*

i) $P\{rk(XAX') = \min(rk(\Sigma),rk(\Phi A \Phi))\} = 1$ *and*

ii) $P\{$*the nonzero eigenvalues of* XAX' *are distinct*$\} = 1$.

PROOF. Let $rk(\Sigma) = q$, $rk(\Phi) = m$ *and let* $X \approx rCUD'$ be the stochastic representation of X. Then, $vec(U')$ is uniformly distributed on S_{qm}. Using Theorem 1.3.9 we can write $C = G\begin{bmatrix} B \\ O \end{bmatrix}$, where $G \in O(p)$ and B is a $q \times q$ positive definite matrix. Then

$$XAX' \approx r^2 G \begin{pmatrix} BUD'ADU'B' & O \\ O & O \end{pmatrix} G',$$

where O's denote zero matrices of appropriate dimensions. Since $P(X = O) = 0$, we have $P(r = 0) = 0$. Moreover, the nonzero eigenvalues of $G\begin{pmatrix} BUD'ADU'B & O \\ O & O \end{pmatrix} G'$ are the same as those of $BUD'ADU'B$. Hence,

$$P\{rk(XAX') = rk(BUD'ADU'B')\} = 1 \tag{5.36}$$

and

$$\begin{aligned} &P\{\text{the nonzero eigenvalues of } XAX' \text{ are distinct}\} \\ &= P\{\text{the nonzero eigenvalues of } BUD'ADUB' \text{ are distinct}\}. \end{aligned} \tag{5.37}$$

Let $\Sigma^* = BB'$, $A^* = D'AD$ and define $Y \sim N_{q,m}(O,\Sigma^* \otimes I_m)$. Since B is nonsingular, $\Sigma^* > O$ and so Y is absolutely continuous. Let $Y \approx r^*BU^*$ be the stochastic representation of Y. Then, $vec(U^{*\prime})$ is uniformly distributed on S_{qm}. Now, $YA^*Y' \approx r^{*2}BU^*D'ADU^*B'$ and therefore,

$$P\{rk(YA^*Y') = rk(BU^*D'ADU^*B')\} = 1 \tag{5.38}$$

and

$$\begin{aligned}&P\{\text{the nonzero eigenvalues of } YA^*Y' \text{ are distinct}\}\\&= P\{\text{the nonzero eigenvalues of } BU^*D'ADU^*B' \text{ are distinct}\}.\end{aligned} \tag{5.39}$$

However, from Lemma 5.2.1 we know that

$$P\{rk(YA^*Y') = \min(q, rk(A^*))\} = 1 \tag{5.40}$$

and

$$P\{\text{the nonzero eigenvalues of } YA^*Y' \text{ are distinct}\} = 1. \tag{5.41}$$

Moreover, $rk(A^*) = rk(D'AD) \geq rk(DD'ADD') \geq rk(D^-DD'ADD'D'^-) = rk(D'AD)$ where we used $D^-D = I_m$ which follows from Theorem 1.3.23. Hence,

$$rk(A^*) = rk(DD'ADD') = rk(\Phi A \Phi).$$

Since $vec(U') \approx vec(U^{*\prime})$, we have $U \approx U^*$ and then (i) follows from (5.36), (5.38), and (5.40) and (ii) follows from (5.37), (5.39) and (5.41). ■

If $\Sigma > O$ and $\Phi > O$ in Theorem 5.2.1, we obtain the following result.

THEOREM 5.2.2. *Let* $X \sim E_{p,n}(O, \Sigma \otimes \Phi, \psi)$ *with* $\Sigma > O$, $\Phi > O$ and $P(X = O) = 0$. *Let* A *be an* $n \times n$ *symmetric matrix. Then,*

i) $P\{rk(XAX') = \min(p, rk(A))\} = 1$ *and*

ii) $P\{$*the nonzero eigenvalues of* XAX' *are distinct*$\} = 1$.

PROOF. It follows, directly, from Theorem 5.2.1. ■

COROLLARY 5.2.2.1. *Let* $X \sim E_{p,n}(\mu e_n', \Sigma \otimes \Phi, \psi)$, *where* $p < n$, $\mu \in \mathbb{R}^p$, $\Sigma > O$ *and* $\Phi > O$. *Assume* $P(X = \mu e_n') = 0$. *Let* x_i *be the columns of* X, $i = 1,\dots,n$ *and define* $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$. *Then* $S(X) = \sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})'$ *is positive definite and its characteristic roots are distinct with probability one.*

PROOF. Here $S(X) = X\left(I_n - \frac{e_n e_n'}{n}\right)X' = (X - \mu e_n')\left(I_n - \frac{e_n e_n'}{n}\right)(X - \mu e_n')'$.

Now let $Y = X - \mu e_n'$ and $A = I_n - \frac{e_n e_n'}{n}$. Then, $Y \sim E_{p,n}(O, \Sigma \otimes \Phi, \psi)$, $P(Y = O) = 0$, $S(X) = YAY'$ and from Theorem 5.2.2, we obtain the desired result. ■

5.3. DISTRIBUTIONS OF INVARIANT MATRIX VARIATE FUNCTIONS

In this section, we will derive the distributions of invariant functions of random matrices with m.e.c. distributions. In order to do this, we will need the following theorem (Gupta and Varga, 1992f).

THEOREM 5.3.1. *Let* $X \sim E_{p,n}(O, \Sigma \otimes \Phi, \psi)$ *with* $P(X = O) = 0$. *Assume* $Y \sim N_{p,n}(O, \Sigma \otimes \Phi)$. *Let* $\mathcal{F}$ *be a subset of the* $p \times n$ *real matrices, such that if* $Z \in \mathbb{R}^{p \times n}$, $Z \in \mathcal{F}$, *and* $a > 0$ *then* $aZ \in \mathcal{F}$ *and* $P(X \notin \mathcal{F}) = P(Y \notin \mathcal{F}) = 0$. *Let* $K(Z)$ *be a function defined on* $\mathcal{F}$, *such that if* $Z \in \mathcal{F}$ *and* $a > 0$, *then* $K(Z) = K(aZ)$. *Then,* $K(X)$ *and* $K(Y)$ *are defined with probability one and* $K(X)$ *and* $K(Y)$ *are identically distributed.*

PROOF. $K(X)$ and $K(Y)$ are defined if $X \in \mathcal{F}$ and $Y \in \mathcal{F}$. Since $P(X \notin \mathcal{F}) = P(Y \notin \mathcal{F}) = 0$ we see that $K(X)$ and $K(Y)$ are defined with probability one. Let $r_1 A U_1 B'$ be the stochastic representation of X and $r_2 A U_2 B'$, the stochastic representation of Y. It follows, from the conditions of the theorem, that if $aZ \in \mathcal{F}$ and $a > 0$ then $Z \in \mathcal{F}$. Since $P(X = O) = 0$, we have $P(r_1 = 0) = 0$. Since $P(r_1 A U_1 B' \in \mathcal{F}) = 1$, we get $P(A U_1 B' \in \mathcal{F}) = 1$. So, $K(A U_1 B')$ is defined with probability one. Moreover,

$P\{K(r_1AU_1B') = K(AU_1B')\} = 1$.

Similarly, $P\{K(r_2AU_2B') = K(AU_2B')\} = 1$. But $AU_1B' \approx AU_2B'$. Hence, $K(AU_1B') \approx K(AU_2B')$. Therefore, $K(r_1AU_1B') \approx K(r_2AU_2B)$, which means $K(X) \approx K(Y)$. ■

REMARK 5.3.1. The significance of Theorem 5.3.1 is the following. Assume the conditions of the theorem are satisfied. Then, it is enough to determine the distribution of the function for the normal case, in order to get the distribution of the function, when the underlying distribution is elliptically contoured.

Now, we apply Theorem 5.3.1 to special cases.

THEOREM 5.3.2. *Let* $X \sim E_{p,n}(O, I_p \otimes I_n, \psi)$, *with* $P(X = O) = 0$. *Let* G: $n \times m$ *be such that* $n - m \geq p$ *and* $G'G = I_m$. *Then,*

$$(X(I_n - GG')X')^{-\frac{1}{2}} XG \sim T_{p,m}(n - (m + p) + 1, O, I_p, I_m).$$

PROOF. Let $K(Z) = (Z(I_n - GG')Z')^{-\frac{1}{2}} ZG$. Let $\mathcal{F} = \{Z \mid Z$ is $p \times n$ matrix, such that $Z(I_p - GG')Z'$ is nonsingular$\}$. Clearly, if $Z \in \mathbb{R}^{p\times n}$, $Z \in \mathcal{F}$ and $a > 0$, then $aZ \in \mathcal{F}$. If $Z \in \mathcal{F}$ and $a > 0$, then

$$K(aZ) = (aZ(I_n - GG')(aZ)')^{-\frac{1}{2}} aZG = (Z(I_n - GG')Z')^{-\frac{1}{2}} ZG = K(Z).$$

Let E: $n \times m$ be defined as $E = \begin{pmatrix} I_m \\ O \end{pmatrix}$. Then, $E'E = G'G$. Now, Theorem 1.3.11 says that there exists an $n \times n$ matrix H, such that $HH' = I_n$ and $G'H = E'$. That means, H is orthogonal and $H'G = E$. So, we have

$$\begin{aligned} I_n - GG' &= I_n - HEE'H' \\ &= H(I_n - EE')H' \\ &= H\left(I_n - \begin{pmatrix} I_m & O \\ O & O \end{pmatrix}\right)H' \end{aligned}$$

$$= H\begin{pmatrix} O & O \\ O & I_{n-m} \end{pmatrix} H' = HDH',$$

where $D = \begin{pmatrix} O & O \\ O & I_{n-m} \end{pmatrix}$. Clearly, $D = BB'$, where B is an $n \times (n-m)$ matrix defined as $B = \begin{pmatrix} O \\ I_{n-m} \end{pmatrix}$. Using Theorem 5.2.2, we get

$$\begin{aligned} P\{rk(X(I_n - GG')X') &= \min(rk(I_n - GG'),p)\} \\ &= 1. \end{aligned}$$

However, $rk(I_n - GG') = rk(HDH') = rk(D) = n - m$, and since $n - m \geq p$, we see that rk $(X(I_n - GG')X') = p$, with probability one. Therefore, $X(I_n - GG')X'$ is of full rank with probability one. So, $P(X \notin \mathcal{F}) = 0$.

Hence, we can use Theorem 5.3.1. That means we can assume $X \sim N_{p,n}(O, I_p \otimes I_n)$. Let $V = XH$. Then, $V \sim N_{p,n}(O, I_p \otimes I_n)$. Partition V as $V = (V_1, V_2)$, where V_1 is $p \times m$. Then, $V_1 \sim N_{p,m}(O, I_p \otimes I_m)$, and $V_2 \sim N_{p,n-m}(O, I_p \otimes I_{n-m})$, where V_1 and V_2 are independent. Now,

$$\begin{aligned} (X(I_n - GG')X')^{-\frac{1}{2}} XG &= (XHDH'X')^{-\frac{1}{2}} XHE \\ &= (VBB'V')^{-\frac{1}{2}} VE \\ &= \left(V \begin{pmatrix} O \\ I_{n-m} \end{pmatrix} (O, I_{n-m}) V'\right)^{-\frac{1}{2}} V \begin{pmatrix} I_m \\ O \end{pmatrix} \\ &= (V_2 V_2')^{-\frac{1}{2}} V_1. \end{aligned}$$

Here, $V_2 V_2' \sim W_p(n - m, I_p)$, $V_1 \sim N_{p,m}(O, I_p \otimes I_m)$ and V_1 and $V_2 V_2'$ are independent. From Dickey (1967), we get that under these conditions, $(V_2 V_2')^{-\frac{1}{2}} V_1 \sim T_{p,m}(n - m - p + 1, O, I_p, I_m)$ (also see Javier and Gupta, 1985a). ∎

THEOREM 5.3.3. *Let* $X \sim E_{p,n}(O, I_p \otimes I_n, \psi)$ *with* $P(X = O) = 0$. *Let* B: $n \times n$ *be a symmetric, idempotent matrix of rank* m *where* $m \geq p$ *and* $n - m \geq p$. *Then,* $(XX')^{-\frac{1}{2}}(XBX')(XX')^{-\frac{1}{2}} \sim B_p^I\left(\frac{m}{2}, \frac{n-m}{2}\right)$.

PROOF. Let $K(Z) = (ZZ')^{-\frac{1}{2}}(ZBZ')(ZZ')^{-\frac{1}{2}}$. Let $\mathcal{F} = \{Z \mid Z$ is $p \times n$ matrix, such that ZZ' is nonsingular$\}$. Clearly, if $Z \in \mathbb{R}^{p \times n}$, $Z \in \mathcal{F}$ and $a > 0$, then $aZ \in \mathcal{F}$. If $Z \in \mathcal{F}$ and $a > 0$, then

$$\begin{aligned} K(aZ) &= (a^2ZZ')^{-\frac{1}{2}}(a^2ZBZ')(a^2ZZ')^{-\frac{1}{2}} \\ &= (ZZ')^{-\frac{1}{2}}(ZBZ')(ZZ')^{-\frac{1}{2}} \\ &= K(Z). \end{aligned}$$

Using Theorem 5.2.2, we get

$$P\{rk(XX') = \min(rk(I_n), p)\} = 1.$$

Since $m \geq p$ and $n - m \geq p$ we have $n \geq 2p$ and hence, $n \geq p$. Thus, $\min(rk(I_n), p) = p$. Therefore, XX' is of full rank with probability one. So, $P(X \notin \mathcal{F}) = 0$. Similarly, $P(Y \notin \mathcal{F}) = 0$. Hence, we can use Theorem 5.3.1. That means we can assume $X \sim N_{p,n}(O, I_p \otimes I_n)$. Since B is a symmetric, idempotent matrix of rank m, there exists an $n \times n$ orthogonal matrix H, such that $B = H\begin{pmatrix} I_q & O \\ O & O \end{pmatrix}H'$. We can write $\begin{pmatrix} I_q & O \\ O & O \end{pmatrix} = CC'$ where C is an $n \times m$ matrix defined as $C = \begin{pmatrix} I_q \\ O \end{pmatrix}$. Let $V = XH$. Then, $V \sim N_{p,n}(O, I_p \otimes I_n)$. Partition V as $V = (V_1, V_2)$ where V_1 is $p \times m$. Then, $V_1 \sim N_{p,m}(O, I_p \otimes I_m)$, and $V_2 \sim N_{p,n-m}(O, I_p \otimes I_{n-m})$, where V_1 and V_2 are independent.

Now,

$$
\begin{aligned}
(XX')^{-\frac{1}{2}}(XBX')(XX')^{-\frac{1}{2}} &= (XHH'X')^{-\frac{1}{2}}(XHCC'H'X')(XHH'X')^{-\frac{1}{2}} \\
&= (VV')^{-\frac{1}{2}}(VCC'V')(VV')^{-\frac{1}{2}} \\
&= (V_1V_1' + V_2V_2')^{-\frac{1}{2}}(V_1V_1')(V_1V_1' + V_2V_2')^{-\frac{1}{2}}.
\end{aligned}
$$

Here $V_1V_1' \sim W_p(m,I_p)$, $V_2V_2' \sim W_p(n-m,I_p)$ and V_1V_1' and V_2V_2' are independent. Finally from Olkin and Rubin (1964), we get that under these conditions

$$
(V_1V_1' + V_2V_2')^{-\frac{1}{2}}(V_1V_1'))(V_1V_1' + V_2V_2')^{-\frac{1}{2}} \sim B_p^I\left(\frac{m}{2}, \frac{n-m}{2}\right). \blacksquare
$$

THEOREM 5.3.4. *Let* $X \sim E_{p,n}(O, I_p \otimes I_n, \psi)$, *with* $P(X = O) = 0$. *Let* A *and* B *be* $n \times n$ *symmetric, idempotent matrices,* $rk(A) = n_1$, $rk(B) = n_2$, *such that* $n_1, n_2 \geq p$ *and* $AB = O$. *Then,* $(XAX')^{-\frac{1}{2}}(XBX')(XAX')^{-\frac{1}{2}} \sim B_p^{II}\left(\frac{n_2}{2}, \frac{n_1}{2}\right)$.

PROOF. Let $K(Z) = (ZAZ')^{-\frac{1}{2}}(ZBZ')(ZAZ')^{-\frac{1}{2}}$. Let $\mathcal{F} = \{Z \mid Z$ is $p \times n$ matrix, such that ZAZ' is nonsingular$\}$. Clearly, if $Z \in \mathbb{R}^{p\times n}$, $Z \in \mathcal{F}$ and $a > 0$, then $aZ \in \mathcal{F}$. If $Z \in \mathcal{F}$ and $a > 0$, then

$$
\begin{aligned}
K(aZ) &= (a^2ZAZ')^{-\frac{1}{2}}(a^2ZBZ')(a^2ZAZ')^{-\frac{1}{2}} \\
&= (ZAZ')^{-\frac{1}{2}}(ZBZ')(ZAZ')^{-\frac{1}{2}} \\
&= K(Z).
\end{aligned}
$$

Using Theorem 5.2.2, we get

$$
P\{rk(XAX') = \min(rk(A),p)\} = 1.
$$

However $rk(A) = n_1 \geq p$, and hence $\min(rk(A),p) = p$. Therefore, XAX' is of full rank with probability one. So, $P(X \notin \mathcal{F}) = 0$. Similarly, $P(Y \notin \mathcal{F}) = 0$.

Hence, we can use Theorem 5.3.1. That means we can assume

$X \sim N_{p,n}(O, I_p \otimes I_n)$. Since A and B are symmetric, idempotent matrices and $AB = O$, there exists an $n \times n$ orthogonal matrix H, such that

$$H'AH = \begin{pmatrix} I_{n_1} & O & O \\ O & O & O \\ O & O & O \end{pmatrix} \text{ and } H'BH = \begin{pmatrix} O & O & O \\ O & I_{n_2} & O \\ O & O & O \end{pmatrix}$$

(see Hocking, 1985).

We can write $H'AH = CC'$ and $H'BH = DD'$, where $C' = (I_{n_1}, O, O)$ and $D' = (O, I_{n_2}, O)$. Let $V = XH$. Then, $V \sim N_{p,n}(O, I_p \otimes I_n, \psi)$. Partition V as $V = (V_1, V_2, V_3)$ where V_1 is $p \times n_1$ and V_2 is $p \times n_2$. Then, $V_1 \sim N_{p,m}(O, I_p \otimes I_n)$ and$V_2 \sim N_{p,n-m}(O, I_p \otimes I_{n-m})$ where V_1 and V_2 are independent.

Now,

$$\begin{aligned}
&(XAX')^{-\frac{1}{2}}(XBX')(XAX')^{-\frac{1}{2}} \\
&= (XHCC'H'X')^{-\frac{1}{2}}(XHDD'H'X')(XHCC'H'X')^{-\frac{1}{2}} \\
&= (VCC'V')^{-\frac{1}{2}}(VDD'V')(VCC'V')^{-\frac{1}{2}} \\
&= (V_1V_1')^{-\frac{1}{2}}(V_2V_2')(V_1V_1')^{-\frac{1}{2}}.
\end{aligned}$$

Here, $V_1V_1' \sim W_p(n_1, I_p)$, $V_2V_2' \sim W_p(n_2, I_p)$ and V_1V_1' and V_2V_2' are independent. From Olkin and Rubin (1964), we get that under these conditions

$$(V_1V_1')^{-\frac{1}{2}}(V_2V_2')(V_1V_1')^{-\frac{1}{2}} \sim B_p^{II}\left(\frac{n_2}{2}, \frac{n_1}{2}\right). \blacksquare$$

The next theorem shows that under some general conditions, Hotelling's T^2 statistic has the same distribution in the elliptically contoured case, as in the normal case; that is, we get an F distribution.

THEOREM 5.3.5. *Let* $X \sim E_{p,n}(O, \Sigma \otimes I_n, \psi)$ *with* $P(X = O) = 0$. *Assume* $p < n$. *Let* x_i *be the* i^{th} *column of* X, $i = 1,\dots,n$. *Let* $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$. *Define* $S(x) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})'$ *and* $T^2(x) = n\bar{x}' S(x)^{-1}\bar{x}$. *Then* $\frac{T^2(X)}{n-1}\frac{n-p}{p} \sim F_{p,n-p}$.

PROOF. Let $A = I_n - \frac{e_n e_n'}{n}$. Then, $S(X) = XAX'$. We also have $\bar{x} = \frac{1}{n}Xe_n$. Thus, $T^2(X) = e_n'X'(XAX')^{-1}Xe_n$.

Let $K(Z) = e_n' Z'(ZAZ')^{-1}Ze_n$. Let $\mathcal{F} = \{Z \mid Z$ is $p \times n$ matrix, such that ZAZ' is nonsingular$\}$. Clearly, if $Z \in \mathbb{R}^{p\times n}$, $Z \in \mathcal{F}$, and $a > 0$ then $aZ \in \mathcal{F}$. If $Z \in \mathcal{F}$ and $a > 0$, then

$$\begin{aligned} K(aZ) &= ne_n' aZ'(a^2 ZAZ')^{-1} aZe_n \\ &= ne_n' Z'(ZAZ')^{-1}Ze_n. \end{aligned}$$

From Corollary 5.2.2.1, we see that XAX' is of full rank with probability one. So, $P(X \notin \mathcal{F}) = 0$. Similarly, $P(Y \notin \mathcal{F}) = 0$.

Hence, we can use Theorem 5.3.1. That means we can assume $X \sim N_{p,n}(O, \Sigma \otimes I_n)$. However, for the normal case this is a well known result (see Corollary 5.2.1 of Anderson, 1984, p. 163). ■

THEOREM 5.3.6. *Let* $X \sim E_{p,n}(\mu e_n', \Sigma \otimes I_n, \psi)$ *with* $P(X = \mu e_n') = 0$ *and* $\mu \in \mathbb{R}^p$. *Assume* $n > p$. *Let* $Y \sim N_{p,n}(\mu e_n', \Sigma \otimes I_n)$, *and* $S(X) = X\left(I_n - \frac{e_n e_n'}{n}\right)X'$. *Then, the principal components of* $S(X)$ *have the same joint distribution as the principal components of* $S(Y)$.

PROOF. Let $A = I_n - \frac{e_n e_n'}{n}$ and $S(Z) = ZAZ'$ for $Z \in \mathbb{R}^{p\times n}$. First, note that $S(Z) = S(Z - \mu e_n')$ therefore, without loss of generality, we can assume $\mu = \mathbf{o}$.

Let $K(Z) = \{$normalized characteristic vectors of $ZAZ'\}$. Let

$\mathcal{F} = \{Z \mid Z$ is $p \times n$ matrix, such that the characteristic roots of ZAZ' are nonzero and distinct$\}$. Clearly, if $Z \in \mathbb{R}^{n \times m}$, $Z \in \mathcal{F}$, and $a > 0$, then $aZ \in \mathcal{F}$. If $Z \in \mathcal{F}$ and $a > 0$. Then obviously, $K(aZ) = K(Z)$. Using Corollary 5.2.2.1, we find that the characteristic roots of XAX' are nonzero and distinct with probability one. So, $P(X \notin \mathcal{F}) = 0$. Similarly, $P(Y \notin \mathcal{F}) = 0$.

Now applying Theorem 5.3.1, we obtain the desired result. ■

CHAPTER 6

CHARACTERIZATION RESULTS

6.1. CHARACTERIZATIONS BASED ON INVARIANCE

In this section, we characterize the parameters of m.e.c. distributions which are invariant under certain linear transformations. First we prove the following lemma.

LEMMA 6.1.1. *The* $p \times p$ *matrix* Σ *defined by*

$$\Sigma = aI_p + be_pe_p',$$

is positive semidefinite if and only if $a \geq 0$ *and* $a \geq -pb$.

PROOF: From part (vi) of Theorem 1.3.2, we have to show that the characteristic roots of Σ are nonnegative. From Theorem 1.3.5 we obtain

$$\begin{aligned} |\Sigma - \lambda I_p| &= |(a-\lambda)I_p + be_pe_p'| \\ &= ((a-\lambda+b)-b)^{p-1}\,(a-\lambda+b+(p-1)b) \\ &= (a-\lambda)^{p-1}\,(a+pb-\lambda). \end{aligned}$$

Hence, the characteristic roots of Σ are $\lambda_1 = a$ and $\lambda_2 = a + pb$. Therefore, the characteristic roots of Σ are nonnegative if and only if $a \geq 0$ and $a + pb \geq 0$. ■

THEOREM 6.1.1. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ *with* $p > 1$ *and* $\Phi \neq O$. *Define*

$$\mathcal{P} = \{P\colon\ P \text{ is } p \times p \text{ permutation matrix}\},$$

$\mathcal{R} = \{R\colon$ R *is* $p \times p$ *signed permutation matrix*$\}$, *and*

$\mathcal{G} = \{G\colon$ G *is* $p \times p$ *orthogonal matrix*$\}$.

Then,

a) *for every* $P \in \mathcal{P}$, $PX \approx X$ *if and only if* $M = \mathbf{e}_p\mu'$ *where* $\mu \in \mathbb{R}^n$, $\Sigma = aI_p + b\mathbf{e}_p\mathbf{e}_p'$, $a,b \in \mathbb{R}$, $a \geq 0$, *and* $a \geq -pb$,

b) *for every* $R \in \mathcal{R}$, $RX \approx X$ *if and only if* $M = O$ *and* $\Sigma = aI_p$, *where* $a \geq 0$,

c) *for every* $G \in \mathcal{G}$, $GX \approx X$ *if and only if* $M = O$ *and* $\Sigma = aI_p$, *where* $a \geq 0$.

PROOF: a) First, assume that $X \sim E_{p,n}(\mathbf{e}_p\mu', (aI_p + b\mathbf{e}_p\mathbf{e}_p') \otimes \Phi, \psi)$ with $a \geq 0$ and $a \geq -pb$. Then, from Lemma 6.1.1, Σ is positive semidefinite.

Let $P \in \mathcal{P}$. Then, $PX \sim E_{p,n}(P\mathbf{e}_p\mu', (P(aI_p + b\mathbf{e}_p\mathbf{e}_p')P') \otimes \Phi, \psi)$. Since $P\mathbf{e}_p = \mathbf{e}_p$ and $PP' = I_p$, we get $PX \sim E_{p,n}(\mathbf{e}_p\mu', (aI_p + b\mathbf{e}_p\mathbf{e}_p') \otimes \Phi, \psi)$, which proves that $PX \approx X$.

Next, assume $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ and $PX \approx X$ for every $P \in \mathcal{P}$. Then, $PX \sim E_{p,n}(PM, P\Sigma P' \otimes \Phi, \psi)$ and hence, $PM = M$ and $P\Sigma P' = \Sigma$ for every $P \in \mathcal{P}$. Now, we introduce the following notation. Let $P(k,\ell)$, $1 \leq k,\ell \leq p$ denote a $p \times p$ symmetric matrix, whose $(i,j)^{\text{th}}$ element is

$$\begin{cases} 1 \text{ if } i = j, & i \neq k,\ i \neq \ell \\ 1 \text{ if } i = \ell, & j = k \\ 1 \text{ if } i = k, & j = \ell \\ 0 \text{ elsewhere.} & \end{cases}$$

Then, it is easy to see that $P(k,\ell) \in \mathcal{P}$.

From $P(1,i)M = M$, $i = 2,\ldots,p$, we get $M = \mathbf{e}_p\mu'$, where $\mu \in \mathbb{R}^n$. From $P(1,i)\Sigma P(1,i) = \Sigma$, $i = 2,\ldots,p$ we get

$$\sigma_{ii} = \sigma_{11} \tag{6.1}$$

and

$$\sigma_{ij} = \sigma_{1j} \text{ if } j > i. \tag{6.2}$$

If $p \geq 3$, then from $P(2,i)\Sigma P(2,i) = \Sigma$, $i = 3,\dots,p$, we get

$$\sigma_{1j} = \sigma_{12} \text{ if } j \geq 3. \tag{6.3}$$

From (6.1), it is clear that the diagonal elements of Σ are equal, whereas (6.2) and (6.3) show that the off-diagonal elements of Σ are also equal. Therefore, $\Sigma = aI_p + be_pe_p'$, and $|\Sigma| = (a + pb)a^{p-1}$. From Lemma 6.1.1, in order for $aI_p + be_pe_p'$ to be positive semidefinite, we must have $a \geq 0$ and $a \geq -pb$.

b) First assume that $X \sim E_{p,n}(O,aI_p \otimes \Phi,\psi)$ with $a \geq 0$. Let $R \in \mathcal{R}$, then $RX \sim E_{p,n}(O,aRI_pR' \otimes \Phi,\psi)$. Since $RR' = I_p$, we have $RX \approx X$.

Next, assume that $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ and $RX \approx X$ for every $R \in \mathcal{R}$. Let $R = -I_p$, then $RX \sim E_{p,n}(-M,\Sigma \otimes \Phi,\psi)$. Therefore $M = -M$ i.e. $M = O$. Since $\mathcal{P} \subset \mathcal{R}$, Σ must be of the form $aI_p + be_pe_p'$, $a \geq 0$ and $a \geq -pb$. So, $X \sim E_{p,n}(O,(aI_p + be_pe_p') \otimes \Phi,\psi)$.

Let R be a $p \times p$ symmetric matrix, whose $(i,j)^{th}$ element is
$$\begin{cases} -1 & \text{if } i = j = 1 \\ 1 & \text{if } i = j > 1 \\ 0 & \text{otherwise.} \end{cases}$$
Then, $R \in \mathcal{R}$ and $RX \sim E_{p,n}(O,R(aI_p + be_pe_p')R' \otimes \Phi,\psi)$. Since, $R(aI_p + be_pe_p')R' = aI_p + bRe_pe_p'R$, we must have $bRe_pe_p' = be_pe_p'$, which can also be written as

$$b(e_pe_p' - Re_pe_p'R) = O. \tag{6.4}$$

Now, $e_pe_p' - Re_pe_p'R$ is a matrix whose $(i,j)^{th}$ element is $\begin{cases} 2 & \text{if } i = 1, \quad j \geq 2 \\ 2 & \text{if } i \geq j, \quad j = 1 \\ 0 & \text{elsewhere} \end{cases}$.

Hence, $e_pe_p' - Re_pe_p'R \neq O$ and thus, from (6.4), we conclude that $b = 0$. Therefore, $X \sim E_{p,n}(O,aI_p \otimes \Phi,\psi)$ with $a \geq 0$.

c) First assume $X \sim E_{p,n}(O,aI_p \otimes \Phi,\psi)$ with $a \geq 0$. Let $G \in \mathcal{G}$, then $GX \sim E_{p,n}(O,aGG' \otimes \Phi,\psi)$ and since $GG' = I_p$, we have $GX \approx X$.

On the other hand, if $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ and $GX \approx X$ for every $G \in \mathcal{G}$, then $RX \approx X$ for $R \in \mathcal{R}$ must also hold, since $\mathcal{R} \subset \mathcal{G}$. Then, using part (b), we obtain $X \sim E_{p,n}(O,aI_p \otimes \Phi,\psi)$ with $a \geq 0$. ■

DEFINITION 6.1.1. Let $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$. Then, X is called left-spherical, if $GX \approx X$, for every $G \in O(p)$, right-spherical if $XH \approx X$ for every $H \in O(n)$, and spherical if it is both left- and right-spherical.

THEOREM 6.1.2. *Let* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$. *Then,* X *is left-spherical if and only if* $M = O$ *and* $\Sigma = aI_p$, *with* $a \geq 0$; *right-spherical if and only if* $M = O$ *and* $\Phi = bI_n$, *with* $b \geq 0$; *and spherical if and only if* $M = O$, $\Sigma = aI_p$ *and* $\Phi = bI_n$, *with* $a \geq 0$.

PROOF: It follows, from Theorem 6.1.1. ∎

6.2. CHARACTERIZATIONS OF NORMALITY

In this section, it is shown that if m.e.c. distributions possess certain properties, they must be normal. The first result shows that the normality of one element of a random matrix with m.e.c. distribution implies the normality of the whole random matrix.

THEOREM 6.2.1. *Let* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ *and assume that there exist* i *and* j *such that* x_{ij} *is nondegenerate normal. Then,* $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$.

PROOF: It follows, from Theorems 2.3.3, that $x_{ij} \sim E_1(m_{ij},\sigma_{ii}\phi_{jj},\psi)$. Since x_{ij} is normal, we have $\psi(z) = \exp\left(-\frac{z}{2}\right)$. Then $E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ becomes $N_{p,n}(M,\Sigma \otimes \Phi)$. ∎

The following characterization results are based on independence.

THEOREM 6.2.2. *Let* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$. *If* X *has two elements that are nondegenerate and independent, then* $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$.

PROOF: Without loss of generality we can assume $M = O$. Let us denote the two independent elements of X by y and z. Then, we get

$$\begin{pmatrix} y \\ z \end{pmatrix} \sim E_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} a & c \\ c & b \end{pmatrix}, \psi\right).$$

From Theorem 2.3.1, we have $y \sim E_1(0,a,\psi)$ and from Theorem 2.6.2, we obtain

$$y|z \sim E_1\left(\frac{c}{b}z, a - \frac{c^2}{b}, \psi^*\right).$$

Since, y and z are independent, $y|z \approx y$. Hence, $y \sim E_1\left(\frac{c}{b}z, a - \frac{c^2}{b}, \psi^*\right)$ and in view of Theorem 2.1.4 we must have $\frac{c}{b}z = 0$ for every z real number. This is possible, only if $c = 0$. Therefore,

$$\begin{pmatrix} y \\ z \end{pmatrix} \sim E_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \psi\right).$$

Now let $y_1 = \sqrt{a}\,y$, $z_1 = \sqrt{b}\,z$, and $\mathbf{w} = \begin{pmatrix} y_1 \\ z_1 \end{pmatrix}$. Then, $\mathbf{w} \sim E_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \psi\right)$, and its characteristic function at $\mathbf{t} = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}$ is

$$\phi_{\mathbf{w}}(\mathbf{t}) = \psi\left((t_1 \quad t_2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} t_1 \\ t_2 \end{pmatrix}\right) = \psi(t_1^2 + t_2^2).$$

The characteristic functions of y_1 and z_1 are $\phi_{y_1}(t_1) = \psi(t_1^2)$ and $\phi_{z_1}(t_2) = \psi(t_2^2)$. Since y and z are independent, so are y_1 and z_1. Therefore, $\phi_{\mathbf{w}}(\mathbf{t}) = \phi_{y_1}(t_1)\phi_{z_1}(t_2)$; that is $\psi(t_1^2 + t_2^2) = \psi(t_1^2)\psi(t_2^2)$, or equivalently

$$\psi(t_1 + t_2) = \psi(t_1)\,\psi(t_2), \qquad t_1 \geq 0,\ t_2 \geq 0 \tag{6.5}$$

Now, using Corollary 1.4.1.1 we see that $\psi(t) = e^{-kt}$, $k \geq 0$, $t \geq 0$. Moreover, $k = 0$ is impossible since this would make y degenerate. Therefore, $k > 0$ and hence X is normal. ■

COROLLARY 6.2.2.1. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ *and* x_i, $i = 1,\dots,n$ *denote the columns of* X. *If* $x_1, x_2, \dots, x_n$ *are all nondegenerate and independent, then* $X \sim N_{p,n}(M, \Sigma \otimes \Phi)$, *where* Φ *is diagonal.*

PROOF: This follows, from Theorem 6.2.1. Since if two columns are independent, then any two elements, picked one from each of these columns will also be independent. The structure of Φ is implied by the fact that $x_1, x_2, \dots, x_n$ are independent and normal. ■

REMARK 6.2.1. For the case $p = 1$, a result similar to Corollary 6.2.1.1, was given by Kelker (1970). However, he had stronger conditions since he made the diagonality of Φ an assumption of the theorem.

The following theorem, although not a characterization result gives the idea for a further characterization of normality.

THEOREM 6.2.3. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ *be nondegenerate, with finite second order moment. Assume* A: $q \times p$, B: $n \times k$ C: $r \times p$, *and* D: $n \times \ell$ *are constant matrices. Then,* AXB *and* CXD *are uncorrelated if and only if either* $A\Sigma C' = O$ *or* $B'\Phi D = O$.

PROOF: Without loss of generality, we can assume $M = O$. Then, using Theorem 1.3.17, we can write

$$\begin{aligned}
\mathrm{Cov}(\mathrm{vec}(AXB)', \mathrm{vec}(CXD)') &= \mathcal{E}(\mathrm{vec}(AXB)'(\mathrm{vec}(CXD)')') \\
&= \mathcal{E}((A \otimes B')\, \mathrm{vec}(X')(\mathrm{vec}(X'))'(C' \otimes D)) \\
&= -\, 2\psi'(0)(A \otimes B')(\Sigma \otimes \Phi)(C' \otimes D) \\
&= -2\psi'(0)(A\Sigma C') \otimes (B'\Phi D).
\end{aligned}$$

Here, $\psi'(0) \neq 0$, since X is nondegenerate, so we must have $(A\Sigma C') \otimes (B'\Phi D) = O$. This holds iff $A\Sigma C' = O$ or $B'\Phi D = O$. ■

REMARK 6.2.2. Since, in the normal case, uncorrelatedness and independence are equivalent, Theorem 6.2.3 implies that for $X \sim N_{p,n}(M, \Sigma \otimes \Phi)$, AXB and CXD are independent iff $A\Sigma C' = O$ or

$B'\Phi D = O$. This property of the matrix variate normal distribution was obtained by Nel (1977).

Theorem 6.2.3 shows that under certain conditions, two linear expressions in a random matrix with m.e.c. distribution are uncorrelated, whereas Remark 6.2.2 says that if the underlying distribution is normal the two linear transforms are independent. The question arises whether the independence of the linear transforms characterizes normality in the class of m.e.c. distributions. The answer is yes, as the next theorem shows.

THEOREM 6.2.4. *Let* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ *and* A: $q \times p$, B: $n \times k$, C: $r \times p$, and D: $n \times \ell$ *be constant nonzero matrices. If* AXB *and* CXD *are nondegenerate and independent, then* X *is normal.*

PROOF: Without loss of generality, we can assume $M = O$. Since AXB and CXD are independent, so are $\mathrm{vec}(AXB)' = (A \otimes B')\,\mathrm{vec}(X')$ and $\mathrm{vec}(CXD)' = (C \otimes D')\,\mathrm{vec}(X')$. Let $\mathbf{x} = \mathrm{vec}(X)$, then $\mathbf{x} \sim E_{pn}(\mathbf{o},\Sigma \otimes \Phi,\psi)$. Let $\mathbf{v}'$ be a nonzero row of $A \otimes B'$ and $\mathbf{w}'$ be a nonzero row of $C \otimes D'$. Then, $\mathbf{v}'\mathbf{x}$ and $\mathbf{w}'\mathbf{x}$ are independent. Let $H = \begin{pmatrix} \mathbf{v}' \\ \mathbf{w}' \end{pmatrix}$. Then

$H\mathbf{x} = \begin{pmatrix} \mathbf{v}'\mathbf{x} \\ \mathbf{w}'\mathbf{x} \end{pmatrix} \sim E_2(\mathbf{o},H(\Sigma \otimes \Phi)H',\psi)$. Since $\mathbf{v}'\mathbf{x}$ and $\mathbf{w}'\mathbf{x}$ are independent and their joint distribution is elliptically contoured, from Theorem 6.2.2 we conclude that $H\mathbf{x}$ is normal. Therefore, X is normal. ∎

The following characterization results are based on conditional distributions (see Gupta and Varga, 1990, 1991a, 1992a, 1992e).

THEOREM 6.2.5. *Let* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ *be nondegenerate. Let* X, M, *and* Σ *be partitioned as* $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, $M = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}$, $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, *where* X_1 *and* M_1 *are* $q \times n$ *and* Σ_{11} *is* $q \times q$. *Assume* $\mathrm{rk}(\Sigma_{22}) \geq 1$ *and* $\mathrm{rk}(\Sigma) - \mathrm{rk}(\Sigma_{22}) \geq 1$. *Let* $X_1 | X_2 \sim E_{q,n}(M_1 + \Sigma_{12}\Sigma_{22}^{-1}(X_2 - M_2),\Sigma_{11\cdot 2} \otimes \Phi,\psi_{q(X_2)})$ *with*

$q(X_2) = tr((X_2 - M_2)' \Sigma_{11}^{-}(X_2 - M_2)\Phi^{-})$. *Then,* $\psi_{q(X_2)}$ *does not depend on* X_2, *with probability one if and only if* $X \sim N_{p,n}(M, \Sigma \otimes \Phi)$.

PROOF: It is known that if $X \sim N_{p,n}(M, \Sigma \otimes \Phi)$, then $\psi_{q(X_2)}(z) = \exp\left(-\frac{z}{2}\right)$ and hence, $\psi_{q(X_2)}$ does not depend on X_2 (see, e.g. Anderson, 1984, p. 37).

Conversely, assume $\psi_{q(X_2)}$ does not depend on X_2 for $X_2 \in A$, where $P(X_2 \in A) = 1$. Thus, for $X_2 \in A$, we can write $\psi_{q(X_2)}(z) = \psi_0(z)$, where ψ_0 does not depend on X_2. It follows from the definition of $q(X_2)$, that it suffices to consider the case $M = O$, $\Sigma = I_p$, $\Phi = I_n$ (see Theorem 2.6.5). Let T be a $p \times n$ matrix and partition it as $T = \begin{pmatrix} T_1 \\ T_2 \end{pmatrix}$, where T_1 is $q \times n$. Then, the characteristic function of X is

$$\begin{aligned} \phi_X(T) &= \psi(tr(T'T)) \\ &= \psi(tr(T_1'T_1 + T_2'T_2)) \\ &= \psi(tr(T_1'T_1) + tr(T_2'T_2)). \end{aligned} \tag{6.6}$$

On the other hand,

$$\begin{aligned} \phi_X(T) &= \mathcal{E}(etr(iT'X)) \\ &= \mathcal{E}(etr(iXT')) \\ &= \mathcal{E}(etr(i(X_1T_1' + X_2T_2'))) \\ &= \mathcal{E}(E(etr(i(X_1T_1' + X_2T_2')) \mid X_2)) \\ &= \mathcal{E}(etr(iX_2T_2')\, E(etr(iX_1T_1') \mid X_2)) \\ &= \mathcal{E}(etr(iX_2T_2')\, \psi_0(tr(T_1'T_1)) \\ &= \psi(tr(T_2'T_2))\, \psi_0(tr(T_1'T_1)). \end{aligned} \tag{6.7}$$

Let $u = tr(T_1'T_1)$ and $v = tr(T_2'T_2)$. Then, from (6.6) and (6.7), we obtain

$$\psi(u + v) = \psi(u)\, \psi_0(v) \tag{6.8}$$

for $u,v \geq 0$. Taking $u = 0$ in (6.8), we get $\psi(v) = \psi_0(v)$. Hence, (6.8) gives

$$\psi(u + v) = \psi(u)\,\psi(v) \tag{6.9}$$

for $u,v \geq 0$. Now from Corollary 1.4.1.1 we see that $\psi(z) = e^{-kz}$, $k \geq 0$, $t \geq 0$. Moreover, k cannot be zero, since this would make X degenerate. Therefore, $k > 0$ and hence X is normal. ■

The next theorem shows that the normality of the conditional distributional characterizes the normal distribution in the class of m.e.c. distributions.

THEOREM 6.2.6. *Let* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ *be nondegenerate. Let* X *and* Σ *be partitioned as* $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, *where* X_1 *is* $q \times n$ *and* Σ_{11} *is* $q \times q$. *Assume* $\mathrm{rk}(\Sigma_{22}) \geq 1$ *and* $\mathrm{rk}(\Sigma) - \mathrm{rk}(\Sigma_{22}) \geq 1$. *Then,*
$P(X_1 \mid X_2$ *is nondegenerate normal*$) = 1$ *if and only if* $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$.

PROOF: It has already been mentioned in the proof of Theorem 6.2.5 that if $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$, then $X_1 \mid X_2$ is nondegenerate normal with probability one.

Conversely, assume $X_1 \mid X_2$ is nondegenerate normal with probability one. Then, from the definition of $q(X_2)$, it follows that it suffices to consider the case $M = O$, $\Sigma = I_p$, $\Phi = I_n$, where now $q(X_2) = \mathrm{tr}(X_2'X_2)$. So $q(X_2) = 0$ if and only if $X_2 = O$. Thus, $P(q(X_2) = 0) = P(X_2 = O)$.

On the other hand, from Corollary 3.1.5.1 it follows that $P(X_2 = O) = P(X = O)$. Hence, $P(q(X_2) = 0) = P(X = O)$. However $P(X = O) > 0$ is not possible since this would imply that $X_1 \mid X_2$ is degenerate with positive probability. Therefore, $P(q(X_2) = 0) = 0$ must hold. Hence, there exists a set $A \subset \mathbb{R}^{q \times n}$ such that $P(X_2 \in A) = 1$. If $X_2 \in A$, then $q(X_2) > 0$ and $X_1 \mid X_2$ is nondegenerate normal. So if $X_2 \in A$, then we get

$$X_1 \mid X_2 \sim E_{q,n}(O,I_q \otimes I_n,\psi_{q(X_2)}) \text{ with } \psi_{q(X_2)}(z) = \exp\left(-\frac{c(q(X_2))z}{2}\right). \tag{6.10}$$

Here, c denotes a function of $q(X_2)$, such that $c(q(X_2)) > 0$ for $X_2 \in A$. Let $r_{q(X_2)}U_1$ be the stochastic representation of $X_1 | X_2$. Since $r^2_{q(X_2)} \approx tr((X_1 | X_2)' (X_1 | X_2))$, from (6.10), we get

$$r^2_{q(X_2)} \sim c(q(X_2))\chi^2_{qn}. \tag{6.11}$$

The p.d.f. of χ^2_{qn} is

$$g(y) = \frac{1}{2^{\frac{qn}{2}}\Gamma\left(\frac{qn}{2}\right)} y^{\frac{qn}{2}-1} e^{-\frac{y}{2}}, \quad y > 0. \tag{6.12}$$

Let $v^2 = c(q(X_2))y$ for fixed X_2 with $v \geq 0$ and $J(y \to v) = \frac{1}{c(q(X_2))} 2v$. Then, $v \approx r_{q(X_2)}$, hence the p.d.f. of v, say p(v), is the same as the p.d.f. of $r_{q(X_2)}$. Therefore, from (6.12) we get

$$p(v) = \frac{2v}{2^{\frac{qn}{2}}\Gamma\left(\frac{qn}{2}\right)c(q(X_2))} \left(\frac{v^2}{c(q(X_2))}\right)^{\frac{qn}{2}-1} e^{-\frac{v^2}{2c(q(X_2))}}$$

$$= \frac{1}{\Gamma\left(\frac{qn}{2}\right)2^{\frac{qn}{2}-1}(c(q(X_2)))^{\frac{qn}{2}}} v^{qn-1} e^{-\frac{v^2}{2c(q(X_2))}}. \tag{6.13}$$

Let rU be the stochastic representation of X and F be the distribution function of r. By appealing to Corollary 2.6.4.1, from (6.13), we obtain

$$1 - F(z) = K_{q(X_2)} \int_{\sqrt{z^2-q(X_2)}}^{\infty} (v_2 + q(X_2))^{\frac{pn}{2}-1} v^{-(qn-2)} p(v)dv$$

$$= L_{X_2} \int\limits_{\sqrt{z^2-q(X_2)}}^{\infty} (v^2 + q(X_2))^{\frac{pn}{2}-1}\, v\, e^{-\frac{v^2}{2c(q(X_2))}}\, dv, \tag{6.14}$$

for $z \geq q(X_2)$ where

$$L_{X_2} = \frac{K_{q(X_2)}}{\Gamma\left(\frac{qn}{2}\right) 2^{\frac{qn}{2}-1} (c(q(X_2)))^{\frac{qn}{2}}}.$$

Substituting $t^2 = v^2 + q(X_2)$ and $J(v \to t) = \frac{t}{v}$ in (6.14), we get

$$1 - F(z) = L_{X_2} \int\limits_{z}^{\infty} t^{pn-2}\, t\, e^{-\frac{t^2-q(X_2)}{2c(q(X_2))}}\, dt$$

$$= L_{X_2}\, e^{\frac{q(X_2)}{2c(q(X_2))}} \int\limits_{z}^{\infty} t^{pn-1}\, e^{-\frac{t^2}{2c(q(X_2))}}\, dt. \tag{6.15}$$

Now, (6.15) holds for $z \geq q(X_2)$. Differentiating (6.15), with respect to z, we get

$$F'(z) = J_{X_2}\, z^{pn-1}\, e^{-\frac{z^2}{2c(q(X_2))}} \quad \text{for } z \geq q(X_2), \tag{6.16}$$

where $J_{X_2} = -L_{X_2}\, e^{\frac{q(X_2)}{2c(q(X_2))}}$.

Let $X_{2,1} \in A$ and $X_{2,2} \in A$ and let $z \in [\max(q(X_{2,1}),q(X_{2,2})),\infty)$. Then, from (6.16), it follows that $c(q(X_{2,1})) = c(q(X_{2,2}))$. Therefore, (6.10) shows that $\phi_{q(X_{2,1})} = \phi_{q(X_{2,2})}$. Since $P(A) = 1$, this means that $\phi_{q(X_2)}$ does not depend on X_2, with probability one. Hence, from Theorem 6.2.5, it is easily seen that X is normally distributed. ■

The next characterization result shows that if a m.e.c. distribution is absolutely continuous and one of its marginals has the p.d.f. whose functional form coincides with that of the p.d.f. of the original distribution up to a constant multiplier, then the distribution must be normal. More precisely, we can prove the following theorem.

THEOREM 6.2.7. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ *be absolutely continuous with p.d.f.*

$$|\Sigma|^{-\frac{n}{2}} |\Phi|^{-\frac{p}{2}} h(\mathrm{tr}((X - M)'\Sigma^{-1}(X - M)\Phi^{-1})).$$

Let X, M *and* Σ *be partitioned as* $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, $M = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}$, *and* $\Sigma = \begin{pmatrix} \Sigma_{11} \Sigma_{12} \\ \Sigma_{21} \Sigma_{22} \end{pmatrix}$, *where* X_1 *and* M_1 *are* $q \times n$, Σ_{11} *is* $q \times q$ *with* $1 \le q < p$. *Let the p.d.f. of* X_1 *be*

$$|\Sigma_{11}|^{-\frac{n}{2}} |\Phi|^{-\frac{q}{2}} h_1(\mathrm{tr}((X_1 - M_1)'\Sigma_{11}^{-1}(X_1 - M_1)\Phi^{-1})).$$

Then, h *and* h_1 *agree up to a constant multiplier; that is,*

$$h(z) = ch_1(z) \tag{6.17}$$

if and only if X *is normal.*

PROOF: If X is normal, then $h(z) = (2\pi)^{-\frac{pn}{2}} \exp\left(-\frac{z}{2}\right)$. Moreover, X_1 is also normal with $h_1(z) = (2\pi)^{-\frac{qn}{2}} \exp\left(-\frac{z}{2}\right)$. Thus, $h(z) = (2\pi)^{\frac{(q-p)n}{2}} h_1(z)$, so (6.17) is satisfied with $c = (2\pi)^{\frac{(q-p)n}{2}}$.

Conversely, assume (6.17) holds. Without loss of generality, we can assume $M = O$, $\Sigma = I_p$, $\Phi = I_n$. From (6.17), we get

$$h(\mathrm{tr}(XX')) = c\, h_1(\mathrm{tr}(X_1X_1')). \tag{6.18}$$

We also use the fact that $h_1(\mathrm{tr}(X_1X_1'))$, being marginal p.d.f. of X, can be obtained from $h(\mathrm{tr}(XX'))$ through integration.

Let $X = (X_1, X_2)$, then $h(\mathrm{tr}(XX')) = h(\mathrm{tr}(X_1X_1' + X_2X_2'))$ $= h(\mathrm{tr}(X_1X_1') + \mathrm{tr}(X_2X_2'))$. Hence we have

$$h_1(\mathrm{tr}(X_1X_1')) = \int_{\mathbb{R}^{(p-q)\times n}} h(\mathrm{tr}(X_1X_1') + \mathrm{tr}(X_2X_2'))dX_2. \tag{6.19}$$

From (6.18) and (6.19), we get

$$h_1(\mathrm{tr}(X_1X_1')) = c \int_{\mathbb{R}^{(p-q)\times n}} h_1(\mathrm{tr}(X_1X_1') + \mathrm{tr}(X_2X_2'))dX_2.$$

Hence

$$h_1(z) = c \int_{\mathbb{R}^{(p-q)\times n}} h_1(z + \mathrm{tr}(X_2X_2'))dX_2, \quad z \geq 0.$$

Using (6.17) again, we have

$$h(z) = c^2 \int_{\mathbb{R}^{(p-q)\times n}} h_1(z + \mathrm{tr}(X_2X_2'))dX_2, \quad z \geq 0$$

which can also be written as

$$h(z) = c^2 \int_{\mathbb{R}^{(p-q)\times n}} h_1(z + \mathrm{tr}(Y_2Y_2'))dY_2. \tag{6.20}$$

Define the $(p + q) \times n$ matrix $Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$, with Y_1: $p \times n$. Let $z = \mathrm{tr}(Y_1Y_1')$. Then, from (6.20), we have

$$h(\mathrm{tr}(Y_1Y_1')) = c^2 \int_{\mathbb{R}^{(p-q)\times n}} h_1(\mathrm{tr}(Y_1Y_1') + \mathrm{tr}(Y_2Y_2'))dY_2. \tag{6.21}$$

Now, the left-hand side of (6.21) is a p.d.f., since Y_1 is $p \times n$. Hence,

$\int_{\mathbb{R}^{p\times n}} h(\mathrm{tr}(Y_1Y_1'))dY_1 = 1$. Therefore, integrating the right-hand side of (6.21), with respect to Y_1 we get

$$c^2 \int_{\mathbb{R}^{p\times n}} \int_{\mathbb{R}^{(p-q)\times n}} h_1(\mathrm{tr}(Y_1Y_1') + \mathrm{tr}(Y_2Y_2'))dY_2dY_1 = 1,$$

which can be rewritten as

$$\int_{\mathbb{R}^{(p+(p-q))\times n}} c^2h_1(\mathrm{tr}(Y))dY = 1.$$

Therefore, $c^2h_1(\mathrm{tr}(Y))$ is the p.d.f. of a $(2p - q) \times n$ random matrix Y with m.e.c. distribution. Moreover, it follows from (6.21) and (6.17), that $c\,h_1(\mathrm{tr}(Y_1Y_1'))$ is the p.d.f. of the $p \times n$ dimensional marginal of Y. Since $Y_1 \sim E_{p,n}(O,I_p \otimes I_n,\psi)$, we must have

$$Y \sim E_{2p-q,n}(O,I_{2p-q} \otimes I_n,\psi).$$

By iterating the above procedure we see that for any $j \geq 1$, there exists a $(p + j(p - q)) \times n$ random matrix Y_j with m.e.c. distribution, such that

$$Y_j \sim E_{p+j(p-q),n}(O,I_{p+j(p-q)} \otimes I_n,\psi).$$

Then it follows from the Definition 4.1.1, Theorem 4.1.3, and Remark 4.1.2, that there exists a distribution function G(u) on $(0,\infty)$, such that

$$\psi(s) = \int_0^\infty \exp\left(-\frac{sv}{2}\right) dG(u). \tag{6.22}$$

Therefore,

$$h(z) = \int_0^\infty \frac{1}{(2\pi z)^{\frac{pn}{2}}} \exp\left(-\frac{zu}{2}\right) dG(u)$$

and

$$h_1(z) = \int_0^\infty \frac{1}{(2\pi z)^{\frac{qn}{2}}} \exp\left(-\frac{zu}{2}\right) dG(u).$$

Using (6.17), we get

$$\int_0^\infty \left(\frac{1}{(2\pi z)^{\frac{pn}{2}}} - \frac{c}{(2\pi z)^{\frac{qn}{2}}}\right) \exp\left(-\frac{zu}{2}\right) dG(u) = 0. \tag{6.23}$$

Using the inverse Laplace transform in (6.23), we obtain

$$\left(\frac{1}{(2\pi z)^{\frac{pn}{2}}} - \frac{c}{(2\pi z)^{\frac{qn}{2}}}\right) dG(z) = 0.$$

Hence,

$$(2\pi z)^{-\frac{pn}{2}}\left(1 - (2\pi z)^{\frac{(p-q)n}{2}} c\right) dG(z) = 0.$$

However, this is possible only if G is degenerate at $z_0 = \frac{1}{2\pi} c^{\frac{2}{(p-q)n}}$; that is,

$$G(z) = \begin{cases} 0 & \text{if } z < z_0 \\ 1 & \text{if } z \geq z_0. \end{cases}$$

Now, let $z_0 = \sigma^2$. Then, (6.22) gives $\psi(s) = \exp\left(-\frac{s\sigma^2}{2}\right)$, which implies that X is normal. ∎

Since in the normal case $\psi_{q(X_2)}$ does not depend on X_2, the conditional distribution $r_{q(X_2)} \mid X_2$ is also independent of X_2. Hence, for every k positive integer the conditional moment $\mathcal{E}(r^k_{q(X_2)} \mid X_2)$, is independent of X_2. The next theorem shows that normal distribution is the only m.e.c. distribution which possesses this property.

THEOREM 6.2.8. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \Psi)$ *be nondegenerate. Let* X *and* Σ *be partitioned as* $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, *where* X_1 *is* $q \times n$ *and* Σ_{11} *is* $q \times q$. *Assume* $\mathrm{rk}(\Sigma_{22}) \geq 1$ *and* $\mathrm{rk}(\Sigma) - \mathrm{rk}(\Sigma_{22}) \geq 1$. *Then, there exists positive integer* k *such that* $\mathcal{E}(r^k_{q(X_2)} \mid X_2)$ *is finite and does not depend on* X_2, *with probability one if and only if* X *is normal. Here,* $r_{q(X_2)}$ *is the one-dimensional random variable, appearing in the stochastic representation of* $X_1 \mid X_2$:

$$X_1 \mid X_2 \approx r_{q(X_2)} A U_1 B'.$$

PROOF: If X is normal, then $\psi_{q(X_2)}(z) = \psi\left(-\frac{z}{2}\right)$. Hence, $r^2_{q(X_2)} \sim \chi^2_{q_1 n}$, with $q_1 = \mathrm{rk}(\Sigma) - \mathrm{rk}(\Sigma_{22})$. Hence, $\mathcal{E}(r^k_{q(X_2)} \mid X_2)$ is the $\frac{k}{2}$ th moment of $\chi^2_{q_1 n}$ which is finite and independent of X_2.

Conversely, assume $\mathcal{E}(r^k_{q(X_2)} \mid X_2)$ is finite and does not depend on X_2, with probability one. Without loss of generality, we can assume $M = O$, $\Sigma = I_p$, and $\Phi = I_n$. Then, we have $q(X_2) = \|X_2\|$ and $X_1 \mid X_2 \approx r_{\|X_2\|^2} U_1$. Hence we get $\operatorname{tr}(X_1'X_1) \mid X_2 \approx r^2_{\|X_2\|^2} \operatorname{tr}(U_1'U_1)$, and since $\operatorname{tr}(U_1'U_1) = 1$, we get $\|X_1\| \mid X_2 \approx r_{\|X_2\|^2}$. Therefore,

$$\mathcal{E}(\|X_1\|^k \mid X_2) = \mathcal{E}(r^k_{\|X_2\|^2} \mid X_2).$$

Hence, $\mathcal{E}(\|X_1\|^k \mid X_2)$ is finite (with probability one) and independent of X_2.

Next, we show that $P(X = O) = 0$. Assume this is not the case. Let $0 < P_0 = P(X = O) = P\left(\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = O\right)$. Then $P((X_1 \mid X_2) = O) \geq P_0$ and $P(\mathcal{E}(\|X_1\|^k \mid X_2) = 0) \geq P_0$. Since $\mathcal{E}(X_1^k \mid X_2)$ does not depend on X_2 with probability one, we have $P(\mathcal{E}(\|X_1\|^k \mid X_2) = 0) = 1$. Hence, $P((X_1 \mid X_2) = O) = 1$.

Since,

$$P(X_1 \in B) = \int_{\mathbb{R}^{q \times n}} P((X_1 \mid X_2 \in B)\, dF_{X_2}(X_2),$$

where F_{X_2} is the distribution function of X_2 and $B \in \mathcal{B}(\mathbb{R}^{q_1 \times n})$, we get $P(X_1 = O) = 1$. Then, from Corollary 3.1.5.1, we get $P(X = O) = P(X_1 = O) = 1$. That means X is degenerate, which contradicts the assumptions of the theorem. So, $P(X = O) = 0$.

From Corollary 3.1.5.1, it follows that $P(X_2 = O) = 0$. Let $X \approx rU$ be the stochastic representation of X and F be the distribution function of r. Then, using Theorem 2.6.4, we obtain

$$P((X_1 \mid X_2) = O) = 1 \text{ if } F(\|X_2\|) = 1$$

and

$$F_{\|X_2\|^2}(z) = \frac{1}{K_{\|X_2\|^2}} \int_{(\|X_2\|, \sqrt{z^2+\|X_2\|^2}]} (s^2 - \|X_2\|^2)^{\frac{qn}{2}-1} s^{-(pn-2)} dF(s) \tag{6.24}$$

if $z \geq 0$ and $F(\|X_2\|) < 1$, where $F_{\|X_2\|^2}(z)$ denotes the distribution function of $r_{\|X_2\|^2}$ and $K_{\|X_2\|^2} = \int_{(\|X_2\|, \infty)} (s^2 - \|X_2\|^2)^{\frac{qn}{2}-1} s^{-(pn-2)} dF(s)$. From (6.24), we get

$$dF_{\|X_2\|^2}(z) = \frac{1}{K_{\|X_2\|^2}} (z^2)^{\frac{qn}{2}-1} (z^2 + \|X_2\|^2)^{-\frac{pn-2}{2}} \frac{ds}{dz} dF(s), \tag{6.25}$$

where $z^2 + \|X_2\|^2 = s^2$. Using (6.25), we obtain

$$\mathcal{E}(r^k_{\|X_2\|^2} | X_2) = \int_0^\infty z^k dF_{\|X_2\|^2}(z)$$

$$= \frac{1}{K_{\|X_2\|^2}} \int_{(\|X_2\|, \infty)} (s^2 - \|X_2\|^2)^{\frac{qn+k}{2}-1} s^{-(pn-2)} dF(s).$$

Since, $\mathcal{E}(r^k_{\|X_2\|^2} | X_2)$ does not depend on X_2, it follows that there exists a constant $c(k)$ which does not depend on X_2, such that

$$\int_{(\|X_2\|, \infty)} (s^2 - \|X_2\|^2)^{\frac{qn+k}{2}-1} s^{-(pn-2)} dF(s)$$

$$= c(k) \int_{(\|X_2\|,\infty)} (s^2 - \|X_2\|^2)^{\frac{qn}{2}-1} s^{-(pn-2)} dF(s), \tag{6.26}$$

almost everywhere. Since $P(X = O) = 0$, Theorem 3.1.5 shows that X_2 is absolutely continuous. Therefore, $\|X_2\|$ is also absolutely continuous. Hence, (6.26) implies that

$$\int_{(y,\infty)} (s^2 - y^2)^{\frac{qn+k}{2}-1} s^{-(pn-2)} dF(s) = c(k) \int_{(y,\infty)} (s^2 - y^2)^{\frac{qn}{2}-1} s^{-(pn-2)} dF(s) \tag{6.27}$$

for almost every y (with respect to the Lebesgue measure) in the interval $(0,y_0)$, where $y_0 = \inf_y\{y\colon F(y) = 1\}$. Furthermore, (6.27) is also true if $y > y_0$, because in that case both sides become zero. Therefore, (6.27) holds for almost every $y > 0$.

Define the following distribution function on $[0,\infty)$

$$H(z) = \frac{\int_0^z s^k dF(s)}{\int_0^\infty s^k dF(s)}. \tag{6.28}$$

In order to do this, we have to prove that $\int_0^\infty s^k dF(s)$ is finite and positive.

Now, F is the distribution function of r, where $X \approx rU$ is the stochastic representation of X. Since, $P(X = O) = 0$, we have $P(r > 0) = 1$. Hence, $\int_0^\infty s^k dF(s) = \mathcal{E}(r^k) > 0$. On the other hand, let $X_1 \approx r_1U_1$ be the stochastic representation of X_1. It follows from Corollary 2.5.6.1, that

$$r_1 \approx rt, \tag{6.29}$$

where r and t are independent and $t^2 \sim B\left(\frac{qn}{2}, \frac{(p-q)n}{2}\right)$. Now, $\mathcal{E}(\|X_1\|^k) = \mathcal{E}(\mathcal{E}(\|X_1\|^k \mid X_2))$ is finite, since $\mathcal{E}(\|X_1\|^k \mid X_2)$ is finite and does not depend on X_2, with probability one. Since $r_1 \approx \|X_1\|$, we see that $\mathcal{E}(r_1^k)$ is finite. From (6.29), we get $\mathcal{E}(r_1^k) = \mathcal{E}(r^k)\,\mathcal{E}(t^k)$ and $\mathcal{E}(r_1^k)$ is finite implies that $\mathcal{E}(r^k)$ is finite. Since $P(X = O) = 0$, we have $F(0) = 0$ and so $H(0) = 0$. Now, (6.27) can be rewritten in terms of H as

$$\int_y^\infty (s^2 - y^2)^{\frac{qn+k}{2}-1} s^{-(pn+k-2)}\, dH(s) = c \int_y^\infty (s^2 - y^2)^{\frac{qn}{2}-1} s^{-(pn+k-2)}\, dH(s). \tag{6.30}$$

Let r_0 be a random variable with distribution function H. Further, let $\mathbf{u}_0$ be uniformly distributed over S_{pn+k}, independent of r_0. Define $\mathbf{y} = r_0\mathbf{u}_0$. Then, $\mathbf{y} \sim E_{pn+k}(\mathbf{o}, I_{pn+k}, \psi^*)$. Since $H(0) = 0$, we get $P(r_0 = 0) = 0$. Thus, $P(\mathbf{y} = 0) = 0$. Let $\mathbf{y}_1$ be a $(p - q)n + k$-dimensional subvector of $\mathbf{y}$ and $\mathbf{y}_2$, a $(p - q)n$-dimensional subvector of $\mathbf{y}_1$. Then, it follows from Theorem 3.1.1, that both $\mathbf{y}_1$ and $\mathbf{y}_2$ are absolutely continuous. Let $h_1(\mathbf{y}_1'\mathbf{y}_1)$ be the p.d.f. of $\mathbf{y}_1$ and $h_2(\mathbf{y}_2'\mathbf{y}_2)$ that of $\mathbf{y}_2$. Let $\mathbf{y}_1 \approx r_{0,1}\mathbf{U}_{0,1}$ and $\mathbf{y}_2 \approx r_{0,2}\mathbf{U}_{0,2}$ be the stochastic representations of $\mathbf{y}_1$ and $\mathbf{y}_2$ respectively. Moreover, Theorem 3.1.1 shows that the p.d.f. of $r_{0,1}$ is

$$g_1(y) = c_1\, y^{(p-q)n+k-1} \int_y^\infty (s^2 - y^2)^{\frac{qn}{2}-1} s^{-(pn+k-2)}\, dH(s)$$

and that of $r_{0,2}$ is

$$g_2(y) = c_2\, y^{(p-q)n-1} \int_y^\infty (s^2 - y^2)^{\frac{qn+k}{2}-1} s^{-(pn+k-2)}\, dH(s).$$

Here c_i denotes a positive constant, $i = 1,2,\ldots$. It follows, from Theorem 2.5.5, that

$$h_1(y^2) = c_3\, y^{-((p-q)n+k)+1}\, g_1(y)$$
$$= c_4 \int_y^{\infty} (s^2 - y^2)^{\frac{qn}{2}-1}\, s^{-(pn+k-2)}\, dH(s)$$

and

$$h_2(y^2) = c_5\, y^{-(p-q)n+1}\, g_2(y)$$
$$= c_6 \int_y^{\infty} (s^2 - y^2)^{\frac{qn+k}{2}-1}\, s^{-(pn+k-2)}\, dH(s).$$

Hence,

$$h_1(y^2) = c_4 \int_y^{\infty} (s^2 - y^2)^{\frac{qn}{2}-1}\, s^{-(pn-2)}\, dF(s) \tag{6.31}$$

and

$$h_2(y^2) = c_6 \int_y^{\infty} (s^2 - y^2)^{\frac{qn+k}{2}-1}\, s^{-(pn-2)}\, dF(s). \tag{6.32}$$

Then, (6.27), (6.30), and (6.31) imply that $h_2(y^2) = c_7\, h_1(y^2)$. Therefore, from Theorem 6.2.7 we conclude that y_1 is normal. Since, y_1 is a subvector of y, it is also normal. Hence, $H(z)$ is the distribution of $c_8\, \chi_{pn+k}$. Then,

$$dH(z) = \ell(z)\, dz, \tag{6.33}$$

where

$$\ell(z) = c_9 \left(\frac{z}{c_8}\right)^{pn+k-1} e^{-\frac{z^2}{2c_8^2}}, \quad z \geq 0. \tag{6.34}$$

From (6.28), it follows that $dH(z) = c_{10}\, z^k\, dF(z)$. Using (6.33) and (6.34), we obtain $c_{11}\, z^{pn+k-1}\, e^{-\frac{z}{2c_8^2}}\, dz = z^k\, dF(z)$. Hence,

$$dF(z) = c_{12}\left(\frac{z}{c_7}\right)^{pn-1} e^{-\frac{z^2}{2c_8^2}},\ z \geq 0.$$

Since F is the distribution of r, we obtain that $r \approx c_{12}\, \chi_{pn}$. Therefore, X is normal. ■

The following theorem gives a characterization of normality based on conditional central moments.

THEOREM 6.2.9. *Let* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ *be nondegenerate. Let* X *and* Σ *be partitioned as* $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, $M = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}$, *and* $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, *where* X_1 *and* M_1 *are* $q \times n$ *and* Σ_{11} *is* $q \times q$. *Assume* $rk(\Sigma_{22}) \geq 1$ *and* $rk(\Sigma) - rk(\Sigma_{22}) \geq 1$. *Assume also that there exist nonnegative integers,* k_{ij}, $i = 1,2,\ldots,q$; $j = 1,2,\ldots,n$, *satisfying* $k = \sum_{i=1}^{p} \sum_{j=1}^{n} k_{ij} \geq 1$, *and such that*

$K(X_2) = \mathcal{E}\left(\prod_{i=1}^{q} \prod_{j=1}^{n} (x_{ij} - \mathcal{E}(x_{ij} \mid X_2))^{k_{ij}} \mid X_2\right)$ *is nonzero with positive probability.*

Then $K(X_2)$ *is finite and does not depend on* X_2, *with probability one if and only if* X *is normally distributed.*

PROOF: If X is normal, then $X_1 \mid X_2 \sim N_{q,n}(M_1 + \Sigma_{12}\Sigma_{22}^{-}(X_2 - M_2), \Sigma_{11\cdot 2} \otimes \Phi)$. Hence

$$(X_1 - \mathcal{E}(X_1 \mid X_2)) \mid X_2 \sim N_{q,n}(O, \Sigma_{11\cdot 2} \otimes \Phi).$$

Therefore, $\mathcal{E}\left(\prod_{i=1}^{q} \prod_{j=1}^{n} (x_{ij} - E(x_{ij} \mid X_2))^{k_{ij}} \mid X_2\right)$ is finite and does not depend on X_2.

Conversely, assume $K(X_2)$ is finite and does not depend on X_2. Now,

$$X_1 \mid X_2 \sim E_{q,n}(M_1 + \Sigma_{12}\Sigma_{22}^{-}(X_2 - M_2), \Sigma_{11\cdot 2} \otimes \Phi, \psi_{q(X_2)}),$$

and hence

$$(X_1 - \mathcal{E}(X_1 \mid X_2)) \mid X_2 \sim E_{q,n}(O, \Sigma_{11\cdot 2} \otimes \Phi, \psi_{q(X_2)}).$$

Let $r_{q(X_2)}AUB'$ be the stochastic representation of $(X_1 - \mathcal{E}(X_1 \mid X_2)) \mid X_2$. Then,

$$\prod_{i=1}^{q} \prod_{j=1}^{n} ((X_{ij} - \mathcal{E}(X_{ij} \mid X_2))^{k_{ij}} \mid X_2) \approx r_{q(X_2)}^{k} \prod_{i=1}^{q} \prod_{j=1}^{n} (AUB')_{ij}. \tag{6.35}$$

The expected value of the left-hand side of (6.35) is finite and does not depend on X_2. Since $r_{q(X_2)}$ and U are independent, taking expectation on both sides of (6.35), we obtain

$$K(X_2) = \mathcal{E}(r_{q(X_2)}^{k} \mid X_2)\, \mathcal{E}\left(\prod_{i=1}^{q} \prod_{j=1}^{n} (A^*UB')_{ij} \right). \tag{6.36}$$

Now $P(0 \neq |K(X_2)| < \infty) > 0$, therefore it follows, from (6.36), that

$$P\left(0 \neq \left| \mathcal{E}\left(\prod_{i=1}^{q} \prod_{j=1}^{n} (A^*UB^*)_{ij} \right) \right| < \infty \right) > 0. \tag{6.37}$$

However $\mathcal{E}\left(\prod_{i=1}^{q} \prod_{j=1}^{n} (A^*UB^*)_{ij} \right)$ is a constant (say c) that does not depend X_2, therefore (6.37) implies that

$$\mathcal{E}\left(\prod_{i=1}^{q}\prod_{j=1}^{n}(A^*UB^*)_{ij}\right)\neq 0. \tag{6.38}$$

From (6.36) and (6.38), we get

$$\mathcal{E}(r_{q(X_2)}^k \mid X_2) = \frac{1}{c}K(X_2).$$

Therefore, $\mathcal{E}(r_{q(X_2)}^k \mid X_2)$ is finite and independent of X_2 with probability one. Then, using Theorem 6.2.8, we conclude that X is normally distributed. ■

Theorems 6.2.5-6.2.9 are due to Cambanis, Huang, and Simons (1981). In Theorem 6.2.9, however, the assumption that the conditional central moments are nonzero, with probability one is missing. Without this assumption, the theorem is not correct. To see this, take $k_{11} = 1$ and $k_{ij} = 0$, $(i,j) \neq (1,1)$. Then

$$\begin{aligned} K(X_2) &= \mathcal{E}((X_{11} - \mathcal{E}(X_{11} \mid X_2)) \mid X_2) \\ &= \mathcal{E}(X_{11} \mid X_2) - \mathcal{E}(X_{11} \mid X_2) \\ &= O. \end{aligned}$$

Thus $K(X_2)$ is finite and does not depend on X_2, but X does not have to be normal.

In order to derive further characterization results, we need the following lemma.

LEMMA 6.2.1. *Let* $X \sim E_{p,n}(O,\Sigma \otimes \Phi,\psi)$ *with* $P(X = O) = P_0$ *where* $0 \leq P_0 < 1$. *Then, there exists a one-dimensional random variable* s *and a* $p \times n$ *random matrix* L, *such that* s *and* L *are independent,* $P(s = 0) = P_0$, $P(s = 1) = 1 - P_0$, $L \sim E_{p,n}\left(O,\Sigma \otimes \Phi, \frac{\psi - P_0}{1 - P_0}\right)$ *and* $X \approx sL$. *Moreover,* $P(L = O) = 0$.

PROOF: If $P_0 = 0$, choose s, such that $P(s = 1) = 1$ and the theorem is trivial. If $P_0 > 0$, then let us define L in the following way. Define a measure P_L on $\mathbb{R}^{p\times n}$ in the following way: $P_L(B) = \frac{1}{1 - P_0} P(X \in (B - \{O\}))$, where B is a Borel set in $\mathbb{R}^{p\times n}$. Now,

$P_L(\mathbb{R}^{p\times n}) = \frac{1}{1 - P_0} P(X \in (\mathbb{R}^{p\times n} - \{O\})) = \frac{1 - P_0}{1 - P_0} = 1$. Therefore, P_L defines a probability measure on $\mathbb{R}^{p\times n}$. Let L be a random matrix, whose distribution is defined by P_L; that is, $P(L \in B) = P_L(B)$ for every Borel set B in $\mathbb{R}^{p\times n}$.

Let s be a one-dimensional random variable, with $P(s = 0) = P_0$, and $P(s = 1) = 1 - P_0$ such that s is independent of L. Then we show $X \approx sL$.

Let B be a Borel set in $\mathbb{R}^{p\times n}$. Then

$$\begin{aligned}
P(sL \in B) &= P(sL \in B \mid s = 1)\, P(s = 1) + P(sL \in B \mid s = 0)\, P(s = 0) \\
&= P(L \in B \mid s = 1)\, P(s = 1) + P(O \in B \mid s = 0)\, P(s = 0) \\
&= P(l \in B)(1 - P_0) + P(O \in B)P_0 \\
&= \frac{1}{1 - P_0} P(X \in (B - \{O\}))(1 - P_0) + \chi_B(O)P_0 \\
&= P(X \in (B - \{O\})) + \chi_B(O)P_0.
\end{aligned}$$

If $O \in B$, then $B = (B - \{O\}) \cup \{O\}$ and $\chi_B(O) = 1$. Therefore

$$\begin{aligned}
P(X \in (B - \{O\})) + \chi_B(O)P_0 &= P(X \in (B - \{O\})) + P(X \in \{O\}) \\
&= P(X \in B).
\end{aligned}$$

If $O \notin B$, then $B = B - \{O\}$ and $\chi_B(O) = 0$. So, $P(X \in (B - \{O\})) + \chi_B(O)P_0 = P(X \in B)$.

Hence, $X \approx sL$. Therefore, the characteristic functions of X and sL are equal, i.e. $\phi_X(T) = \phi_{sL}(T)$. Moreover,

$$\begin{aligned}
\phi_{sL}(T) &= \mathcal{E}(\mathrm{etr}(iT'sL)) \\
&= \mathcal{E}(\mathrm{etr}(iT'sL) \mid s = 1)\, P(s = 1) + \mathcal{E}(\mathrm{etr}(iT'sL) \mid s = 0)\, P(s = 0) \\
&= \mathcal{E}(\mathrm{etr}(iT'L) \mid s = 1)(1 - P_0) + \mathcal{E}(1 \mid s = 0)P_0 \\
&= \mathcal{E}(\mathrm{etr}(iT'L))(1 - P_0) + 1 \cdot P_0 \\
&= \phi_L(T)(1 - P_0) + P_0.
\end{aligned}$$

Hence, $\phi_X(T) = \phi_L(T)(1 - P_0) + P_0$, and therefore $\phi_L(T) = \dfrac{\phi_X(T) - P_0}{1 - P_0} = \dfrac{\psi(\mathrm{tr}(T'\Sigma\Phi)) - P_0}{1 - P_0}$. Furthermore, $P(L = O) = \dfrac{1}{1 - P_0} P(X \in (\{O\} - \{O\}) = 0.$ ■

THEOREM 6.2.10. *Let* $X \sim E_{p,n}(O,\Sigma_1 \otimes I_n,\psi_1)$ *with* $p \leq n$. *Assume* $XX' \sim G_{p,1}\left(\Sigma_2, \frac{n}{2}, \psi_2\right)$. *Then there exists* $c > 0$ *such that* $\Sigma_2 = c\Sigma_1$ *and* $\psi_2(z) = \psi_1\left(\frac{z}{c}\right)$.

PROOF: It follows, from Definition 5.1.1, that there exists a $p \times n$ random matrix Y, such that $Y \sim E_{p,n}(O,\Sigma_2 \otimes I_n,\psi_2)$ and $YY' \sim G_{p,1}\left(\Sigma_2, \frac{n}{2}, \psi_2\right)$. So, we have $XX' \approx YY'$. From Theorem 2.1.4, it suffices to show that $X \approx Y$.

First, note that if Z is a $p \times n$ matrix, then $Z = O$ iff $ZZ' = O$. Therefore, $P(X = O) = P(XX' = O)$ and $P(Y = O) = P(YY' = O)$. Since, $XX' \approx YY'$, we have $P(XX' = O) = P(YY' = O)$ and so, $P(X = O) = P(Y = O)$.

Let us denote $P(X = O)$ by P_0. If $P_0 = 1$, then $P(X = O) = P(Y = O) = 1$, and hence, $X \approx Y$.

If $P_0 < 1$, then from Theorem 1.3.8, there exists a $p \times p$ nonsingular matrix H, such that $H\Sigma_2H' = I_p$ and $H\Sigma_1H' = D$, where D is diagonal.

Let $V_1 = HX$ and $V_2 = HY$. Then, $V_1 \sim E_{p,n}(O,D \otimes I_n,\psi_1)$ and $V_2 \sim E_{p,n}(O,I_p \otimes I_n,\psi_1)$. Moreover, $V_1V_1' = HXX'H'$ and $V_2V_2' = HYY'H'$. So, we have $V_1V_1' \approx V_2V_2'$. It suffices to prove that $V_1 \approx V_2$, because if this is true, then $H_1^{-1}V_1 \approx H^{-1}V_2$ and so, $X \approx Y$.

Using Lemma 6.2.1, we can write $V_i \approx s_iL_i$, where s_i and L_i are independent, s_i is a one-dimensional random variable, with $P(s_i = 0) = P_0$, $P(s_i = 1) = 1 - P_0$, $P(L_i = O) = 0$ $(i = 1,2)$. So $s_1 \approx s_2$. Moreover, we have

$$L_1 \sim E_{p,n}\left(O,D \otimes I_n, \frac{\psi_1 - P_0}{1 - P_0}\right) \text{ and } L_2 \sim E_{p,n}\left(O,I_p \otimes I_n, \frac{\psi_2 - P_0}{1 - P_0}\right).$$

If $P_0 = 0$, then $P(s_1 = 1) = P(s_2 = 1) = 1$ and so, $L_1L_1' \approx s_1^2 L_1L_1' \approx V_1V_1' \approx V_2V_2' \approx s_2^2 L_2L_2' \approx L_2L_2'$. If $0 < P_0 < 1$, then for any Borel set B in $\mathbb{R}^{p\times n}$ we have

$$\begin{aligned} & P(V_iV_i' \in B) \\ &= P(s_iL_iL_i' \in B) \\ &= P(s_iL_iL_i' \in B \mid s_i = 1)P(s_i = 1) + P(s_iL_iL_i' \in B \mid s_i = 0)P(s_i = 0) \\ &= P(L_iL_i' \in B)(1 - P_0) + \chi_B(O)P_0. \end{aligned}$$

Therefore, $P(L_iL_i' \in B) = \dfrac{P(V_iV_i' \in B) - \chi_B(O)P_0}{1 - P_0}$ $(i = 1,2)$. Since, $V_1V_1' \approx V_2V_2'$, we have $P(V_1V_1' \in B) = P(V_2V_2' \in B)$ and so, $P(L_1L_1' \in B) = P(L_2L_2' \in B)$. Hence, $L_1L_1' \approx L_2L_2'$. Let $r_1D^{\frac{1}{2}}U_1$ and r_2U_2 be the stochastic representations of L_1 and L_2, respectively. Then we have $L_1L_1' \approx r_1D^{\frac{1}{2}}U_1U_1'D^{\frac{1}{2}}$ and $L_2L_2' \approx r_2^2U_2U_2'$. Thus,

$$r_1^2D^{\frac{1}{2}}U_1U_1'D^{\frac{1}{2}} \approx r^2U_2U_2'. \tag{6.39}$$

Let $W_1 = U_1U_1'$ and $W_2 = U_2U_2'$. Since, $U_1 \approx U_2$, we have $W_1 \approx W_2$. Note that (6.39) can be rewritten as

$$r_1^2D^{\frac{1}{2}}W_1D^{\frac{1}{2}} \approx r_2^2W_2. \tag{6.40}$$

From Theorem 5.2.2, it follows that $P(\mathrm{rk}(L_1L_1') = p) = 1$. Since, L_1L_1' is a positive semidefinite $p \times p$ matrix, we get $P(L_1L_1' > O) = 1$. Now $L_1L_1' \approx L_2L_2'$, and hence we have $P(L_2L_2' > O) = 1$. Therefore,

$P(r_2^2 W_2 > O) = 1$ and $P(r_1^2 D^{\frac{1}{2}} W_1 D^{\frac{1}{2}} > O) = 1$. Thus, the diagonal elements of $r_2^2 W_2$ and $r_1^2 D^{\frac{1}{2}} W_1 D^{\frac{1}{2}}$ are positive, with probability one. If $p = 1$, then D is a scalar; $D = c$. If $p > 1$, then it follows from (6.40) that

$$\frac{(r_2^2 W_2)_{11}}{(r_2^2 W_2)_{ii}} \approx \frac{(r_1^2 D^{\frac{1}{2}} W_1 D^{\frac{1}{2}})_{11}}{(r_1^2 D^{\frac{1}{2}} W_1 D^{\frac{1}{2}})_{ii}}, \quad i = 2,\ldots,p, \text{ or equivalently,}$$

$$\frac{(W_2)_{11}}{(W_2)_{ii}} \approx \frac{(D^{\frac{1}{2}} W_1 D^{\frac{1}{2}})_{11}}{(D^{\frac{1}{2}} W_1 D^{\frac{1}{2}})_{ii}}. \tag{6.41}$$

However, $D^{\frac{1}{2}}$ is diagonal so, $(D^{\frac{1}{2}} W_1 D^{\frac{1}{2}})_{jj} = (W_1)_{jj}\, d_{jj}$, $j = 1,\ldots,p$ and (6.41) becomes

$$\frac{(W_2)_{11}}{(W_2)_{ii}} \approx \frac{(W_1)_{11}}{(W_1)_{ii}} \frac{d_{11}}{d_{ii}}.$$

Since $W_1 \approx W_2$, we have $\frac{(W_1)_{11}}{(W_1)_{ii}} \approx \frac{(W_2)_{11}}{(W_2)_{ii}}$ and so, $\frac{(W_1)_{11}}{(W_1)_{ii}} \approx \frac{(W_1)_{11}}{(W_1)_{ii}} \frac{d_{11}}{d_{ii}}$. Since $P\left(\frac{(W_1)_{11}}{(W_1)_{ii}} > 0\right) = 1$, this is possible only if $\frac{d_{11}}{d_{ii}} = 1$, $i = 2,\ldots,p$. So, we get $D = c\, I_p$ where c is a scalar constant. From (6.39), we get

$$c\, r_1^2\, U_1 U_1' \approx r_2^2 U_2 U_2'.$$

Taking trace on both sides, we get

$$\operatorname{tr}(c\, r_1^2 U_1 U_1') \approx \operatorname{tr}(r_2^2 U_2 U_2')$$

and hence,

$$c\, r_1^2\, \mathrm{tr}(U_1U_1') \approx r_2^2 \mathrm{tr}(U_2U_2').$$

Now, $\mathrm{tr}(U_1U_1') = \mathrm{tr}(U_2U_2') = 1$ and therefore, $c\, r_1^2 \approx r_2^2$ and $r_2 \approx \sqrt{c}\, r_1$. Let $r_3 \approx r_2$, such that r_3 is independent of U_1 and U_2. Then, we have $L_1 \approx r_1 D^{\frac{1}{2}} U_1 \approx \frac{1}{\sqrt{c}} r_3 (cI_p)^{\frac{1}{2}} U_1 = r_3U_1 \approx r_3U_2$. Since $L_2 \approx r_2U_2 \approx r_3U_2$, we have $L_1 \approx L_2$. Since $s_1 \approx s_2$, we get $s_1L_1 \approx s_2L_2$. Therefore $V_1 \approx V_2$. ■

Now we can prove a result on the characterization of normality.

THEOREM 6.2.11. *Let* $X \sim E_{p,n}(O,\Sigma_1 \otimes I_n,\psi_1)$, with $p \le n$. Assume $XX' \sim W_p(\Sigma_2,n)$. *Then,* $X \sim N_{p,n}(O,\Sigma_1 \otimes I_n)$ *and* $\Sigma_1 = \Sigma_2$.

PROOF: The result follows immediately by taking $\psi_2(z) = \exp\left(-\frac{z}{2}\right)$, in Theorem 6.2.10. ■

The following theorem is an extension of Theorem 6.2.10.

THEOREM 6.2.12. *Let* $X \sim E_{p,n}(O,\Sigma_1 \otimes \Phi,\psi_1)$ *and* $Y \sim E_{p,n}(O,\Sigma_2 \otimes \Phi,\psi_2)$, *with* $\Sigma_1 > O$, $\Sigma_2 \ge O$, *and* $\Phi > O$. Assume $XAX' \approx YAY'$ *where* A *is an* $n \times n$ *positive semidefinite matrix, with* $\mathrm{rk}(A) \ge p$. *Suppose that all moments of* X *exist. Let* $x = \mathrm{tr}(X\Sigma_1^{-1}X\Phi^{-1})$ *and define* $m_k = \mathcal{E}(x^k)$, $k = 1,2,\ldots$. *If* $\sum_{k=1}^{\infty} \left(\frac{1}{m_{2k}}\right)^{\frac{1}{2k}} = \infty$, *then* $X \approx Y$.

PROOF: Without loss of generality, we can assume that $\Phi = I_n$. Indeed, if $\Phi \ne I_n$ then, define $X_1 = X\Phi^{-\frac{1}{2}}$, $Y_1 = Y\Phi^{-\frac{1}{2}}$, and $A_1\ \Phi^{\frac{1}{2}}A\Phi^{\frac{1}{2}}$. Then, $X_1 \sim E_{p,n}(O,\Sigma_1 \otimes I_n,\psi_1)$, $Y_1 \sim E_{p,n}(O,\Sigma_2 \times I_n,\psi_2)$, $\mathrm{rk}(A_1) = \mathrm{rk}(A) \ge p$, $XAX' = X_1A_1X_1'$, and $YAY' = Y_1A_1Y_1'$. So $XAX' \approx YAY'$ if and only if $X_1A_1X_1' \approx Y_1A_1Y_1'$ and $X \approx Y$ if and only if $X_1 \approx Y_1$.

Hence, we can assume $\Phi = I_n$. From Theorem 1.3.8, there exists a $p \times p$ nonsingular matrix H such that $H\Sigma_1H' = I_p$ and $H\Sigma_2H' = D$, where D is diagonal. Let $Z_1 = HX$ and $Z_2 = HY$. Then, $Z_1 \sim E_{p,n}(O,I_p \otimes I_n,\psi_1)$ and

$Z_2 \sim E_{p,n}(O, D \otimes I_n, \psi_2)$. Moreover, $Z_1AZ_1' = HXAX'H'$ and $Z_2AZ_2' = HYAY'H'$. Since $XAX' \approx YAY'$, we have $Z_1AZ_1' \approx Z_2AZ_2'$. It suffices to prove that $Z_1 \approx Z_2$, because if this is true, then $H^{-1}Z_1 \approx H^{-1}Z_2$ and therefore $X \approx Y$. Since $Z_1AZ_1' \approx Z_2AZ_2'$, we get$P(Z_1AZ_1' = O) = P(Z_2AZ_2' = O)$.

If $P(Z_1 = O) < 1$ then, using Lemma 6.2.1, we can write $Z_1 = s_1L_1$ where s_1 and L_1 are independent, s_1 is a one-dimensional random variable with $P(s_1 = 0) = P(Z_1 = O)$, $P(s_1 = 1) = 1 - P(Z_1 = O)$, $L_1 \sim E_{p,n}\left(O, I_p \otimes I_n, \frac{\psi - P(Z_1 = O)}{1 - P(Z_1 = O)}\right)$ and $P(L_1 = O) = 0$. Then, from Theorem 5.2.2, it follows that $P(rk(L_1AL_1') = p) = 1$. Since, A is positive definite, this implies that

$$P(L_1AL_1' > O) = 1. \tag{6.42}$$

Consequently $P(L_1AL_1' = O) = 0$. Moreover,

$$\begin{aligned} P(Z_1AZ_1' = O) &= P(s_1^2 L_1AL_1' = O) \\ &= P(s_1^2 = 0) \\ &= P(s_1 = 0) \\ &= P(Z_1 = O). \end{aligned}$$

Hence, $P(Z_1AZ_1' = O) < 1$.

If $P(Z_2 = O) < 1$, then, using Lemma 6.2.1, we can write $Z_2 = s_2L_2$ where s_2 and L_2 are independent, s_2 is a one-dimensional random variable with $P(s_2 = 0) = P(Z_2 = O)$, $P(s_2 = 1) = 1 - P(Z_2 = O)$, $L_2 \sim E_{p,n}\left(O, D \otimes I_n, \frac{\psi - P(Z_2 = O)}{1 - P(Z_2 = O)}\right)$ and $P(L_2 = O) = 0$. Then, from Theorem 5.2.1, it follows that $P(rk(L_2AL_2') = \min(rk(D), p)) = 1$. From $P(Z_2 = 0) < 1$, it follows that $rk(D) \geq 1$. Hence, $P(rk(L_2AL_2') \geq 1) = 1$. Thus, $P(L_2AL_2' = O) = 0$. Therefore,

$$P(Z_2AZ_2' = O) = P(s_2^2 L_2AL_2' = O)$$

$$= P(s_2^2 = 0)$$
$$= P(s_2 = 0)$$
$$= P(Z_2 = O).$$

Hence, $P(Z_2AZ_2' = O) < 1$. Therefore, if either $P(Z_1 = O) < 1$ or $P(Z_2 = O) < 1$, then we get $P(Z_1AZ_1' = O) = P(Z_2AZ_2' = O) < 1$ and hence, $P(Z_i = O) < 1$ must hold for $i = 1,2$. However, then we get

$$\begin{aligned} P(Z_1 = O) &= P(Z_1AZ_1' = O) \\ &= P(Z_2AZ_2' = O) \\ &= P(Z_2 = O) \end{aligned}$$

and $P(s_1 = 0) = P(s_2 = 0)$. Hence, $s_1 \approx s_2$.

If either $P(Z_1 = O) = 1$ or $P(Z_2 = O) = 1$, then we must have $P(Z_i = O) = 1$ for $i = 1,2$. Therefore, $P(Z_1 = O) = P(Z_2 = O)$ is always true. Let $P_p = P(Z_1 = O)$.

If $P_0 = 1$, then $Z_1 = O = Z_2$ and the theorem is proved.

If $P_0 < 1$, then we first prove that

$$L_1AL_1' \approx L_2AL_2'. \tag{6.43}$$

If $P_0 = 0$, then $P(s_1 = 1) = P(s_2 = 1) = 1$ and so,

$$L_1AL_1' \approx s_1L_1AL_1' \approx Z_1AZ_1' \approx Z_2AZ_2' \approx s_2L_2AL_2' \approx L_2AL_2'.$$

If $0 < P_0 < 1$, then for any Borel set B in $\mathbb{R}^{p\times n}$, we have

$$\begin{aligned} &P(Z_1AZ_1' \in B) \\ &= P(s_1L_1AL_1' \in B) \\ &= P(s_1L_1AL_1' \in B \mid s_1 = 1)P(s_1 = 1) + P(s_1L_1AL_1' \in B \mid s_1 = 0)P(s_1 = 0) \\ &= P(L_1AL_1' \in B)(1 - P_0) + \chi_B(O)P_0. \end{aligned}$$

Therefore, $P(L_1AL_1' \in B) = \dfrac{P(Z_1AZ_1' \in B) - \chi_B(O)P_0}{1 - P_0}$.

Similarly, $P(L_2AL_2' \in B) = \dfrac{P(Z_2AZ_2' \in B) - \chi_B(O)P_0}{1 - P_0}$. Since, $Z_1AZ_1' \approx Z_2AZ_2'$, we get $P(Z_1AZ_1' \in B) = P(Z_2AZ_2' \in B)$ and so, $P(L_1AL_1' \in B) = P(L_2AL_2' \in B)$. Therefore, $L_1AL_1' \approx L_2AL_2'$ which establishes (6.43).

Let r_1U_1 and $r_2D^{\frac{1}{2}}U_2$ be the stochastic representations of L_1 and L_2. Here $D^{\frac{1}{2}}$ is diagonal. Then, we have $L_1AL_1' \approx r_1^2U_1AU_1'$ and

$L_2AL_2' \approx r_2^2D^{\frac{1}{2}}U_2AU_2'D^{\frac{1}{2}}$. Hence,

$$r_1^2U_1AU_1' \approx r_2^2D^{\frac{1}{2}}U_2AU_2'D^{\frac{1}{2}}. \tag{6.44}$$

Let $W_1 = U_1AU_1'$ and $W_2 = U_2AU_2'$. Since $U_1 \approx U_2$, we have $W_1 \approx W_2$. So, (6.44) can be rewritten as

$$r_1^2W_1 \approx r_2^2D^{\frac{1}{2}}W_2D^{\frac{1}{2}}. \tag{6.45}$$

If $p = 1$, then D is a scalar; $D = c\,I_1$. If $p > 1$, then from (6.42) it follows that the diagonal elements of L_1AL_1' are positive, with probability one. From (6.43), it is seen that $P(L_2AL_2' > O) = 1$ and the diagonal elements of L_2AL_2' are also positive with probability one. Using (6.45), we obtain

$$\frac{(r_1^2W_1)_{11}}{(r_1^2W_1)_{ii}} \approx \frac{(r_2^2D^{\frac{1}{2}}W_2D^{\frac{1}{2}})_{11}}{(r_2^2D^{\frac{1}{2}}W_2D^{\frac{1}{2}})_{ii}}, \quad i = 2,\dots,p,$$

or equivalently,

$$\frac{(W_1)_{11}}{(W_1)_{ii}} \approx \frac{(D^{\frac{1}{2}}W_2D^{\frac{1}{2}})_{11}}{(D^{\frac{1}{2}}W_2D^{\frac{1}{2}})_{ii}}. \tag{6.46}$$

However $D^{\frac{1}{2}}$ is diagonal and hence, $(D^{\frac{1}{2}}W_2D^{\frac{1}{2}})_{jj} = (W_2)_{jj}\, d_{jj}$, $j = 1,\ldots,p$. Therefore, (6.46) becomes

$$\frac{(W_1)_{11}}{(W_1)_{ii}} \approx \frac{(W_2)_{11}}{(W_2)_{ii}}\frac{d_{11}}{d_{ii}}.$$

Since $W_1 \approx W_2$, we have $\frac{(W_1)_{11}}{(W_1)_{ii}} \approx \frac{(W_2)_{11}}{(W_2)_{ii}}$ and so, $\frac{(W_2)_{11}}{(W_2)_{ii}} \approx \frac{(W_2)_{11}}{(W_2)_{ii}}\frac{d_{11}}{d_{ii}}$. Now $P\left(\frac{(W_2)_{11}}{(W_2)_{ii}} > 0\right) = 1$, which is possible only if $\frac{d_{11}}{d_{ii}} = 1$, $i = 2,\ldots,p$. Consequently we get $D = cI_p$, where c is a scalar constant. From (6.45), we get

$$r_1^2 W_1 \approx r_2^2\, cW_2.$$

Taking trace on both sides, we have $\operatorname{tr}(r_1^2 W_1) \approx \operatorname{tr}(r_2^2 cW_2)$, and hence,

$$r_1^2 \operatorname{tr}(W_1) \approx r_2^2\, c \operatorname{tr}(W_2). \tag{6.47}$$

Since, $\operatorname{tr}(U_1'U_1) = 1$, all the elements of U_1 are less than 1. Therefore, there exists a positive constant K such that $\operatorname{tr}(U_1AU_1') < K$. From Theorem 5.2.1, it follows that $P(\operatorname{rk}(U_1AU_1') = p) = 1$. Consequently $U_1AU_1' > O$ with probability one. Therefore, $\mathcal{E}((\operatorname{tr}(W_1))^k)$ is a finite positive number for $k = 1,2,\ldots$. From (6.47), it follows that

$$\mathcal{E}(r_1^{2k})\, \mathcal{E}((\operatorname{tr}(W_1))^k) = \mathcal{E}((cr_2^2)^k)\, \mathcal{E}((\operatorname{tr}(W_2))^k),\ k = 1,2,\ldots .$$

Hence,

$$\mathcal{E}(r_1^{2k}) = \mathcal{E}((cr_2^2)^k), \ k = 1,2,\dots . \tag{6.48}$$

Since, $Z_1 \approx s_1 L_1$ and $L_1 \approx r_1 U_1$, we can write $Z_1 \approx s_1 r_1 U_1$, where s_1, r_1, and U_1 are independent.

Similarly, $Z_2 \approx s_2 \sqrt{c}\, r_2 U_2$, with s_2, r_2, U_2 independent. Since $s_1 \approx s_2$, we have

$$\mathcal{E}(s_1^{2k}) = \mathcal{E}(s_2^{2k}) = (1 - P_0), \ k = 0,1,2,\dots . \tag{6.49}$$

From (6.48) and (6.49), it follows that

$$\mathcal{E}\left(\left(\frac{1}{c} s_1^2 r_1^2\right)^k\right) = \mathcal{E}((s_2^2 r_2^2)^k).$$

Now,

$$\begin{aligned} x &= \mathrm{tr}(X'\Sigma_1^{-1}X) \\ &= \mathrm{tr}(Z_1'H'^{-1}\Sigma_1^{-1}H^{-1}Z_1) \\ &= \mathrm{tr}(Z_1'Z_1) \\ &= s_1^2 r_1^2\, \mathrm{tr}(U_1'U_1) \\ &= s_1^2 r_1^2 . \end{aligned}$$

Thus, $m_k = \mathcal{E}((s_1^2 r_1^2)^k) = \mathcal{E}((cs_2^2 r_2^2)^k)$. However, if m_k is the k^{th} moment of a random variable and $\sum_{k=1}^{\infty} \left(\frac{1}{m_{2k}}\right)^{\frac{1}{2k}} = \infty$, then the distribution of the random variable is uniquely determined (see Rao, 1973, p. 106). Thus we have $s_1^2 r_1^2 \approx cs_2^2 r_2^2$. Therefore, $s_1 r_1 \approx s_2 \sqrt{c}\, r_2$, and hence $Z_1 \approx Z_2$. ■

CHAPTER 7

ESTIMATION

7.1. MAXIMUM LIKELIHOOD ESTIMATORS OF THE PARAMETERS

Let $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$, with p.d.f.

$$f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(\mathrm{tr}(X - M)' \Sigma^{-1}(X - M)\Phi^{-1}).$$

Let $n \geq p$ and assume h is known. We want to estimate M, Σ, and Φ based on a single observation from X.

First, we show that, without imposing some restrictions on M, Σ, and Φ, indeed the maximum likelihood estimators (MLE's) do not exist.

Let a be a positive number for which $h(a) \neq 0$, and Φ be any $n \times n$ positive definite matrix. Let s be a positive number, and define $k = \sqrt{\frac{sa}{p}}$. Let V be a $p \times n$ matrix such that $V = k(I_p, O)$, $M = X - V\Phi^{\frac{1}{2}}$, and $\Sigma = sI_p$. Then, $f(X)$ can be expressed as

$$f(X) = \frac{1}{(s^p)^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(\mathrm{tr}(\Phi^{\frac{1}{2}} V'\Sigma^{-1} V \Phi^{\frac{1}{2}} \Phi^{-1}))$$

$$= \frac{1}{s^{\frac{pn}{2}} |\Phi|^{\frac{p}{2}}} h(\mathrm{tr}(k^2(sI_p)^{-1} I_p))$$

$$= \frac{1}{s^{\frac{pn}{2}} |\Phi|^{\frac{p}{2}}} h\left(\frac{k^2}{s} p\right)$$

$$= \frac{h(a)}{s^{\frac{pn}{2}} \, |\Phi|^{\frac{p}{2}}}.$$

Therefore, if $s \to 0$, then $f(X) \to \infty$. Hence, the MLE's do not exist. This example shows that even if Φ is known, there are no MLE's for M and Σ. Thus, we have to restrict the parameter space, in order to get the MLE's. The following lemma will be needed to obtain the estimation results.

LEMMA 7.1.1. *Let* X *be a* $p \times n$ *matrix,* μ *a p-dimensional vector,* $\mathbf{v}$ *an n-dimensional vector and* A *an* $n \times n$ *symmetric matrix, such that* $\mathbf{v}'A\mathbf{v} \neq 0$. *Then,*

$$(X - \mu\mathbf{v}')\, A(X - \mu\mathbf{v}')' = X\left(A - \frac{A\mathbf{v}\mathbf{v}'A}{\mathbf{v}'A\mathbf{v}}\right)X'$$

$$+ (\mathbf{v}'A\mathbf{v})\left(X\frac{A\mathbf{v}}{\mathbf{v}'A\mathbf{v}} - \mu\right)\left(X\frac{A\mathbf{v}}{\mathbf{v}'A\mathbf{v}} - \mu\right)'. \tag{7.1}$$

PROOF: The right hand side of (7.1) equals

$$XAX' - X\frac{A\mathbf{v}\mathbf{v}'A}{\mathbf{v}'A\mathbf{v}}X'$$

$$+ (\mathbf{v}'A\mathbf{v})\left(X\frac{A\mathbf{v}\mathbf{v}'A}{(\mathbf{v}'A\mathbf{v})^2}X' - X\frac{A\mathbf{v}}{\mathbf{v}'A\mathbf{v}}\mu' - \mu\,\frac{\mathbf{v}'A}{\mathbf{v}'A\mathbf{v}}X' + \mu\mu'\right)$$

$$= XAX' - XA\mathbf{v}\mu' - \mu\mathbf{v}'AX' + (\mathbf{v}'A\mathbf{v})\mu\mu'$$

$$= (X - \mu\mathbf{v}')A(X - \mu\mathbf{v}')',$$

which is the left hand side of (7.1). ■

Now denote by $\mathbf{x}_i$ the i^{th} column of the matrix X, then

$$\sum_{i=1}^{n} (x_i - \mu)(x_i - \mu)' = (X - \mu e_n')(X - \mu e_n')'$$

and since $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n} X e_n$, we have

$$\sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})' = (X - \bar{x}e_n')(X - \bar{x}e_n')'$$

$$= \left(X - \frac{1}{n} X e_n e_n'\right)\left(X - \frac{1}{n} X e_n e_n'\right)'$$

$$= X\left(I_n - \frac{1}{n} e_n e_n'\right)\left(I_n - \frac{1}{n} e_n e_n'\right)' X'$$

$$= X\left(I_n - \frac{1}{n} e_n e_n'\right) X'.$$

Choosing $v = e_n$, and $A = I_n$, we have $v'Av = n$, $A - \frac{Avv'A}{v'Av} = I_n - \frac{1}{n} e_n e_n'$, and $X \frac{Av}{v'Av} = \frac{1}{n} e_n$. Then, from Lemma 7.1.1, we get

$$\sum_{i=1}^{n} (x_i - \mu)(x_i - \mu)' = \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})' + n(\bar{x} - \mu)(\bar{x} - \mu)',$$

which is a well-known identity.

THEOREM 7.1.1. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$, *with the p.d.f.* $f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(\mathrm{tr}((X - M)' \Sigma^{-1}(X - M)\Phi^{-1}))$, *where* $h(z)$ *is monotone decreasing on* $[0,\infty)$. *Suppose* h, Σ, *and* Φ *are known and we want to find the MLE of* M *(say* $\hat{M}$*), based on a single observation* X. *Then,*

a) $\hat{M} = X$,

b) *if* $M = \mu v'$, *where* μ *is p-dimensional,* $\mathbf{v}$ *is n-dimensional vector and* $\mathbf{v} \neq \mathbf{o}$ *is known, the MLE of* μ *is* $\hat{\mu} = X\frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}}$, *and*

c) *if* M *is of the form* $M = \mu \mathbf{e}_n'$, *the MLE of* μ *is* $\hat{\mu} = X\frac{\Phi^{-1}\mathbf{e}_n}{\mathbf{e}_n'\Phi^{-1}\mathbf{e}_n}$.

PROOF:

a) $$f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(\mathrm{tr}((X - M)' \Sigma^{-1}(X - M)\Phi^{-1}))$$

$$= \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(\mathrm{tr}((\Sigma^{-\frac{1}{2}}(X - M)\Phi^{-\frac{1}{2}})' (\Sigma^{-\frac{1}{2}}(X - M)\Phi^{-\frac{1}{2}})).$$

Let $\mathbf{y} = \mathrm{vec}(\Sigma^{-\frac{1}{2}}(X - M)\Phi^{-\frac{1}{2}})'$. Then, we have $\mathrm{tr}((\Sigma^{-\frac{1}{2}}(X - M)\Phi^{-\frac{1}{2}})' (\Sigma^{-\frac{1}{2}}(X - M)\Phi^{-\frac{1}{2}})) = \mathbf{y}'\mathbf{y}$. Since h is monotone decreasing in $[0,\infty)$ the last expression attains its minimum when $\mathbf{y}'\mathbf{y}$ is minimum. Now $\mathbf{y}'\mathbf{y} \geq 0$ and $\mathbf{y}'\mathbf{y} = 0$ iff $\mathbf{y} = \mathbf{o}$, therefore $f(X)$ is minimized for $\mathbf{y} = \mathbf{o}$. This means that $X = M$. Hence, $\hat{M} = X$.

b) We have

$$f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(\mathrm{tr}(\Sigma^{-1}(X - \mu\mathbf{v}') \Phi^{-1}(X - \mu\mathbf{v}')')).$$

Using Lemma 7.1.1, we can write

$$f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h\left(\mathrm{tr}\left(\Sigma^{-1}\left[X\left(\Phi^{-1} - \frac{\Phi^{-1}\mathbf{v}\mathbf{v}'\Phi^{-1}}{\mathbf{v}'\Phi^{-1}\mathbf{v}}\right)X'\right.\right.\right.$$

$$\left.\left.\left. + (\mathbf{v}'\Phi^{-1}\mathbf{v})\left(X\frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}} - \mu\right)\left(X\frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}} - \mu\right)'\right]\right)\right)$$

$$= \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h\left(\mathrm{tr}\left(\Sigma^{-1} X \left(\Phi^{-1} - \frac{\Phi^{-1} \mathbf{v}\mathbf{v}' \Phi^{-1}}{\mathbf{v}'\Phi^{-1}\mathbf{v}}\right) X'\right)\right.$$

$$\left. + (\mathbf{v}'\Phi^{-1}\mathbf{v})\, \mathrm{tr}\left(\left(X\frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}} - \mu\right)' \Sigma^{-1}\left(X\frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}} - \mu\right)\right)\right).$$

Again, since h is monotone decreasing in $[0,\infty)$ the last expression attains its minimum when $\mathrm{tr}\left(\left(X\frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}} - \mu\right)' \Sigma^{-1}\left(X\frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}} - \mu\right)\right)$ is minimum. Writing $\mathbf{y} = \mathrm{vec}\left(\Sigma^{-\frac{1}{2}}\left(X\frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}} - \mu\right)\right)'$, we have to minimize $\mathbf{y}\mathbf{y}'$. Therefore, minimum is attained at $\mathbf{y} = \mathbf{o}$. So we must have $X\frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}} = \mu$. Hence,

$$\hat{\mu} = X\frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi\mathbf{v}}.$$

c) This is a special case of (b), with $\mathbf{v} = \mathbf{e}_n$. ■

The next result is based on a theorem due to Anderson, Fang and Hsu (1986).

THEOREM 7.1.2. *Assume we have an observation* X *from the distribution* $E_{p,n}(M,\Sigma \otimes \Phi,\psi)$, *where* $(M,\Sigma \otimes \Phi) \in \Omega \subset \mathbb{R}^{p\times n} \times \mathbb{R}^{pn\times pn}$. *Suppose* Ω *has the property that if* $(Q,S) \in \Omega$ $(Q \in \mathbb{R}^{p\times n}$, $S \in \mathbb{R}^{pn\times pn})$, *then* $(Q,cS) \in \Omega$ *for any* $c > 0$ *scalar. Moreover, let* X *have the p.d.f.*

$$f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(\mathrm{tr}((X - M)' \Sigma^{-1}(X - M)\Phi^{-1})),$$

where $\ell(z) = z^{\frac{pn}{2}} h(z)$, $z \geq 0$ *has a finite maximum at* $z = z_h > 0$. *Furthermore, suppose that under the assumption that* X *has the distribution* $N_{p,n}(M,\Sigma \otimes \Phi)$,

$(M, \Sigma \otimes \Phi) \in \Omega$, *the MLE's of* M *and* $\Sigma \otimes \Phi$ *are* M^* *and* $(\Sigma \otimes \Phi)^*$, *which are unique and* $P((\Sigma \otimes \Phi)^* > O) = 1$. *Then, under the condition* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$, $(M, \Sigma \otimes \Phi) \in \Omega$, *the MLE's of* M *and* $\Sigma \otimes \Phi$ *are* $\hat{M} = M^*$, $(\widehat{\Sigma \otimes \Phi}) = \frac{np}{z_h}(\Sigma \otimes \Phi)^*$ *and the maximum of the likelihood is*

$$|(\widehat{\Sigma \otimes \Phi})|^{-\frac{1}{2}} h(z_h).$$

PROOF: Define

$$\Sigma_1 = \frac{\Sigma}{|\Sigma|^{\frac{1}{p}}}, \quad \Phi_1 = \frac{\Phi}{|\Phi|^{\frac{1}{n}}} \tag{7.2}$$

and $z = \operatorname{tr}((X - M)'\Sigma^{-1}(X - M)\Phi^{-1}))$. Then, we have

$$z = \frac{\operatorname{tr}(X - M)' \Sigma_1^{-1}(X - M)\Phi_1^{-1}}{|\Sigma|^{\frac{1}{p}} |\Phi|^{\frac{1}{n}}}. \tag{7.3}$$

Therefore, we can write

$$\begin{aligned} f(X) &= \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(z) \\ &= z^{\frac{np}{2}} h(z)(\operatorname{tr}((X - M)'\Sigma_1^{-1}(X - M)\Phi_1^{-1}))^{-\frac{np}{2}}. \end{aligned} \tag{7.4}$$

Hence, maximizing $f(X)$ is equivalent to maximizing $\ell(z)$ and $(\operatorname{tr}((X - M)' \Sigma_1^{-1}(X - M)\Phi_1^{-1}))^{-\frac{np}{2}}$. If $X \sim N_{p,n}(M, \Sigma \otimes \Phi)$, then $h(z) = (2\pi)^{-\frac{pn}{2}} e^{-\frac{z}{2}}$, and $\ell(z) = (2\pi)^{-\frac{pn}{2}} z^{\frac{pn}{2}} e^{-\frac{z}{2}}$. Therefore,

$$\frac{d\ell(z)}{dz} = (2\pi)^{-\frac{pn}{2}}\left(\frac{pn}{2}z^{\frac{pn}{2}-1}e^{-\frac{z}{2}} + z^{\frac{pn}{2}}\left(-\frac{1}{2}\right)e^{-\frac{z}{2}}\right)$$

$$= \frac{1}{2}(2\pi)^{-\frac{pn}{2}} z^{\frac{pn}{2}-1} e^{-\frac{z}{2}}(pn - z).$$

Consequently, $\ell(z)$ attains its maximum at $z = pn$. From the conditions of the theorem it follows that under normality, $(\mathrm{tr}((X - M)'\Sigma_1^{-1}(X - M)\Phi^{-1}))^{-\frac{np}{2}}$ is maximized for $M = M^*$ and $\Sigma_1 \otimes \Phi_1 = (\Sigma_1 \otimes \Phi_1)^* = \frac{(\Sigma \otimes \Phi)^*}{|(\Sigma \otimes \Phi)^*|^{\frac{1}{np}}}$. Since, $(\mathrm{tr}((X - M)'\Sigma_1^{-1}(X - M)\Phi_1^{-1}))^{-\frac{np}{2}}$ does not depend on h, in the case of $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$, it also attains its maximum for $M = \hat{M} = M^*$ and $\Sigma_1 \otimes \Phi_1 = (\Sigma_1 \hat{\otimes} \Phi_1) = (\Sigma_1 \otimes \Phi_1)^*$. On the other hand, $\ell(z)$ is maximized for $z = z_h$. Then, using (7.2) and (7.3) we get

$$(\Sigma \hat{\otimes} \Phi) = |\Sigma \otimes \Phi|^{\frac{1}{np}}(\Sigma_1 \hat{\otimes} \Phi_1)$$

$$= |\hat{\Sigma}|^{\frac{1}{p}} |\hat{\Phi}|^{\frac{1}{n}}(\Sigma_1 \hat{\otimes} \Phi_1)$$

$$= \frac{\mathrm{tr}((X - M)'\hat{\Sigma}_1^{-1}(X - M)\hat{\Phi}_1^{-1})}{z_h}(\Sigma_1 \hat{\otimes} \Phi_1)$$

$$= \frac{np}{z_h}\,\frac{\mathrm{tr}((X - M^*)'\,\Sigma_1^{*-1}(X - M^*)\Phi_1^{*-1})}{np}(\Sigma_1 \otimes \Phi_1)^*$$

$$= \frac{np}{z_h}(\Sigma \otimes \Phi)^*.$$

The maximum of the likelihood is

$$\frac{1}{|(\Sigma \hat{\otimes} \Phi)|^{\frac{1}{2}}} h(z_h) = |(\Sigma \hat{\otimes} \Phi)|^{-\frac{1}{2}} h(z_h). \blacksquare$$

REMARK 7.1.1. Assume $X \sim N_{p,n}(M, \Sigma \otimes \Phi)$. Then $h(z) = (2\pi)^{-\frac{pn}{2}} e^{-\frac{z}{2}}$, $\ell(z) = \left(\frac{z}{2\pi}\right)^{\frac{pn}{2}} e^{-\frac{z}{2}}$, and $\ell(z)$ attains its maximum at $z_h = pn$. Moreover, $h(z_h) = (2\pi e)^{-\frac{pn}{2}}$.

It is natural to ask whether z_h, as defined in Theorem 7.1.2, exists in a large class of m.e.c. distributions. The following lemma, essentially due to Anderson, Fang, and Hsu (1986), gives a sufficient condition for the existence of z_h.

LEMMA 7.1.2. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ *have the p.d.f.*

$$f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(\mathrm{tr}((X - M)'\Sigma^{-1}(X - M)\Phi^{-1})).$$

Assume $h(z)$, $(z \geq 0)$ *is continuous and monotone decreasing, if* z *is sufficiently large. Then, there exists* $z_h > 0$, *such that* $\ell(z) = z^{\frac{np}{2}} h(z)$ *attains its maximum at* $z = z_h$.

PROOF: From Theorem 2.5.5, $r = (\mathrm{tr}((X - M)'\Sigma^{-1}(X - M)\Phi^{-1}))^{\frac{1}{2}}$ has the p.d.f.

$$p_1(r) = \frac{2\pi^{\frac{np}{2}}}{\Gamma\left(\frac{np}{2}\right)} r^{np-1} h(r^2), \quad r \geq 0.$$

Let $y = r^2$. Then, $J(r \to y) = \frac{1}{2r}$, and hence the p.d.f. of y is

$$p_2(y) = \frac{\pi^{\frac{np}{2}}}{\Gamma\left(\frac{np}{2}\right)} y^{\frac{np}{2}-1} h(y), \ y \geq 0.$$

Consequently,

$$\int_0^\infty \frac{\pi^{\frac{np}{2}}}{\Gamma\left(\frac{np}{2}\right)} y^{\frac{np}{2}-1} h(y)dy = 1.$$

Therefore, $\int_z^\infty y^{\frac{np}{2}-1} h(y)dy \to 0$ as $z \to \infty$.

Now, let $t = \frac{z}{2}$. Then we prove that for sufficiently large z,

$g(z) \leq c \int_t^\infty y^{\frac{np}{2}-1} h(y)dy$, where c is a constant.

If $np > 1$ then

$$\begin{aligned}
\ell(z) &= \ell(2t) \\
&= 2^{\frac{np}{2}} t^{\frac{np}{2}} h(2t) \\
&= 2^{\frac{np}{2}} t^{\frac{np}{2}-1} th(2t) \\
&\leq 2^{\frac{np}{2}} t^{\frac{np}{2}-1} \int_t^{2t} h(y)dy \\
&\leq 2^{\frac{np}{2}} \int_t^{2t} y^{\frac{np}{2}-1} h(y)dy
\end{aligned}$$

$$\leq 2^{\frac{np}{2}} \int_t^{\infty} y^{\frac{np}{2}-1}\, h(y)dy.$$

If $np = 1$, then

$$\begin{aligned}
\ell(z) &= \ell(2t) \\
&= 2^{\frac{1}{2}} t^{\frac{1}{2}} h(2t) \\
&= 2^{\frac{1}{2}} t^{\frac{1}{2}-1}\, t\, h(2t) \\
&\leq 2^{\frac{1}{2}} t^{\frac{1}{2}-1} \int_t^{2t} h(y)dy \\
&= 2(2t)^{\frac{1}{2}-1} \int_t^{2t} h(y)dy \\
&\leq 2 \int_t^{2t} y^{\frac{1}{2}-1}\, h(y)dy \\
&\leq 2 \int_t^{\infty} y^{\frac{1}{2}-1}\, h(y)dy.
\end{aligned}$$

However, $\int_t^{\infty} y^{\frac{np}{2}-1}\, h(y)dy \to 0$ as $t \to \infty$, thus $\ell(z) \to 0$ as $z \to \infty$. Moreover, $\ell(0) = 0h(0) = 0$. Since, $\ell(z)$ is continuous, nonnegative and $\lim_{z\to 0} \ell(z) = \lim_{z\to\infty} \ell(z) = 0$, $\ell(z)$ attains its minimum at a positive number z_h. ■

For the next estimation result we need the following lemma, given in Anderson (1984), p. 62.

LEMMA 7.1.3. *Let* A *be a* $p \times p$ *positive definite matrix and define a function* g *on the set of* $p \times p$ *positive definite matrices as*

$$g(B) = -n \log |B| - \mathrm{tr}(B^{-1}A).$$

Then, $g(B)$ *attains its maximum at* $B = \frac{A}{n}$ *and its maximum value is*

$$g\left(\frac{A}{n}\right) = pn(\log n - 1) - n \log |A|.$$

PROOF: We can write

$$\begin{aligned} g(B) &= -n \log |B| - \mathrm{tr}(B^{-1}A) \\ &= n(\log |A| - \log |B|) - \mathrm{tr}(B^{-1}A) - n \log |A| \\ &= n \log |B^{-1}A| - \mathrm{tr}(B^{-1}A) - n \log |A|. \end{aligned}$$

Now, $g(B)$ is maximized for the same B as $h(B) = n \log |B^{-1}A| - \mathrm{tr}(B^{-1}A)$. We have

$$\begin{aligned} h(B) &= n \log |B^{-1}A^{\frac{1}{2}}A^{\frac{1}{2}}| - \mathrm{tr}(B^{-1}A^{\frac{1}{2}}A^{\frac{1}{2}}) \\ &= n \log |A^{\frac{1}{2}}B^{-1}A^{\frac{1}{2}}| - \mathrm{tr}(A^{\frac{1}{2}}B^{-1}A^{\frac{1}{2}}). \end{aligned}$$

Now, from Theorem 1.3.1, it follows that $A^{\frac{1}{2}}B^{-1}A^{\frac{1}{2}}$ is also positive definite and from Theorem 1.3.7, it can be written as

$$A^{\frac{1}{2}}B^{-1}A^{\frac{1}{2}} = GDG',$$

where G is $p \times p$ orthogonal and D is $p \times p$ diagonal with positive diagonal elements $\lambda_1, \lambda_2, \ldots, \lambda_p$. We obtain

$$h(B) = n \log\left(\prod_{i=1}^{p} \lambda_i\right) - \sum_{i=1}^{p} \lambda_i$$

$$= \sum_{i=1}^{p} (n \log \lambda_i - \lambda_i)\,.$$

Now, $\frac{\partial h(B)}{\partial \lambda_i} = \frac{n}{\lambda_i} - 1$. Thus from $\frac{\partial h(B)}{\partial \lambda_i} = 0$, we get $\lambda_i = n$, $i = 1,2,\ldots,p$. Hence, $h(B)$ attains its maximum when $D = nI_p$ and so

$$A^{\frac{1}{2}}B^{-1}A^{\frac{1}{2}} = GnI_pG' = nI_p.$$

Therefore, $B^{-1} = nA^{-1}$, and then $B = \frac{A}{n}$. Moreover,

$$g\left(\frac{A}{n}\right) = -n \log \left(\frac{A}{n}\right) - \operatorname{tr}(n\, I_p)$$

$$= n(pn - 1) - n \log(A). \blacksquare$$

THEOREM 7.1.3. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$ *have the p.d.f.* $f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(\operatorname{tr}(X - M)'\Sigma^{-1}(X - M)\Phi^{-1}))$, *where* $n \geq p$ *and the function* $\ell(z) = z^{\frac{pn}{2}} h(z)$, $z \in [0,\infty)$, *attains its maximum for a positive* z *(say* z_h*). Suppose* h, M, *and* Φ *are known, and we want to find the MLE of* Σ *(say* $\hat{\Sigma}$*), based on a single observation* X. *Then,* $\hat{\Sigma} = \frac{p}{z_h}(X - M)\Phi^{-1}(X - M)'$.

PROOF: Step 1. First, we prove the result for normal distribution; that is, when $X \sim N_{p,n}(M, \Sigma \otimes \Phi)$. Here, $h(z) = \exp\left(-\frac{z}{2}\right)$, and

$$f(X) = \frac{1}{(2\pi)^{\frac{pn}{2}} |\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} \operatorname{etr}\left\{-\frac{(\Sigma^{-1}(X - M)\Phi^{-1}(X - M)')}{2}\right\}.$$

Taking logarithm of both sides of the last equation and applying Lemma 7.1.3, we obtain that $f(X)$ attains its maximum in Σ, if

$$\hat{\Sigma} = \frac{(X - M)\Phi^{-1}(X - M)'}{n}.$$

Step 2. Let $X \sim E_{p,n}(\mu v', \Sigma \otimes \Phi, \psi)$. Since we proved, in Step 1, that for the normal case, the MLE of Σ is $\frac{(X - M)\Phi^{-1}(X - M)'}{n}$, by using Theorem 7.1.2 we get

$$\hat{\Sigma} = \frac{np}{z_h} \frac{(X - M)\Phi^{-1}(X - M)'}{n} = \frac{p}{z_h}(X - M)\Phi^{-1}(X - M)'.$$

It follows, from Theorem 5.2.2, that $\operatorname{rk}((X - M)\Phi^{-1}(X - M)') = p$ with probability one. Hence, $P(\hat{\Sigma} > O) = 1.$ ■

The next result is an extension of a result of Anderson, Fang, and Hsu (1986).

THEOREM 7.1.4. *Let* $X \sim E_{p,n}(\mu v', \Sigma \otimes \Phi, \psi)$ *have the p.d.f.* $f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(\operatorname{tr}((X - \mu v')'\Sigma^{-1}(X - \mu v')\Phi^{-1}))$, *where* $n > p$, μ *is a p-dimensional vector,* $\mathbf{v}$ *is an n-dimensional nonzero vector, and the function* $\ell(z) = z^{\frac{pn}{2}} h(z)$, $z \in [0,\infty)$, *attains its maximum for a positive* z *(say* z_h*). Suppose* h, $\mathbf{v}$, *and* Φ *are known and we want to find the MLEs of* μ *and* Σ *(say* $\hat{\mu}$ *and* $\hat{\Sigma}$*) based on a single observation* X. *Then*

$$\hat{\mu} = X\frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}} \quad \textit{and} \quad \hat{\Sigma} = \frac{p}{z_h} X\left(\Phi^{-1} - \frac{\Phi^{-1}\mathbf{v}\mathbf{v}'\Phi^{-1}}{\mathbf{v}'\Phi^{-1}\mathbf{v}}\right)X'.$$

In the special case, when $\mathbf{v} = \mathbf{e}_n$, *we have*

$$\hat{\mu} = X\frac{\Phi^{-1}\mathbf{e}_n}{\mathbf{e}_n'\Phi^{-1}\mathbf{e}_n} \text{ and } \hat{\Sigma} = \frac{p}{z_h}X\left(\Phi^{-1} - \frac{\Phi^{-1}\mathbf{e}_n\mathbf{e}_n'\Phi^{-1}}{\mathbf{e}_n'\Phi^{-1}\mathbf{e}_n}\right)X'.$$

PROOF: Step 1. First we prove the result for normal distribution; that is, when $X \sim N_{p,n}(\mu\mathbf{v}', \Sigma \otimes \Phi)$. Here, $h(z) = \exp\left(-\frac{z}{2}\right)$.

Using Lemma 7.1.1, the p.d.f. of X can be written as

$$f(X) = \frac{1}{(2\pi)^{\frac{pn}{2}} |\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} \exp\left\{-\frac{1}{2}\left(\operatorname{tr} \Sigma^{-1}\left[X\left(\Phi^{-1} - \frac{\Phi^{-1}\mathbf{v}\mathbf{v}'\Phi^{-1}}{\mathbf{v}'\Phi^{-1}\mathbf{v}}\right)X'\right.\right.\right.$$

$$\left.\left.\left. + (\mathbf{v}'\Phi^{-1}\mathbf{v})\left(X\frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}} - \mu\right)\left(X\frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}} - \mu\right)'\right]\right)\right\}.$$

Minimizing the expression on the right-hand side, we get

$$\hat{\mu} = X\frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}} \text{ and } \hat{\Sigma} = \frac{X\left(\Phi^{-1} - \frac{\Phi^{-1}\mathbf{v}\mathbf{v}'\Phi^{-1}}{\mathbf{v}'\Phi^{-1}\mathbf{v}}\right)X'}{n}.$$

As noted in Remark 7.1.1, $\ell(z)$ is maximized for $z_h = pn$. Thus, $\hat{\Sigma} = \frac{p}{z_h}X\left(\Phi^{-1} - \frac{\Phi^{-1}\mathbf{v}\mathbf{v}'\Phi^{-1}}{\mathbf{v}'\Phi^{-1}\mathbf{v}}\right)X'$.

Step 2. Let $X \sim E_{p,n}(\mu\mathbf{v}', \Sigma \otimes \Phi, \psi)$. We found the MLE's of μ and Σ for the normal case in Step 1. Now using Theorem 7.1.2, we get

$$\hat{\mu} = X\frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}}$$

and

$$\hat{\Sigma} = \frac{np}{z_h} \frac{X\left(\Phi^{-1} - \dfrac{\Phi^{-1}vv'\Phi^{-1}}{v'\Phi^{-1}v}\right)X'}{n} = \frac{p}{z_h} X\left(\Phi^{-1} - \frac{\Phi^{-1}vv'\Phi^{-1}}{v'\Phi^{-1}v}\right)X'.$$

From Theorem 5.2.2 it follows that with probability one,

$$\text{rk}\left(X\left(\Phi^{-1} - \frac{\Phi^{-1}vv'\Phi^{-1}}{v'\Phi^{-1}v}\right)X'\right) = \min\left(\text{rk}\left(\Phi^{-1} - \frac{\Phi^{-1}vv'\Phi^{-1}}{v'\Phi^{-1}v}\right), p\right).$$

Since, $\Phi^{-1} - \dfrac{\Phi^{-1}vv'\Phi^{-1}}{v'\Phi^{-1}v} = \Phi^{-\frac{1}{2}}\left(I_n - \dfrac{\Phi^{-\frac{1}{2}}vv'\Phi^{-\frac{1}{2}}}{v'\Phi^{-1}v}\right)\Phi^{-\frac{1}{2}}$ and Φ is of full rank, we have

$$\begin{aligned}
\text{rk}\left(\Phi^{-1} - \frac{\Phi^{-1}vv\Phi^{-1}}{v'\Phi^{-1}v}\right) &= \text{rk}\left(I_n - \frac{\Phi^{-\frac{1}{2}}vv'\Phi^{-\frac{1}{2}}}{v'\Phi^{-1}v}\right) \\
&\geq \text{rk}(I_n) - \text{rk}\left(\frac{\Phi^{-\frac{1}{2}}vv'\Phi^{-\frac{1}{2}}}{v'\Phi^{-1}v}\right) \\
&= n - \text{rk}\left(\frac{v'\Phi^{-1}v}{v'\Phi^{-1}v}\right) \\
&= n - 1
\end{aligned}$$

where we used part (ii) of Theorem 1.3.3. Hence, $P(\hat{\Sigma} > O) = 1.$ ■

7.2. PROPERTIES OF THE ESTIMATORS

Now, we derive the distributions of the estimators of μ and Σ. These theorems are based on Anderson and Fang (1982a).

THEOREM 7.2.1. *Let* $X \sim E_{p,n}(\mu v', \Sigma \otimes \Phi, \psi)$ *have the p.d.f.*
$f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(\mathrm{tr}((X - \mu v')' \Sigma^{-1}(X - \mu v')\Phi^{-1}))$, *where* $n > p$, μ *is a p-dimensional and* $\mathbf{v}$ *is an n-dimensional vector,* $\mathbf{v} \neq \mathbf{o}$. *Let* $\hat{\mu} = X \frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}}$ *and*
$A = X\left(\Phi^{-1} - \frac{\Phi^{-1}\mathbf{v}\mathbf{v}'\Phi^{-1}}{\mathbf{v}'\Phi^{-1}\mathbf{v}}\right)X'$.

a) *Then, the joint density of* $\hat{\mu}$ *and* A *is*

$$p(\hat{\mu}, A) = \frac{(\mathbf{v}'\Phi^{-1}\mathbf{v})^{\frac{p}{2}} |A|^{\frac{n-p}{2}-1} \pi^{\frac{p(n-1)}{2}}}{\Gamma_p\left(\frac{n-1}{2}\right) |\Sigma|^{\frac{n}{2}}} h(\mathbf{v}'\Phi^{-1}\mathbf{v}(\hat{\mu} - \mu)'\Sigma^{-1}(\hat{\mu} - \mu) + \mathrm{tr}(\Sigma^{-1}A)). \tag{7.5}$$

b) $\hat{\mu} \sim E_p\left(\mu, \frac{1}{\mathbf{v}'\Phi^{-1}\mathbf{v}}\Sigma, \psi\right)$ *and* $\hat{\mu}$ *has the p.d.f.*

$$p_1(\hat{\mu}) = \frac{2(\mathbf{v}'\Phi^{-1}\mathbf{v})^{\frac{p}{2}} \pi^{\frac{p(n-1)}{2}}}{\Gamma\left(\frac{p(n-1)}{2}\right) |\Sigma|^{\frac{1}{2}}} \int_0^\infty r^{p(n-1)-1} h(r^2 + \mathbf{v}'\Phi^{-1}\mathbf{v}(\hat{\mu} - \mu)'\Sigma^{-1}(\hat{\mu} - \mu))dr. \tag{7.6}$$

c) $A \sim G_{p,1}\left(\Sigma, \frac{n-1}{2}, \psi\right)$ *and its p.d.f. is given by*

$$p_2(A) = \frac{2\pi^{\frac{pn}{2}} |A|^{\frac{n-p}{2}-1}}{\Gamma\left(\frac{p}{2}\right)\Gamma_p\left(\frac{n-1}{2}\right) |\Sigma|^{\frac{n-1}{2}}} \int_0^\infty r^{p-1} h(r^2 + \mathrm{tr}(\Sigma^{-1}A))dr. \tag{7.7}$$

PROOF: a) First, we derive the result for the case $\Phi = I_n$.

Let $G \in O(n)$, whose first column is $\frac{\mathbf{v}}{\sqrt{\mathbf{v}'\mathbf{v}}}$ and let us use the notation $G = \left(\frac{\mathbf{v}}{\sqrt{\mathbf{v}'\mathbf{v}}}, G_1\right)$. Since G is orthogonal, we have $GG' = I_n$; that is,

$\frac{\mathbf{v}}{\sqrt{\mathbf{v}'\mathbf{v}}}\frac{\mathbf{v}'}{\sqrt{\mathbf{v}'\mathbf{v}}} + G_1G_1' = I_n$. Thus,

$$G_1G_1' = I_n - \frac{\mathbf{v}\mathbf{v}'}{\mathbf{v}'\mathbf{v}}. \tag{7.8}$$

Define $Y = XG$. Then,

$$Y \sim E_{p,n}(\mu\mathbf{v}'G, \Sigma \otimes I_n, \psi). \tag{7.9}$$

Partition Y as $Y = (\mathbf{y}_1, Y_2)$, where $\mathbf{y}_1$ is a p-dimensional vector. Since, G is orthogonal, $\mathbf{v}'G_1 = \mathbf{o}$. Therefore,

$$\mathbf{v}'G = \mathbf{v}'\left(\frac{\mathbf{v}}{\sqrt{\mathbf{v}'\mathbf{v}}}, G_1\right) = (\sqrt{\mathbf{v}'\mathbf{v}}, \mathbf{o}).$$

Now, (7.9) can be written as

$$(\mathbf{y}_1, Y_2) \sim E_{p,n}((\sqrt{\mathbf{v}'\mathbf{v}}\,\mu, O), \Sigma \otimes I_n, \psi). \tag{7.10}$$

Moreover, $(\mathbf{y}_1, Y_2) = X\left(\frac{\mathbf{v}}{\sqrt{\mathbf{v}'\mathbf{v}}}, G_1\right) = \left(X\frac{\mathbf{v}}{\sqrt{\mathbf{v}'\mathbf{v}}}, XG_1\right)$, hence $\hat{\mu} = \frac{\mathbf{y}_1}{\sqrt{\mathbf{v}'\mathbf{v}}}$ and using (7.8) we get $A = Y_2Y_2'$.

Now, the density of Y; that is, the joint density of $\mathbf{y}_1$ and Y_2 is

$$p_3(\mathbf{y}_1, Y_2) = \frac{1}{|\Sigma|^{\frac{n}{2}}} h((\mathbf{y}_1 - \sqrt{\mathbf{v}'\mathbf{v}}\,\mu)'\Sigma^{-1}(\mathbf{y} - \sqrt{\mathbf{v}'\mathbf{v}}\,\mu) + \operatorname{tr}(Y_2'\Sigma^{-1}Y_2))$$

$$= \frac{1}{|\Sigma|^{\frac{n}{2}}} h((\mathbf{y}_1 - \sqrt{\mathbf{v}'\mathbf{v}}\,\mu)'\Sigma^{-1}(\mathbf{y}_1 - \sqrt{\mathbf{v}'\mathbf{v}}\,\mu) + \operatorname{tr}(\Sigma^{-1}Y_2Y_2')). \tag{7.11}$$

Using Lemma 5.1.2, we can write the joint density of $\mathbf{y}_1$ and A as

$$p_4(y_1,A) = \frac{\pi^{\frac{p(n-1)}{2}}}{\Gamma_p\left(\frac{n-1}{2}\right)} |A|^{\frac{n-1-p-1}{2}} \frac{1}{|\Sigma|^{\frac{n}{2}}} h((y_1 - \sqrt{v'v}\ \mu)'$$

$$\cdot\ \Sigma^{-1}(y_1 - \sqrt{v'v}\ \mu) + \mathrm{tr}(\Sigma^{-1}A)). \tag{7.12}$$

We have $y_1 = \sqrt{v'v}\ \hat{\mu}$. Hence, $J(y_1 \to \hat{\mu}) = (v'v)^{\frac{p}{2}}$. Therefore, the joint p.d.f. of $\hat{\mu}$ and A is

$$p(\hat{\mu},A) = \frac{(v'v)^{\frac{p}{2}} \pi^{\frac{p(n-1)}{2}}}{\Gamma_p\left(\frac{n-1}{2}\right) |\Sigma|^{\frac{n}{2}}} |A|^{\frac{n-p}{2}-1} h(v'v(\hat{\mu} - \mu)'\Sigma^{-1}(\hat{\mu} - \mu) + \mathrm{tr}(\Sigma^{-1}A)). \tag{7.13}$$

Now, for $\Phi \neq I_n$, define $X^* = X\Phi^{-\frac{1}{2}}$. Then, $X^* \sim E_{p,n}(\mu v^{*\prime}, \Sigma \otimes I_n, \psi)$, with $v^* = \Phi^{-\frac{1}{2}} v$. Thus, we get $v^{*\prime}v^* = v'\Phi^{-1}v$,

$$\hat{\mu} = X^* \frac{v^*}{v^{*\prime}v^*} \text{ and } A = X^*\left(I_n - \frac{v^*v^{*\prime}}{v^{*\prime}v^*}\right) X^{*\prime}.$$

So, using the first part of the proof, from (7.13), we obtain (7.5).

b) Since, $\hat{\mu} = X\dfrac{\Phi^{-1}v}{v'\Phi^{-1}v}$, we get

$$\hat{\mu} \sim E_p\left(\mu \frac{v'\Phi^{-1}v}{v'\Phi^{-1}v}, \Sigma \otimes \frac{v'\Phi^{-1}\Phi\Phi^{-1}v}{(v'\Phi^{-1}v)^2}, \psi\right) = E_p\left(\mu, \frac{1}{v'\Phi^{-1}v}\Sigma, \psi\right).$$

Now, assume $\Phi = I_n$. Then, from (7.11), we derive the p.d.f. of $y_1 = \sqrt{v'v}\ \hat{\mu}$, as

$$p_5(y_1) = \int_{\mathbb{R}^{p\times(n-1)}} p_3(y_1, Y_2)dY_2. \tag{7.14}$$

Let $W = \Sigma^{-\frac{1}{2}} Y_2$. Then, $J(Y_2 \rightarrow W) = |\Sigma|^{\frac{n-1}{2}}$, and from (7.11) and (7.14) we get

$$p_5(y_1) = \frac{1}{|\Sigma|^{\frac{n}{2}}} |\Sigma|^{\frac{n-1}{2}} \int_{\mathbb{R}^{p\times(n-1)}} h((y_1 - \sqrt{v'v}\,\mu)'\Sigma^{-1}(y_1 - \sqrt{v'v}\,\mu)$$

$$+ \operatorname{tr}(WW'))dW. \tag{7.15}$$

Writing $w = \operatorname{vec}(W')$, (7.15) becomes

$$p_5(y_1) = \frac{1}{|\Sigma|^{\frac{1}{2}}} \int_{\mathbb{R}^{p(n-1)}} h((y_1 - \sqrt{v'v}\,\mu)'\Sigma^{-1}(y_1 - \sqrt{v'v}\,\mu) + w'w)dw.$$

Using Lemma 2.2.1, we get

$$p_5(y_1) = \frac{1}{|\Sigma|^{\frac{1}{2}}} \frac{2\pi^{\frac{p(n-1)}{2}}}{\Gamma\left(\frac{p(n-1)}{2}\right)} \int_0^\infty r^{p(n-1)-1} h(r^2 + (y_1 - \sqrt{v'v}\,\mu)'$$

$$\cdot \Sigma^{-1}(y_1 - \sqrt{v'v}\,\mu))dr.$$

Since $\hat{\mu} = (v'v)^{-\frac{1}{2}} y_1$ and $J(y_1 \rightarrow \hat{\mu}) = (v'v)^{\frac{p}{2}}$, the p.d.f. of $\hat{\mu}$ is given by

$$p_1(\hat{\mu}) = \frac{2(v'v)^{\frac{p}{2}} \pi^{\frac{p(n-1)}{2}}}{\Gamma\left(\frac{p(n-1)}{2}\right) |\Sigma|^{\frac{1}{2}}} \int_0^\infty r^{p(n-1)-1} h(r^2 + v'v(\hat{\mu} - \mu)'\Sigma^{-1}(\hat{\mu} - \mu))dr. \tag{7.16}$$

For $\Phi \neq I_n$, define $X^* = X\Phi^{-\frac{1}{2}}$. Then, $X^* \sim E_{p,n}(\mu v^{*\prime}, \Sigma \otimes I_n, \psi)$ with

$v^* = \Phi^{-\frac{1}{2}} v$. Thus, we get $v^{*\prime}v^* = v'\Phi^{-1}v$ and $\hat{\mu} = X^* \dfrac{v^*}{v^{*\prime}v^*}$ and from (7.16) we obtain (7.6).

c) First, assume $\Phi = I_n$. Then from (7.10) we get

$$Y_2 \sim E_{p,n-1}(O, \Sigma \otimes I_{n-1}, \psi).$$

Hence, by Definition 5.1.1, $A = Y_2 Y_2' \sim G_p\left(\Sigma, \frac{n-1}{2}, \Psi\right)$, and its p.d.f., using (7.12), is given by

$$p_2(A) = \int_{\mathbb{R}^p} p_4(y_1, A) dy_1. \tag{7.17}$$

Let $w = \Sigma^{-\frac{1}{2}}(y_1 - \sqrt{v'v}\,\mu)$. Then, $J(y_1 \to w) = |\Sigma|^{\frac{1}{2}}$, and using (7.12) and (7.17) we can write

$$p_2(A) = \frac{\pi^{\frac{p(n-1)}{2}}}{\Gamma_p\left(\frac{n-1}{2}\right) |\Sigma|^{\frac{n}{2}}} |\Sigma|^{\frac{1}{2}} |A|^{\frac{n-p}{2}-1} \int_{\mathbb{R}^p} h(w'w + \mathrm{tr}(\Sigma^{-1}A)) dw.$$

Using polar coordinates, we get

$$p_2(A) = \frac{\pi^{\frac{p(n-1)}{2}}}{\Gamma_p\left(\frac{n-1}{2}\right) |\Sigma|^{\frac{n-1}{2}}} |A|^{\frac{n-p}{2}-1} \frac{2\pi^{\frac{p}{2}}}{\Gamma\left(\frac{p}{2}\right)} \int_0^\infty r^{p-1} h(r^2 + \mathrm{tr}(\Sigma^{-1}A)) dr$$

$$= \frac{2\pi^{\frac{pn}{2}} |A|^{\frac{n-p}{2}-1}}{\Gamma\left(\frac{p}{2}\right)\Gamma_p\left(\frac{n-1}{2}\right) |\Sigma|^{\frac{n-1}{2}}} \int_0^\infty r^{p-1} h(r^2 + \mathrm{tr}(\Sigma^{-1}A)) dr.$$

For $\Phi \neq I_n$, define $X^* = X\Phi^{-\frac{1}{2}}$. Then, $X^* \sim E_{p,n}(\mu v^{*\prime}, \Sigma \otimes I_n, \psi)$ with $v^* = \Phi^{-\frac{1}{2}} v$. Since, $A = X^*\left(I_n - \frac{v^* v^{*\prime}}{v^{*\prime} v^*}\right) X^{*\prime}$ and the distribution of A does not depend on v^*, it has the same distribution under the m.e.c. distribution with $\Phi = I_n$ as under the m.e.c. distribution with $\Phi \neq I_n$. ■

THEOREM 7.2.2. *Let* $X \sim E_{p,n}(O, \Sigma \otimes \Phi, \psi)$ *have the p.d.f.* $f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(\mathrm{tr}(X'\Sigma^{-1}X\Phi^{-1}))$, *where* $n \geq p$. *Let* $B = X\Phi^{-1}X'$, *then* $B \sim G_{p,1}\left(\Sigma, \frac{n}{2}, \psi\right)$, *and the p.d.f. of* B *is* $p(B) = \frac{\pi^{\frac{pn}{2}} |B|^{\frac{n-p-1}{2}}}{\Gamma_p\left(\frac{n}{2}\right) |\Sigma|^{\frac{n}{2}}} h(\mathrm{tr}(\Sigma^{-1}W))$.

PROOF: First, assume $\Phi = I_n$. Then, by Definition 5.1.1, we have $B = XX' \sim G_{p,1}\left(\Sigma, \frac{n}{2}, \psi\right)$. Moreover, we have $f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}}} h(\mathrm{tr}(\Sigma^{-1}XX'))$.

Using Lemma 5.1.1, we obtain the p.d.f. of B as

$$p(B) = \frac{\pi^{\frac{pn}{2}}}{\Gamma_p\left(\frac{n}{2}\right) |\Sigma|^{\frac{n}{2}}} |B|^{\frac{n-p-1}{2}} h(\mathrm{tr}\, \Sigma^{-1}B). \tag{7.18}$$

For $\Phi \neq I_n$, define $X^* = X\Phi^{-\frac{1}{2}}$. Then, $X^* \sim E_{p,n}(O, \Sigma \otimes I_n, \psi)$. Since, $B = X^*X^{*\prime}$, the distribution of B is the same under the m.e.c. distribution with $\Phi = I_n$ as under the m.e.c. distribution with $\Phi \neq I_n$. ■

Before we derive the joint density of the characteristic roots of the estimators of Σ, we give a lemma taken from Anderson (1984).

LEMMA 7.2.1. *Let* A *be a symmetric random matrix. Let*

$\lambda_1 > \lambda_2 > \ldots > \lambda_p$ *be the characteristic roots of* A, *and assume that the density of* A *is a function of* $(\lambda_1,\lambda_2,\ldots,\lambda_p)$, i.e. $f(A) = g(\lambda_1,\lambda_2,\ldots,\lambda_p)$. *Then, the p.d.f. of* $(\lambda_1,\lambda_2,\ldots,\lambda_p)$ *is*

$$p(\lambda_1,\lambda_2,\ldots,\lambda_p) = \frac{\pi^{\frac{p^2}{2}} g(\lambda_1,\lambda_2,\ldots,\lambda_p) \prod_{i<j} (\lambda_i - \lambda_j)}{\Gamma_p\left(\frac{p}{2}\right)}.$$

PROOF: See Anderson (1984), p. 532. ■

THEOREM 7.2.3. *Let* $X \sim E_{p,n}(\mu v', I_p \otimes \Phi, \psi)$ *have the p.d.f.* $f(X) = \frac{1}{|\Phi|^{\frac{n}{2}}} h(\mathrm{tr}((X - \mu v')' (X - \mu v')\Phi^{-1}))$, *where* $n > p$, μ *is a p-dimensional vector and* $\mathbf{v}$ *is an n-dimensional vector,* $\mathbf{v} \neq \mathbf{o}$. *Let* $A = X\left(\Phi^{-1} - \frac{\Phi^{-1}\mathbf{v}\mathbf{v}\Phi^{-1}}{\mathbf{v}'\Phi^{-1}\mathbf{v}}\right)X'$.

Further let $\lambda_1 > \lambda_2 > \ldots > \lambda_p$ *be the characteristic roots of* A. *Then, the p.d.f. of* $(\lambda_1,\lambda_2,\ldots,\lambda_p)$ *is*

$$p(\lambda_1,\lambda_2,\ldots,\lambda_p) = \frac{2\pi^{\frac{p(p+n)}{2}}}{\Gamma\left(\frac{p}{2}\right)\Gamma_p\left(\frac{p}{2}\right)\Gamma_p\left(\frac{n-1}{2}\right)} \left(\prod_{i=1}^{p} \lambda_i\right)^{\frac{n-p}{2}-1} \prod_{i<j} (\lambda_i - \lambda_j)$$

$$\int_0^\infty r^{p-1} h\left(r^2 + \sum_{i=1}^{p} \lambda_i\right) dr. \tag{7.19}$$

PROOF: From (7.7), the p.d.f. of A is

$$p_2(A) = \frac{2\pi^{\frac{pn}{2}}}{\Gamma\left(\frac{p}{2}\right)\Gamma_p\left(\frac{n-1}{2}\right)} \left(\prod_{i=1}^{p} \lambda_i\right)^{\frac{n-p}{2}-1} \int_0^\infty r^{p-1} h\left(r^2 + \sum_{i=1}^{p} \lambda_i\right) dr$$

and then, using Lemma 7.2.1, we obtain (7.19). ■

THEOREM 7.2.4. *Let* $X \sim E_{p,n}(O, I_p \otimes \Phi, \psi)$ *have the p.d.f.* $f(X) = \frac{1}{|\Phi|^{\frac{p}{2}}} h(\mathrm{tr}(X'X\Phi^{-1}))$, *where* $n \geq p$. *Let* $B = X\Phi^{-1}X'$. *Further let* $\lambda_1 > \lambda_2 > \ldots > \lambda_p$ *be the characteristic roots of* B. *Then, the p.d.f. of* $(\lambda_1, \lambda_2, \ldots, \lambda_p)$ *is*

$$p(\lambda_1, \lambda_2, \ldots, \lambda_p) = \frac{\pi^{\frac{p(p+n)}{2}}}{\Gamma_p\left(\frac{p}{2}\right)\Gamma_p\left(\frac{n}{2}\right)} \left(\prod_{i=1}^{p} \lambda_i\right)^{\frac{n-p-1}{2}} \prod_{i<j} (\lambda_i - \lambda_j)\, h\left(\sum_{i=1}^{p} \lambda_i\right). \tag{7.20}$$

PROOF: From (7.18), the p.d.f. of B is

$$p(B) = \frac{\pi^{\frac{pn}{2}} \left(\prod_{i=1}^{p} \lambda_i\right)^{\frac{n-p-1}{2}}}{\Gamma_p\left(\frac{n}{2}\right)} h\left(\sum_{i=1}^{p} \lambda_i\right)$$

and using Lemma 7.2.1, we obtain (7.20). ■

Next, we want to give a representation of the generalized variance of the estimator of Σ, for which we need the following result.

LEMMA 7.2.2. *Let* $X \sim N_{p,n}(O, I_p \otimes I_n)$, *with* $n \geq p$. *Then, there exists a lower triangular random matrix* T *(that is,* $t_{ij} = 0$ *if* $i < j$*), such that*

$$t_{ii}^2 \sim \chi_{n-i+1}^2, \quad i = 1, \ldots, p, \tag{7.21}$$

$$t_{ij} \sim N(0,1), \quad i > j, \tag{7.22}$$

$$t_{ij},\ i \geq j \text{ are independent}, \tag{7.23}$$

and $XX' \approx TT'$.

PROOF: See Anderson (1984), p. 247. ■

THEOREM 7.2.5. *Let* $X \sim E_{p,n}(O,\Sigma \otimes \Phi,\psi)$ *have the p.d.f.*
$f(X) = \dfrac{1}{|\Phi|^{\frac{p}{2}} |\Sigma|^{\frac{n}{2}}} h(tr(X'\Sigma^{-1}X\Phi^{-1}))$, *where* $n \geq p$. *Let* $X \approx r\Sigma^{\frac{1}{2}} U\Phi^{\frac{1}{2}}$ *be the stochastic representation of* X. *Let* $B = X\Phi^{-1}X'$, *then*

$$|B| \approx r^{2p} |\Sigma| \frac{\prod_{i=1}^{p} y_i}{\left(\sum_{i=1}^{p+1} y_i\right)^p}, \tag{7.24}$$

where

$$y_i \sim \chi^2_{n-i+1}, \ i = 1,\dots,p, \tag{7.25}$$

$$y_{p+1} \sim \chi^2_{\frac{p(p-1)}{2}}, \ \textit{and} \tag{7.26}$$

$$y_i, \ i = 1,\dots,p+1 \ \textit{are independent.} \tag{7.27}$$

PROOF: It follows, from Theorem 7.2.2, that the distribution of B does not depend on Φ. So, we can assume $\Phi = I_n$.

Let $V \sim N_{p,n}(O,I_p \otimes I_n)$. Then, from Theorem 2.5.4, it follows that $U \approx \dfrac{V}{\sqrt{tr(VV')}}$. Assume X and V are independent. Then, we get

$$X \approx r\Sigma^{\frac{1}{2}} \frac{V}{\sqrt{tr(VV')}},$$

and hence,

$$|B| = |XX'| \approx \left| r^2\Sigma^{\frac{1}{2}} \frac{VV'}{tr(VV')} \Sigma^{\frac{1}{2}} \right| = r^{2p} |\Sigma| \frac{|VV'|}{(tr(VV'))^p}.$$

From Lemma 7.2.2, we can find a lower triangular matrix T, such that $VV' \approx TT'$, and T satisfies (7.21), (7.22) and (7.23). Then,

$$|VV'| \approx \sum_{i=1}^{p} t_{ii}^2$$

and

$$\mathrm{tr}(VV') \approx \sum_{i=1}^{p} t_{ii}^2 + \sum_{i>j} t_{ij}.$$

Define $y_i = t_{ii}^2$, $i = 1,\dots,p$ and $y_{p+1} = \sum_{i>j} t_{ij}$. Then, (7.24)-(7.28) are satisfied. ■

Theorems 7.2.4 and 7.2.5 are adapted from Anderson and Fang (1982b). Now, we study the question of unbiasedness of the estimators of μ and Σ.

THEOREM 7.2.6. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$.

a) *Then,* $\hat{M} = X$ *is an unbiased estimator of* X.

b) *If* M, Φ, *and* ψ *are known then* $\hat{\Sigma}_1 = \frac{-1}{2n\psi'(0)}(X - M)\Phi^{-1}(X - M)'$ *is an unbiased estimator of* Σ.

c) *If* $M = \mu\mathbf{v}'$, *where* μ *is a p-dimensional vector,* $\mathbf{v}$ *is an n-dimensional vector,* $\mathbf{v} \neq \mathbf{o}$, $\mathbf{v}$ *and* Φ *are known, then* $\hat{\mu} = X\frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}}$ *is an unbiased estimator of* μ. *Moreover, if* ψ *is also known, then*

$$\hat{\Sigma}_2 = \frac{-1}{2(n-1)\psi'(0)} X \left(\Phi^{-1} - \frac{\Phi^{-1}\mathbf{v}\mathbf{v}'\Phi^{-1}}{\mathbf{v}'\Phi^{-1}\mathbf{v}} \right) X'$$

is an unbiased estimator of Σ.

We assume that the first-order moment of X *exists when we state the unbiasedness of the estimators of* μ *and that the second-order moment exists when we consider the unbiasedness of the estimators of* Σ.

PROOF: a) $\mathcal{E}(\hat{M}) = \mathcal{E}(X) = M$.

b) Let $Y = (X - M)\Phi^{-\frac{1}{2}}$. Then, $Y \sim E_{p,n}(O, \Sigma \otimes I_n, \psi)$. So $\hat{\Sigma}_1 = \frac{-1}{2n\psi'(0)}$ YY' and it follows, from Theorem 3.2.8 that

$$\mathcal{E}(\hat{\Sigma}_1) = -\frac{1}{2n\psi'(0)}(-2\psi'(0))\,\Sigma\,\mathrm{tr}(I_n) = \Sigma.$$

c) Now we have

$$\mathcal{E}(\hat{\mu}) = E\left(X\frac{\Phi^{-1}v}{v'\Phi^{-1}v}\right) = \frac{\mu v'\Phi^{-1}v}{v'\Phi^{-1}v} = \mu.$$

From Theorem 3.2.8, we obtain

$$\mathcal{E}(\hat{\Sigma}_2) = -\frac{1}{2(n-1)\psi'(0)}(-2\psi'(0))\,\Sigma\,\mathrm{tr}\left(\left(\Phi^{-1} - \frac{\Phi^{-1}vv'\Phi^{-1}}{v'\Phi^{-1}v}\right)\Phi\right)$$

$$+ \mu v'\left(\Phi^{-1} - \frac{\Phi^{-1}vv'\Phi^{-1}}{v'\Phi^{-1}v}\right)v\mu'$$

$$= \frac{1}{n-1}\,\Sigma\left(\mathrm{tr}(I_n) - \mathrm{tr}\left(\frac{\Phi^{-1}vv'}{v'\Phi^{-1}v}\right)\right) + \mu\left(v'\Phi^{-1}v - \frac{v'\Phi^{-1}vv'\Phi^{-1}v}{v'\Phi^{-1}v}\right)\mu'$$

and since $\mathrm{tr}\left(\frac{\Phi^{-1}vv'}{v'\Phi^{-1}v}\right) = \mathrm{tr}\left(\frac{v'\Phi^{-1}v}{v'\Phi^{-1}v}\right) = \frac{v'\Phi^{-1}v}{v'\Phi^{-1}v} = 1$, we get

$$\mathcal{E}(\hat{\Sigma}_2) = \frac{1}{n-1}\,\Sigma(n-1) + O = \Sigma. \blacksquare$$

THEOREM 7.2.7. *Let* $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$, *with the p.d.f.*
$f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(\mathrm{tr}((X - M)' \Sigma^{-1}(X - M)\Phi^{-1}))$ *and assume* Φ *is known.*

a) *If* Σ *is known, then* $\hat{M} = X$ *is sufficient for* M.

b) *If* M *is known, then* $A = (X - M)\Phi^{-1}(X - M)'$ *is sufficient for* Σ.

c) *If* $M = \mu\mathbf{v}'$, *where* μ *is a p-dimensional vector, and* $\mathbf{v} \neq \mathbf{o}$ *is an n-dimensional known vector, then* $(\hat{\mu}, B)$ *is sufficient for* (μ, Σ), *where*

$$\hat{\mu} = X \frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}} \quad \textit{and} \quad B = X\left(\Phi^{-1} - \frac{\Phi^{-1}\mathbf{v}\mathbf{v}\Phi^{-1}}{\mathbf{v}'\Phi^{-1}\mathbf{v}}\right)X'.$$

PROOF: a) Trivial.

b) This statement follows, if we write

$$f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(\mathrm{tr}(\Sigma^{-1}[(X - M)\Phi^{-1}(X - M)']))$$

$$= \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(\mathrm{tr}(\Sigma^{-1}A)).$$

c) Using Lemma 7.1.1, we can write

$$f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h\left(\mathrm{tr}\left(\Sigma^{-1}\left[X\left(\Phi^{-1} - \frac{\Phi^{-1}\mathbf{v}\mathbf{v}'\Phi^{-1}}{\mathbf{v}'\Phi^{-1}\mathbf{v}}\right)X'\right.\right.\right.$$

$$\left.\left.\left. + (\mathbf{v}'\Phi^{-1}\mathbf{v})\left(X\frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}} - \mu\right)\left(X\frac{\Phi^{-1}\mathbf{v}}{\mathbf{v}'\Phi^{-1}\mathbf{v}} - \mu\right)'\right]\right)\right)$$

$$= \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(\mathrm{tr}(\Sigma^{-1}[B + (\mathbf{v}'\Phi^{-1}\mathbf{v})(\hat{\mu} - \mu)(\hat{\mu} - \mu)'])).$$

This proves that $(\hat{\mu}, B)$ is sufficient for (μ, Σ). ■

REMARK 7.2.1. Let $X \sim E_{p,n}(\mu e_n', \Sigma \otimes I_n, \psi)$ have the p.d.f. $f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}}} h(\mathrm{tr}((X - \mu e_n')'\Sigma^{-1}(X - \mu e_n')))$, where μ is a p-dimensional vector and $h: \mathbb{R}_0^+ \to \mathbb{R}_0^+$ a decreasing function. Let $g: \mathbb{R}_0^+ \to \mathbb{R}_0^+$ be an increasing function. Assume we have an observation from X and want to estimate μ. Then, $\hat{\mu} = X\frac{e_n}{n} = \bar{x}$ is a minimax estimator of μ under the loss function $\ell(z) = g((z - \mu)'\Sigma^{-1}(z - \mu))$. This result has been proved by Fan and Fang (1985a). In another paper, Fan and Fang (1985b) showed that if in addition to the above mentioned conditions, we assume that $n > p \geq 4$, and g is a convex function whose second derivative exists almost everywhere with respect to the Lebesgue measure and $P(X = \mu e_n') = 0$, then $\bar{x}$ is an inadmissible estimator of μ. More precisely, they showed that the estimator $\hat{\mu}_c = \left(1 - \frac{c}{\bar{x}'\hat{\Sigma}^{-1}\bar{x}}\right)\bar{x}$, where $\hat{\Sigma} = X\left(I_n - \frac{e_n e_n'}{n}\right)X'$ and

$$0 < c \leq \frac{2(pn - p + 2)(n - 1)(p - 2)(p - 3)}{n(n - p + 2)(p - 1)(n^2 - 2pn + 2n + p - 2)}$$

dominates $\bar{x}$. As a consequence, $\hat{\mu}_c$ is also a minimax estimator of μ.

CHAPTER 8

HYPOTHESIS TESTING

8.1. GENERAL RESULTS

Before studying concrete hypotheses, we derive some general theorems. These results are based on Anderson, Fang, and Hsu (1986) and Hsu (1985b).

THEOREM 8.1.1. *Assume we have an observation* X *from the distribution* $E_{p,n}(M, \Sigma \otimes \Phi, \psi)$, *where* $(M, \Sigma \otimes \Phi) \in \Omega \subset \mathbb{R}^{p \times n} \times \mathbb{R}^{pn \times pn}$, *and we want to test*

$$H_0: (M, \Sigma \otimes \Phi) \in \omega \text{ against } H_1: (M, \Sigma \otimes \Phi) \in \Omega - \omega, \tag{8.1}$$

where $\omega \subset \Omega$. *Suppose* Ω *and* ω *have the properties that if* $Q \in \mathbb{R}^{p \times n}$, $S \in \mathbb{R}^{pn \times pn}$, *then* $(Q,S) \in \Omega$ *implies* $(Q,cS) \in \Omega$ *and* $(Q,S) \in \omega$ *implies* $(Q,cS) \in \omega$ *for any positive scalar* c. *Moreover, let* X *have the p.d.f.*

$$f(X) = \frac{1}{|\Sigma|^{\frac{n}{2}} |\Phi|^{\frac{p}{2}}} h(\mathrm{tr}((X - M)' \Sigma^{-1}(X - M)\Phi^{-1})),$$

where $\ell(z) = z^{\frac{pn}{2}} h(z)$ $(z \geq 0)$ *has a finite maximum at* $z = z_h > 0$. *Furthermore, suppose that under the assumption that* $X \sim N_{p,n}(M, \Sigma \otimes \Phi)$, $(M, \Sigma \otimes \Phi) \in \Omega$, *the MLE's of* M *and* $\Sigma \otimes \Phi$ *are* M^* *and* $(\Sigma \otimes \Phi)^*$, *which are unique and* $P((\Sigma \otimes \Phi)^* > O) = 1$.

Assume also that under the assumption that $X \sim N_{p,n}(M, \Sigma \otimes \Phi)$, $(M, \Sigma \otimes \Phi) \in \omega$, *the MLE's of* M *and* $\Sigma \otimes \Phi$ *are* M_0^* *and* $(\Sigma \otimes \Phi)_0^*$, *which are unique and* $P((\Sigma \otimes \Phi)_0^* > O) = 1$.

Then, the likelihood ratio test (LRT) statistic for testing (8.1) under the assumption that $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$, *is the same as under the assumption that* $X \sim N_{p,n}(M, \Sigma \otimes \Phi)$, *namely* $\dfrac{|(\Sigma \otimes \Phi)^*|}{|(\Sigma \otimes \Phi)_0^*|}$.

PROOF: From Theorem 7.1.2 it follows that under the condition $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$, $(M, \Sigma \otimes \Phi) \in \Omega$, the MLE's of M and $\Sigma \otimes \Phi$ are $\hat{M} = M^*$, $(\Sigma \hat{\otimes} \Phi) = \dfrac{np}{z_h}(\Sigma \otimes \Phi)^*$ and the maximum of the likelihood is

$$|(\Sigma \hat{\otimes} \Phi)|^{-\frac{1}{2}} h(z_h).$$

Similarly, under the condition $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$, $(M, \Sigma \otimes \Phi) \in \Omega$, the MLE's of M and $\Sigma \otimes \Phi$ are $\hat{M}_0 = M^*$, $(\Sigma \hat{\otimes} \Phi)_0 = \dfrac{np}{z_h}(\Sigma \otimes \Phi)_0^*$ and the maximum of the likelihood is

$$|(\Sigma \hat{\otimes} \Phi)_0|^{-\frac{1}{2}} h(z_h).$$

Hence, the LRT statistic is

$$\frac{|(\Sigma \hat{\otimes} \Phi)_0|^{-\frac{1}{2}} h(z_h)}{|(\Sigma \hat{\otimes} \Phi)|^{-\frac{1}{2}} h(z_h)} = \frac{\left|\frac{np}{z_h}(\Sigma \otimes \Phi)_0^*\right|^{-\frac{1}{2}}}{\left|\frac{np}{z_h}(\Sigma \otimes \Phi)^*\right|^{-\frac{1}{2}}} = \left(\frac{|(\Sigma \otimes \Phi)^*|}{|(\Sigma \otimes \Phi)_0^*|}\right)^{\frac{1}{2}}$$

which is equivalent to the test statistic $\dfrac{|(\Sigma \otimes \Phi)^*|}{|(\Sigma \otimes \Phi)_0^*|}$. ■

THEOREM 8.1.2. *Assume we have an observation* X *from the absolutely continuous distribution* $E_{p,n}(M, \Sigma \otimes \Phi, \psi)$, *where* $(M, \Sigma \otimes \Phi) \in \Omega = \Omega_1 \times \Omega_2$ *with* $\Omega_1 \subset \mathbb{R}^{p \times n}$, $\Omega_2 \subset \mathbb{R}^{pn \times pn}$. *We want to test*

$$H_0\colon (M,\Sigma \otimes \Phi) \in \omega \; \textit{against} \; H_1\colon (M,\Sigma \otimes \Phi) \in \Omega - \omega,$$

where $\omega = \omega_1 \times \omega_2$, $\omega_1 \subset \Omega_1$, $\omega_2 \subset \Omega_2$. *Assume that* $O \in \omega_1$. *Let* $f(Z)$ *be a test statistic, such that* $f(cZ) = f(Z)$ *for any scalar* $c > 0$. *Then, we have the following*

a) *If*

$$f(Z) = f(Z - M) \tag{8.2}$$

for every $M \in \omega_1$, *then the null distribution of* $f(X)$ *is the same under* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$, *as under* $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$.

b) *If* $f(Z) = f(Z - M)$ *for every* $M \in \Omega_1$, *then the distribution of* $f(X)$, *null as well as the nonnull, is the same under* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$, *as under* $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$.

PROOF: a) Let $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$, $M \in \omega_1$. Define $Y = X - M$. Then $Y \sim E_{p,n}(O,\Sigma \otimes \Phi,\psi)$ and $f(X) = f(X - M) = f(Y)$. Thus, the distribution of $f(X)$ is the same as the distribution of $f(Y)$.

From Theorem 5.3.1, it follows that the distribution of $f(Y)$ is the same under $Y \sim E_{p,n}(O,\Sigma \otimes \Phi,\psi)$, as under $Y \sim N_{p,n}(O,\Sigma \otimes \Phi)$. However, $f(Y) = f(Y + M)$, therefore, the distribution of $f(Y)$ is the same under $Y \sim N_{p,n}(O,\Sigma \otimes \Phi)$, as under $Y \sim N_{p,n}(M,\Sigma \otimes \Phi)$.

b) Let $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$, $M \in \Omega_1$. Define $Y = X - M$. Then, $Y \sim E_{p,n}(O,\Sigma \otimes \Phi,\psi)$ and the proof can be completed in exactly the same way as the proof of part (a). ■

COROLLARY 8.1.2.1. *Assume that the conditions of Theorem 8.1.2 are satisfied, including the condition of part (b). Assume that a test based on* $f(X)$ *is unbiased under* $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$. *Then the same test is also unbiased under* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$.

If a test based on $f(X)$ *is strictly unbiased under* $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$; *that is, the power function is greater under* H_1 *than under* H_0; *then the same test is also strictly unbiased under* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$.

PROOF: Since the distribution of the test statistic does not depend on ψ, neither does the power function. Now, the unbiasedness is determined by the power function. So, if the test is unbiased when $\psi(z) = \exp\left(-\frac{z}{2}\right)$, then it is also unbiased for the other ψ's. The other part of the statement follows similarly. ■

Next, we look at the hypothesis testing problem from the point of view of invariance.

THEOREM 8.1.3. *Assume we have an observation* X *from the absolutely continuous distribution* $E_{p,n}(M,\Sigma \otimes \Phi,\psi)$, *where* $(M,\Sigma \otimes \Phi) \in \Omega = \Omega_1 \times \Omega_2$, *with* $\Omega_1 \subset \mathbb{R}^{p\times n}$, $\Omega_2 \subset \mathbb{R}^{pn\times pn}$, *and we want to test*

$$H_0\text{: } (M,\Sigma \otimes \Phi) \in \omega \textit{ against } H_1\text{: } (M,\Sigma \otimes \Phi) \in \Omega - \omega,$$

where $\omega = \omega_1 \times \omega_2$, $\omega_1 \subset \Omega_1$, $\omega_2 \subset \Omega_2$. *Let* $O \in \omega_1$. *Assume the hypotheses are invariant under a group* G *of transformations* $g\text{: } \mathbb{R}^{p\times n} \to \mathbb{R}^{p\times n}$, *such that* $g(X) = cX$, $c > 0$ *scalar. Let* $f(X)$ *be a test statistic invariant under* G. *Then, we have the following.*

a) *If* $g(X) = X - M$, $M \in \omega_1$ *are all elements of* G, *then the null distribution of* $f(X)$ *is the same under* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ *as under* $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$.

b) *If* $g(X) = X - M$, and $M \in \Omega_1$ *are all elements of* G, *then the distribution of* $f(X)$, *the null as well as the nonnull is the same under* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ *as under* $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$.

PROOF: The results follow from Theorem 8.1.2, since the invariance of $f(X)$, under the transformation $g(X) = cX$, $c > 0$ implies $f(X) = f(cX)$, $c > 0$; the invariance under the transformation $g(X) = X - M$, $M \in \omega_1$, implies $f(X) = f(X - M)$, $M \in \omega_1$, and the invariance under the transformation $g(X) = X - M$, $M \in \Omega_1$, implies $f(X) = f(X - M)$, $X \in \Omega_1$. ■

COROLLARY 8.1.3.1. *Assume that the conditions of Theorem 8.1.3 are satisfied, including the condition of part (b). Assume that a test based on* $f(X)$ *is unbiased under* $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$. *Then, the same test is also unbiased under* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$.

If a test based on $f(X)$ *is strictly unbiased under* $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$*, then the same test is also strictly unbiased under* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$.

PROOF: It follows from the proof of Theorem 8.1.3 that the conditions of Theorem 8.1.2 are satisfied including the conditions of part (b). Thus Corollary 8.1.2.1 can be applied, which completes the proof. ■

Further aspects of invariant statistics are studied in the following theorems.

THEOREM 8.1.4. *Assume we have an observation* X*, from the absolutely continuous distribution* $E_{p,n}(M,\Sigma \otimes \Phi,\psi)$*, where* $(M,\Sigma \otimes \Phi) \in \Omega \subset \mathbb{R}^{p\times n} \times \mathbb{R}^{pn\times pn}$ *and we want to test*

$$H_0\colon (M,\Sigma \otimes \Phi) \in \omega \text{ against } H_1\colon (M,\Sigma \otimes \Phi) \in \Omega - \omega,$$

where $\omega \subset \Omega$*. Let* G *be a group of the linear transformations* $g\colon \mathbb{R}^{p\times n} \to \mathbb{R}^{p\times n}$*, where* $g(X) = C_1XC_2 + C_2$ *with* C_1: $p \times p$, C_2: $n \times n$, and C_3: $p \times n$ *matrices. Then, the hypotheses are invariant under* G *when* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$*, if and only if the hypotheses are invariant under* G *when* $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$.

Also suppose that the sufficient statistic $T(X)$ *for* $(M,\Sigma \otimes \Phi) \in \Omega$ *is the same under* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ *as under* $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$*. Then,* $f(X)$ *is an invariant of the sufficient statistic under* G *when* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$*, if and only if* $f(X)$ *is an invariant of the sufficient statistic under* G *when* $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$.

PROOF: Assume the hypotheses are invariant under G when $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$. Let $g \in G$. Then, $g(X) = C_1XC_2 + C_3$. Now,

$$g(X) \sim E_{p,n}(C_1MC_2 + C_3,(C_1\Sigma C_1') \otimes (C_2'\Phi C_2),\psi). \tag{8.3}$$

Thus, we have

i) if $(M,\Sigma \otimes \Phi) \in \omega$ then $(C_1MC_2 + C_3,(C_1\Sigma C_1') \otimes (C_2'\Phi C_2)) \in \omega$

and

ii) if $(M,\Sigma \otimes \Phi) \in \Omega\text{-}\omega$ then $(C_1MC_2+C_3,(C_1\Sigma C_1') \otimes (C_2'\Phi C_2)) \in \Omega\text{-}\omega$.

Now, assume $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$. Then

$$g(X) \sim N_{p,n}(C_1MC_2 + C_3,(C_1\Sigma C_1') \otimes (C_2'\Phi C_2)) \tag{8.4}$$

and it follows from (i) and (ii) that the hypotheses are invariant under G when X is normal.

Conversely, assume the hypotheses are invariant under G when $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$ and $g \in G$. Let $g(X) = C_1XC_2 + C_3$. Then, we have (8.4), and (i) and (ii) follow. Since in the case of $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$, (8.3) holds, (i) and (ii) imply that the hypotheses are invariant under G when $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$. Therefore, the invariant test statistics are the same for the normal case as for $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$. Since, by assumption, the sufficient statistic is the same for X normal as for $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$, the statement about $f(X)$ follows. ■

The next result follows easily from the above theorem.

COROLLARY 8.1.4.1. *In Theorem 8.1.4 if G is a group of linear transformations then the hypotheses are invariant under G without specifying what particular absolutely continuous m.e.c. distribution we are working with.*

COROLLARY 8.1.4.2. *Assume that the conditions of Theorem 8.1.4 are satisfied. Also assume the hypotheses are invariant under a group G of linear transformations and that $s(X)$ is a maximal invariant of the sufficient statistic under G when $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$. Then $s(X)$, is also a maximal invariant of the sufficient statistic under G when $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$.*

PROOF: From Theorem 8.1.4 it follows that $s(X)$ is an invariant of the sufficient statistic under G when $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$. On the other hand, the maximal invariance of $s(X)$ means that $s(X) = s(X^*)$ iff there exists $g \in G$, such that $g(X) = X^*$ and this property does not depend on the distribution of X. ■

THEOREM 8.1.5. *Assume we have an observation X from the absolutely continuous distribution $E_{p,n}(M,\Sigma \otimes \Phi,\psi)$, where $(M,\Sigma \otimes \Phi) \in \Omega = \Omega_1 \times \Omega_2$ with $\Omega_1 \subset \mathbb{R}^{p\times n}$, $\Omega_2 \subset \mathbb{R}^{pn\times pn}$. We want to test*

$$H_0\text{: } (M,\Sigma \otimes \Phi) \in \omega \textit{ against } H_1\text{: } (M,\Sigma \otimes \Phi) \in \Omega - \omega,$$

where $\omega = \omega_1 \times \omega_2$, $\omega_1 \subset \Omega_1$, $\omega_2 \subset \Omega_2$. *Let* $O \in \omega_1$. *Assume the hypotheses are invariant under a group* G *of linear transformations* g: $\mathbb{R}^{p\times n} \to \mathbb{R}^{p\times n}$, *where* $g(X) = C_1XC_2 + C_3$ *with* C_1: $p \times p$, C_2: $n \times n$, C_3: $p \times n$ *matrices and the transformations* $g(X) = cX - M$, $c > 0$, $M \in \Omega_1$ *are all elements of* G. *Also suppose that the sufficient statistic* T(X) *for* $(M,\Sigma \otimes \Phi) \in \Omega$ *is the same under* $X \sim E_{p,n}(M,\Sigma \otimes,\psi)$ *as under* $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$.

If s(X) *is a maximal invariant of the sufficient statistic under* G *and a test based on* s(X) *is uniformly most powerful invariant (UMPI) among the tests based on the sufficient statistic in the normal case, then the same test is also UMPI among the tests based on the sufficient statistic when* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$.

PROOF: From Theorem 8.1.4, it follows that the invariant of the sufficient under G, in the normal case are the same as when $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$. Corollary 8.1.4.2 implies that s(X) is a maximal invariant statistic of the sufficient statistic when $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$. Since $O \in \omega_1$ and the transformations $g(X) = cX - M$, $c > 0$, $M \in \Omega_1$ are all elements of G, from part (b) of Theorem 8.1.3 it follows, that s(X) has the same distribution under $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ as under $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$. Since, s(X) is maximal invariant, every invariant test can be expressed as a function of s(X) (see e.g. Lehmann, 1959, p. 216). Therefore, the distributions of the invariant statistics are the same under $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$ as under $X \sim N_{p,n}(M,\Sigma \otimes \Phi)$. Thus, if s(X) is uniformly most powerful among invariant tests in the normal case then it has the same property when $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$. ■

8.2. TWO MODELS

Now, we describe the parameter spaces in which we want to study hypothesis testing problems.

8.2.1. Model I

Let $x_1,x_2,\dots,x_n$ be p-dimensional random vectors, such that $n > p$ and

$x_i \sim E_p(\mu,\Sigma,\psi)$, $i = 1,\ldots,n$. Moreover, assume that x_i, $i = 1,\ldots,n$ are uncorrelated and their joint distribution is elliptically contoured and absolutely continuous. This model can be expressed as

$$X \sim E_{p,n}(\mu e_n', \Sigma \otimes I_n, \psi), \tag{8.5}$$

where $X = (x_1,x_2,\ldots,x_n)$. Then the joint p.d.f. of $x_1,x_2,\ldots,x_n$ can be written as

$$f(X) = \frac{1}{|\Sigma|^n} h\left(\sum_{i=1}^{n} (x_i - \mu)' \Sigma^{-1}(x_i - \mu)\right). \tag{8.6}$$

Assume $\ell(z) = z^{\frac{pn}{2}} h(z)$, $z \geq 0$ has a finite maximum at $z = z_h > 0$. Define

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{and} \quad A = \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})'.$$

Then $\bar{x} = X\frac{e_n}{n}$, $A = X\left(I_n - \frac{e_n e_n'}{n}\right)X'$ and from Theorem 7.2.7, the statistic $T(X) = (\bar{x},A)$ is sufficient for (μ,Σ).

If $\psi(z) = \exp\left(-\frac{z}{2}\right)$, then $X \sim N_{p,n}(\mu e_n', \Sigma \otimes I_n)$. In this case $x_1,x_2,\ldots,x_n$ are independent, and identically distributed random vectors each with distribution $N_p(\mu,\Sigma)$. Inference for this structure has been extensively studied in Anderson (1984).

8.2.2. Model II

Let $x_1^{(i)},x_2^{(i)},\ldots,x_{n_i}^{(i)}$ be p-dimensional random vectors, such that $n_i > p$, $i = 1,\ldots,q$, and $x_j^{(i)} \sim E_p(\mu_i,\Sigma_i,\psi)$, $j = 1,\ldots,n_i$, $i = 1,\ldots,q$. Moreover, assume that $x_j^{(i)}$, $i = 1,\ldots,q$, $j = 1,\ldots,n_i$ are uncorrelated and their joint distribution is also elliptically contoured and absolutely continuous. This model can be expressed as

$$x \sim E_{pn}\left(\begin{pmatrix} \mathbf{e}_{n_1} \otimes \mu_1 \\ \mathbf{e}_{n_2} \otimes \mu_2 \\ \vdots \\ \mathbf{e}_{n_q} \otimes \mu_q \end{pmatrix}, \begin{pmatrix} I_{n_1} \otimes \Sigma_1 & & & \\ & I_{n_2} \otimes \Sigma_2 & & \\ & & \ddots & \\ & & & I_{n_q} \otimes \Sigma_q \end{pmatrix}, \psi\right) \tag{8.7}$$

where $n = \sum_{i=1}^{q} n_i$ and

$$x = \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_{n_1}^{(1)} \\ x_1^{(2)} \\ \vdots \\ x_{n_2}^{(2)} \\ \vdots \\ x_1^{(q)} \\ \vdots \\ x_{n_q}^{(q)} \end{pmatrix}.$$

Then, the joint p.d.f. of $x_j^{(i)}$, $i = 1,\dots,q$, $j = 1,\dots,n_i$ can be written as

$$f(x) = \frac{1}{|\Sigma|^n} h\left(\sum_{i=1}^{q} \sum_{j=1}^{n_i} (x_j^{(i)} - \mu_i)' \Sigma_i^{-1} (x_j^{(i)} - \mu_i)\right). \tag{8.8}$$

Assume $\ell(z) = z^{\frac{pn}{2}} h(z)$, $z \geq 0$ has a finite maximum at $z = z_h > 0$. Define

$$\bar{x}^{(i)} = \frac{1}{n_i}\sum_{j=1}^{n_i} x_j^{(i)}, \quad A_i = \sum_{j=1}^{n_i}(x_j^{(i)} - \bar{x}^{(i)})(x_j^{(i)} - \bar{x}^{(i)})', \quad i = 1,2,\dots,q$$

and $A = \sum_{i=1}^{q} A_i$. Also let $\bar{x} = \frac{1}{n}\sum_{i=1}^{q}\sum_{j=1}^{n_i} x_j^{(i)}$ and

$B = \sum_{i=1}^{q}\sum_{j=1}^{n_i}(x_j^{(i)} - \bar{x})(x_j^{(i)} - \bar{x})'$. Then, using Corollary 7.1.1.1 we get

$$\sum_{j=1}^{n_i}(x_j^{(i)} - \mu_i)'\,\Sigma_i^{-1}(x_j^{(i)} - \mu_i)$$

$$= \operatorname{tr}\left(\sum_{j=1}^{n_i}(x_j^{(i)} - \mu_i)'\,\Sigma_i^{-1}(x_j^{(i)} - \mu_i)\right)$$

$$= \operatorname{tr}\left(\Sigma_i^{-1}\left(\sum_{j=1}^{n_i}(x_j^{(i)} - \mu_i)(x_j^{(i)} - \mu_i)'\right)\right)$$

$$= \operatorname{tr}\left(\Sigma_i^{-1}\left(\sum_{j=1}^{n_i}(x_j^{(i)} - \bar{x}^{(i)})(x_j^{(i)} - \bar{x}^{(i)})' + n(\bar{x}^{(i)} - \mu_i)(\bar{x}^{(i)} - \mu_i)'\right)\right)$$

$$= \operatorname{tr}(\Sigma_i^{-1}(A_i + n(\bar{x}^{(i)} - \mu_i)(\bar{x}^{(i)} - \mu_i)').$$

Thus,

$$f(X) = \frac{1}{|\Sigma|^n}\, h\left(\sum_{i=1}^{q}\operatorname{tr}(\Sigma_i^{-1}(A_i + n(\bar{x}^{(i)} - \mu_i)(\bar{x}^{(i)} - \mu_i)'))\right)$$

hence the statistic $(\bar{x}^{(1)},\dots,\bar{x}^{(q)},A_1,\dots,A_q)$ is sufficient for $(\mu_1,\dots,\mu_q,\Sigma_1,\dots,\Sigma_q)$.

If $\psi(z) = \exp\left(-\frac{z}{2}\right)$, then

$$x \sim N_{pn}\left(\begin{pmatrix} e_{n_1}\otimes\mu_1 \\ e_{n_2}\otimes\mu_2 \\ \vdots \\ e_{n_q}\otimes\mu_q \end{pmatrix}, \begin{pmatrix} I_{n_1}\otimes\Sigma_1 & & & \\ & I_{n_2}\otimes\Sigma_2 & & \\ & & \vdots & \\ & & & I_{n_q}\otimes\Sigma_q \end{pmatrix}\right).$$

In this case $x_1^{(i)},x_2^{(i)},\dots,x_{n_i}^{(i)}$ are independent, and identically distributed random variables each with distribution $N_p(\mu_i,\Sigma_i)$, $i = 1,\dots,q$. Moreover, $x_j^{(i)}$, $i = 1,\dots,q$, $j = 1,\dots,n_i$ are jointly independent. Inference for this structure has been studied in Anderson (1984).

A special case of Model II is when $\Sigma_1 = \dots = \Sigma_q = \Sigma$. Then the model can also be expressed as

$$X \sim E_{p,n}((\mu_1 e'_{n_1};\mu_2 e'_{n_2},\dots,\mu_q e'_{n_q}),\Sigma\otimes I_n,\psi),$$

where $n = \sum_{i=1}^{q} n_i$ and

$$X = (x_1^{(1)},\dots,x_{n_1}^{(1)},x_1^{(2)},\dots,x_{n_2}^{(2)},\dots,x_1^{(q)},\dots,x_{n_q}^{(q)}).$$

This leads to the same joint p.d.f. of $x_j^{(i)}$, $i = 1,\dots,q$, $j = 1,\dots,n_i$ as (8.7); that is

$$f(X) = \frac{1}{|\Sigma|^n}\, h\left(\sum_{i=1}^{q}\sum_{j=1}^{n_i}(x_j^{(i)}-\mu_i)'\,\Sigma_i^{-1}(x_j^{(i)}-\mu_i)\right).$$

8.3. TESTING CRITERIA

In this section, we give results on testing of hypotheses for the two models of Section 8.2. We use the notations of that section. We also use the theorems in Section 8.1 which show, that in certain cases, the hypothesis testing results of the theory of normal distributions can be easily extended to the theory of m.e.c. distributions. This section is based on Anderson and Fang (1982a), Hsu (1985a, 1985b), and Gupta and Varga (1992d).

8.3.1. Testing that a Mean Vector is Equal to a Given Vector

In Model I (see Section 8.2.1) we want to test

$$H_0: \mu = \mu_0 \text{ against } H_1: \mu \neq \mu_0. \tag{8.9}$$

We assume that μ and Σ are unknown and μ_0 is given. Note that problem (8.9) is equivalent to testing

$$H_0: \mu = \mathbf{o} \text{ against } H_1: \mu \neq \mathbf{o}. \tag{8.10}$$

Indeed, if $\mu_0 \neq \mathbf{o}$ then define $x_i^* = x_i - \mu_0$, $i = 1,\ldots,n$ and $\mu^* = \mu - \mu_0$. Then, problem (8.9) becomes

$$H_0: \mu^* = \mathbf{o} \text{ against } H_1: \mu^* \neq \mathbf{o}.$$

Problem (8.10) remains invariant under the group G, where

$$G = \{g \mid g(X) = CX, \ C \text{ is } p \times p \text{ nonsingular}\}. \tag{8.11}$$

Now, we can prove the following theorem.

THEOREM 8.3.1. *The likelihood ratio test (LRT) statistic for problem (8.9) is*

$$T^2 = n(n-1)(\bar{x} - \mu_0)' A^{-1}(\bar{x} - \mu_0).$$

The critical region at level α *is* $T^2 \geq T^2_{p,n-1}(\alpha)$, *where* $T^2_{p,n-1}(\alpha) = \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha)$ and $F_{p,n-p}(\alpha)$ *denotes the* $100\alpha\%$ *point of the* $F_{p,n-p}$ *distribution.*

If H_0 *holds, then* $\frac{n-p}{(n-1)p} T^2 \sim F_{p,n-p}$. *Moreover, if* $\mu_0 = \mathbf{o}$ *then* T^2 *is the maximal invariant of the sufficient statistic under* G.

PROOF: From Theorem 8.1.1 it follows that the LRT statistic is the same as in the normal case; that is, Hotelling's T^2 statistic.. Since problems (8.9) and (8.10) are equivalent, we can focus on (8.10). Then, G satisfies the conditions of part (a) of Theorem 8.1.3 (here $\omega_1 = \{O\}$) and T^2 is invariant under G, so the null distribution of T^2 is the same as in the normal case. Thus, the corresponding results of the normal theory can be used here (see Anderson, 1984, Chapter 5). Since T^2 is the maximal invariant in the normal case, from Corollary 8.1.4.2 it follows that it is also maximal invariant for the present model. ■

The nonnull distribution of T^2 depends on ψ. The nonnull p.d.f. of T^2 was derived by Hsu (1985a). He also showed that if $p > 1$, then the T^2-test is locally most powerful invariant (LMPI). On the other hand, Kariya (1981) proved that if h is a decreasing convex function, then the T^2-test is uniformly most powerful invariant (UMPI).

8.3.2. Testing That a Covariance Matrix is Equal to a Given Matrix

In Model I (see Section 8.2.1), assume that h is decreasing. We want to test

$$H_0\colon \Sigma = \Sigma_0 \text{ against } H_1\colon \Sigma \neq \Sigma_0. \tag{8.12}$$

We assume that μ and Σ are unknown and $\Sigma_0 > O$ is given. It is easy to see that, problem (8.12) is equivalent to testing

$$H_0\colon \Sigma = I_p \text{ against } H_1\colon \Sigma \neq I_p. \tag{8.13}$$

THEOREM 8.3.2. *The LRT statistic for the problem (8.12) is*

$$\tau = |\Sigma_0^{-1} A|^{\frac{n}{2}} h(\mathrm{tr}(\Sigma_0^{-1} A)).$$

The critical region at level α *is*

$$\tau \le \tau_\psi(\alpha),$$

where $\tau_\psi(\alpha)$ *depends on* ψ, *but not on* Σ_0. *The null distribution of* τ *does not depend on* Σ_0.

PROOF: From Theorems 7.1.2 and 7.1.4 it follows that

$$\max_{\mu,\Sigma>O} f(X) = \left|\frac{p}{z_h}(A \otimes I_n)\right|^{-\frac{1}{2}} h(z_h)$$

$$= \left(\frac{p}{z_h}\right)^{-\frac{pn}{2}} h(z_h)\, |A|^{-\frac{n}{2}}.$$

On the other hand, from Theorem 7.1.1, we obtain

$$\max_{\mu,\Sigma=\Sigma_0} f(X) = \frac{1}{|\Sigma_0|^{\frac{n}{2}}} h(\mathrm{tr}((X - \bar{x} e_n')' \Sigma_0^{-1} (X - \bar{x} e_n')))$$

$$= \frac{1}{|\Sigma_0|^{\frac{n}{2}}} h(\mathrm{tr}(\Sigma_0^{-1} A)).$$

Thus, the likelihood ratio test statistic is given by

$$\frac{\max\limits_{\mu,\Sigma=\Sigma_0} f(X)}{\max\limits_{\mu,\Sigma>O} f(X)} = \frac{\frac{1}{|\Sigma_0|^{n/2}} h(\mathrm{tr}(\Sigma_0^{-1} A))}{\left(\frac{p}{z_h}\right)^{-\frac{pn}{2}} h(z_h)\, |A|^{-\frac{n}{2}}}$$

$$= |\Sigma_0^{-1}A|^{\frac{n}{2}}\, h(\mathrm{tr}(\Sigma_0^{-1}A)) \left(\frac{p}{z_h}\right)^{\frac{pn}{2}} \frac{1}{h(z_h)}.$$

Hence, the critical region is of the form $\tau \le \tau_\psi(\alpha)$. Since, (8.13) is equivalent to (8.12), it follows that the null distribution of τ does not depend on Σ_0. Hence, $\tau_\psi(\alpha)$ does not depend on Σ_0, either. ■

In this problem, the distribution of the test statistic τ depends on ψ.

8.3.3. Testing That a Covariance Matrix is Proportional to a Given Matrix

In Model I (see Section 8.2.1) we want to test

$$H_0\colon \Sigma = \sigma^2\Sigma_0 \text{ against } H_1\colon \Sigma \neq \sigma^2\Sigma_0, \tag{8.14}$$

where μ, Σ, σ^2 are unknown, $\sigma^2 > 0$ is a scalar, and $\Sigma_0 > O$ is given. Problem (8.13) remains invariant under the group G, where G is generated by the linear transformations

i) $g(X) = cX$, $c > 0$ scalar and

ii) $g(X) = X + \mathbf{v}\mathbf{e}_n'$, $\mathbf{v}$ is p-dimensional vector.

It is easy to see that problem (8.14) is equivalent to testing

$$H_0\colon \Sigma = \sigma^2 I_p \text{ against } H_1\colon \Sigma \neq \sigma^2 I_p. \tag{8.15}$$

THEOREM 8.3.3. *The LRT statistic for problem (8.14) is*

$$\tau = \frac{|\Sigma_0^{-1}A|^{\frac{n}{2}}}{\mathrm{tr}\left(\frac{1}{p}\Sigma_0^{-1}A\right)^{\frac{np}{2}}}.$$

The critical region at level α is

$$\tau \le \tau(\alpha),$$

where $\tau(\alpha)$ is the same as in the normal case and it does not depend on Σ_0.

The distribution of τ is the same as in the normal case. The null distribution of τ does not depend on Σ_0. τ is an invariant of the sufficient statistic under G.

PROOF: From Theorem 8.1.1 it follows that the LRT statistic is the same as in the normal case. It is easy to see that τ is invariant under G. Moreover, G satisfies the conditions of part (b) of Theorem 8.1.3; therefore, the distribution of τ is the same as in the normal case. Hence, the corresponding results of the normal theory can be used here (see Anderson, 1984, Section 10.7). Since (8.15) is equivalent to (8.14), it follows that the null distribution of τ does not depend on Σ_0. Hence, $\tau(\alpha)$ does not depend on Σ_0, either. ∎

REMARK 8.3.1. Since the distribution of τ is the same as in the normal case, its moments and asymptotic distribution under the null hypothesis are those given by the formulas in Anderson (1984, pp. 430-432), Gupta (1977), and Gupta and Nagar (1987, 1988), for the normal case.

REMARK 8.3.2. Nagao's (1973) criterion,

$$\frac{n-1}{2}\operatorname{tr}\left(\frac{pA}{\operatorname{tr} A} - I_p\right)^2,$$

is also an invariant test criterion under G and hence, it has the same distribution as in the normal case (see also Anderson, 1984, pp. 432-433).

8.3.4. Testing That a Mean Vector and Covariance Matrix are Equal to a Given Vector and Matrix

In Model I (see Section 8.2.1) we want to test

$$H_0\colon \mu = \mu_0 \text{ and } \Sigma = \Sigma_0 \text{ against } H_1\colon \mu \neq \mu_0 \text{ or } \Sigma \neq \Sigma_0, \tag{8.16}$$

where μ, Σ are unknown, and μ_0 and $\Sigma_0 > O$ are given. It is easy to see that problem (8.16) is equivalent to testing

$$H_0\colon \mu = \mathbf{o} \text{ and } \Sigma = I_p \text{ against } H_1\colon \mu \neq \mathbf{o} \text{ or } \Sigma \neq I_p. \tag{8.17}$$

THEOREM 8.3.4. *The LRT statistic for problem (8.16) is*

$$\tau = |\Sigma_0^{-1}A|^{\frac{n}{2}}\, h(\mathrm{tr}(\Sigma_0^{-1}A) + n(\bar{x} - \mu_0)'\Sigma_0^{-1}(\bar{x} - \mu_0)).$$

The critical region at level α *is*

$$\tau \leq \tau_\psi(\alpha),$$

where $\tau_\psi(\alpha)$ *depends on* ψ, *but not* μ_0 *or* Σ_0. *The null distribution of* τ *does not depend on* μ_0 *or* Σ_0.

PROOF: From Theorems 7.1.2 and 7.1.4 it follows that

$$\max_{\mu,\Sigma>O} f(X) = \left(\frac{p}{z_h}\right)^{-\frac{pn}{2}} h(z_h)\,|A|^{-\frac{n}{2}}.$$

On the other hand,

$$\max_{\mu=\mu_0,\Sigma=\Sigma_0} f(X) = \frac{1}{|\Sigma_0|^{\frac{n}{2}}} h(\mathrm{tr}(X - \mu_0 e_n')'\Sigma_0^{-1}(X - \mu_0 e_n'))$$

$$= \frac{1}{|\Sigma_0|^{\frac{n}{2}}} h(\mathrm{tr}(\Sigma_0^{-1}A) + n(\bar{x} - \mu_0)'\Sigma_0^{-1}(\bar{x} - \mu_0)).$$

Therefore, the likelihood ratio test statistic is given by

$$\frac{\max\limits_{\mu=\mu_0,\Sigma=\Sigma_0} f(X)}{\max\limits_{\mu,\Sigma>0} f(X)} = \frac{\dfrac{1}{|\Sigma_0|^{\frac{n}{2}}} h(\mathrm{tr}(\Sigma_0^{-1}A) + n(\bar{x} - \mu_0)'\Sigma_0^{-1}(\bar{x} - \mu_0))}{\left(\dfrac{p}{z_h}\right)^{-\frac{pn}{2}} h(z_h)\,|A|^{-\frac{n}{2}}}$$

$$= \frac{|\Sigma_0^{-1} A|^{\frac{n}{2}}}{h(z_h)} h(\mathrm{tr}(\Sigma_0^{-1} A) + n(\bar{x} - \mu_0)' \Sigma_0^{-1} (\bar{x} - \mu_0)) \left(\frac{p}{z_h}\right)^{\frac{pn}{2}} .$$

Thus ,the critical region is of the form $\tau \le \tau_\psi(\alpha)$. Since (8.17) is equivalent to (8.16), it follows that the null distribution of τ does not depend on μ_0 or Σ_0. Hence, $\tau_\psi(\alpha)$ does not depend on μ_0 or Σ_0, either. ■

In this problem, the distribution of the test statistic τ depends on ψ. Nevertheless, Quan and Fang (1987) have proved that the test defined in Theorem 8.3.4 is unbiased.

8.3.5. Testing That a Mean Vector is Equal to a Given Vector and a Covariance Matrix is Proportional to a Given Matrix

In Model I (see Section 8.2.1), we want to test

$$H_0:\ \mu = \mu_0 \text{ and } \Sigma = \sigma^2\Sigma_0 \text{ against } H_1:\ \mu \ne \mu_0 \text{ or } \Sigma \ne \sigma^2\Sigma_0, \tag{8.18}$$

where μ, Σ, σ^2 are unknown, $\sigma^2 > 0$ is a scalar , and $\mu_0, \Sigma_0 > O$ are given.

Note that problem (8.18) is equivalent to testing

$$H_0:\ \mu = \mathbf{o} \text{ and } \Sigma = \sigma^2 I_p \text{ against } H_1:\ \mu \ne \mathbf{o} \text{ or } \Sigma \ne \sigma^2 I_p. \tag{8.19}$$

Problem (8.19) remains invariant under the group G, where

$$G = \{g \mid g(X) = cX,\ c > 0 \text{ scalar}\}.$$

THEOREM 8.3.5. *The LRT statistic for problem (8.18) is*

$$\tau = \frac{|\Sigma_0^{-1}A|^{\frac{n}{2}}}{\left(\operatorname{tr}\left(\frac{1}{p}\Sigma_0^{-1}A\right)+\frac{n}{p}(\bar{x}-\mu_0)'\Sigma_0^{-1}(\bar{x}-\mu_0)\right)^{\frac{pn}{2}}}.$$

The critical region at level α *is*

$$\tau \leq \tau(\alpha),$$

where $\tau(\alpha)$ *is the same as in the normal case and it does not depend on* μ_0 *or* Σ_0. *The null distribution of* τ *is the same as in the normal case and it does not depend on* μ_0 *or* Σ_0. *Moreover, if* $\mu_0 = \mathbf{o}$ *then* τ *is an invariant of the sufficient statistic under* G.

PROOF: From Theorem 8.1.1 it follows that the LRT statistic is the same as in the normal case. However, since Anderson (1984) does not give the corresponding result for the normal case, we derive it here.

So, assume $X \sim N_{p,n}(\mu e_n', \Sigma \otimes I_n)$ and we want to test (8.18). Then, from Theorem 7.1.2, Remark 7.1.1 and Theorem 7.1.4 it follows that

$$\max_{\mu,\Sigma>O} f(X) = \left|\frac{A}{n}\otimes I_n\right|^{-\frac{1}{2}} \frac{1}{(2\pi)^{\frac{pn}{2}}} e^{-\frac{pn}{2}}$$

$$= \left(\frac{n}{2\pi}\right)^{\frac{pn}{2}} e^{-\frac{pn}{2}} |A|^{-\frac{n}{2}}.$$

Next, we want to maximize $f(X)$ under H_0 where

$$f(X) = \frac{1}{(2\pi)^{\frac{pn}{2}} |\sigma^2\Sigma_0|^{\frac{n}{2}}} \exp\left\{-\frac{1}{2}[\operatorname{tr}((\sigma^2\Sigma_0)^{-1}A)+n(\bar{x}-\mu_0)'(\sigma^2\Sigma_0)^{-1}(\bar{x}-\mu_0)]\right\}.$$

Then, $\frac{\partial \log f(X)}{\partial\sigma^2} = -\frac{pn}{2\sigma^2} + \frac{1}{2\sigma^4}[\operatorname{tr}(\Sigma_0^{-1}A) + n(\bar{x}-\mu_0)'\Sigma_0^{-1}(\bar{x}-\mu_0)]$ and from

$\frac{\partial \log f(X)}{\partial \sigma^2} = 0$, we obtain

$$\hat{\sigma}^2 = \frac{\operatorname{tr}(\Sigma_0^{-1}A) + n(\bar{x} - \mu_0)'\Sigma_0^{-1}(\bar{x} - \mu_0)}{pn}.$$

Thus,

$$\max_{\mu=\mu_0, \Sigma=\sigma^2\Sigma_0>O} f(X) = \left(\frac{pn}{2\pi}\right)^{\frac{pn}{2}} e^{-\frac{pn}{2}} |\Sigma_0|^{-\frac{n}{2}} (\operatorname{tr}(\Sigma_0^{-1}A) + n(\bar{x} - \mu_0)'\Sigma_0^{-1}(\bar{x} - \mu_0))^{-\frac{pn}{2}}.$$

Therefore, the likelihood ratio test statistic is given by

$$\frac{\max\limits_{\mu=\mu_0, \Sigma=\sigma^2\Sigma_0>O} f(X)}{\max\limits_{\mu, \Sigma>O} f(X)}$$

$$= \frac{\left(\frac{pn}{2\pi}\right)^{\frac{pn}{2}} e^{-\frac{pn}{2}} |\Sigma_0|^{-\frac{n}{2}} (\operatorname{tr}(\Sigma_0^{-1}A) + n(\bar{x} - \mu_0)'\Sigma_0^{-1}(\bar{x} - \mu_0))^{-\frac{pn}{2}}}{\left(\frac{n}{2\pi}\right)^{\frac{pn}{2}} e^{-\frac{pn}{2}} |A|^{-\frac{n}{2}}}$$

$$= \frac{|\Sigma_0^{-1}A|^{\frac{n}{2}}}{\left(\operatorname{tr}\left(\frac{1}{p}\Sigma_0^{-1}A\right) + \frac{n}{p}(\bar{x} - \mu_0)'\Sigma_0^{-1}(\bar{x} - \mu_0)\right)^{\frac{np}{2}}}.$$

Thus, the critical region is of the form $\tau \leq \tau(\alpha)$. Since problems (8.18) and (8.19) are equivalent, we can focus on (8.19). It is easy to see that τ is invariant under G if $\mu_0 = \mathbf{o}$. Moreover G satisfies the conditions of part (a) of Theorem 8.1.3, so the null distribution of τ is the same as in the normal case and, since problems (8.18) and (8.19) are equivalent, it does not depend on μ_0 or Σ_0. Therefore, $\tau(\alpha)$ is the same as in the normal case and it does not depend on μ_0 or Σ_0. ■

In this case, the nonnull distribution of τ depends on ψ. Nevertheless Quan and Fang (1987) have proved that the LRT is unbiased if h is decreasing.

8.3.6. Testing Lack of Correlation Between Sets of Variates

In Model I (see Section 8.2.1), we partition x_i into q subvectors

$$x_i = \begin{pmatrix} x_i^{(1)} \\ x_i^{(2)} \\ \vdots \\ x_i^{(q)} \end{pmatrix},$$

where $x_i^{(j)}$ is p_j-dimensional, $j = 1,\ldots,q$, $i = 1,\ldots,n$. Partition Σ and A into

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1q} \\ \Sigma_{21} & \Sigma_{22} & \cdots & \Sigma_{2q} \\ & \vdots & & \\ \Sigma_{q1} & \Sigma_{q2} & \cdots & \Sigma_{qq} \end{pmatrix}, \qquad A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1q} \\ A_{21} & A_{22} & \cdots & A_{2q} \\ & \vdots & & \\ A_{q1} & A_{q2} & \cdots & A_{qq} \end{pmatrix}$$

where Σ_{jj} and A_{jj} are $p_j \times p_j$, $j = 1,\ldots,q$. We want to test

$$H_0: \Sigma_{jk} = O \text{ if } 1 \le j < k \le q \text{ against}$$

$$H_1: \text{there exists } j, k \text{ such that } \Sigma_{jk} \ne O. \tag{8.20}$$

Problem (8.20) remains invariant under the group G, where G is generated by the linear transformations

i) $g(X) = CX$ with $C = \begin{pmatrix} C_1 & & & \\ & C_2 & & \\ & & \ddots & \\ & & & C_q \end{pmatrix}$, where C_j is $p_j \times p_j$ nonsingular matrix, $j = 1,\ldots,q$,

ii) $g(X) = X + \mathbf{v}\mathbf{e}_n'$, where $\mathbf{v}$ is p-dimensional vector.

THEOREM 8.3.6. *The LRT statistic for problem (8.20) is*

$$\tau = \frac{|A|}{\prod_{i=1}^{q} |A_{ii}|}.$$

The critical region at level α *is*

$$\tau \le \tau(\alpha),$$

where $\tau(\alpha)$ *is the same as in the normal case. The distribution of* τ *is the same as in the normal case. If* H_0 *holds, then* $\tau \approx \prod_{i=2}^{q} v_i$, *where* $v_2, v_3, \ldots, v_q$ *are independent and* $v_i \sim U_{p_i, \bar{p}_i, n-\bar{p}_i}$, *with* $\bar{p}_i = \sum_{j=1}^{i-1} p_j$, $i = 2,\ldots,q$. *The LRT is strictly unbiased; that is, if* H_1 *holds, then the probability of rejecting* H_0 *is greater than* α. τ *is an invariant of the sufficient statistic under* G.

PROOF: From Theorem 8.1.1 it follows that the LRT statistic is the same as in the normal case. It is easy to see that τ is invariant under G. Moreover G satisfies the conditions of part (b) of Theorem 8.1.3. Therefore, the distribution of τ is the same as in the normal case. Hence, the corresponding results of the normal theory can be used here (see Anderson, 1984, Chapter 9). The strict unbiasedness follows from Corollary 8.1.3.1 and the normal theory. ■

REMARK 8.3.3. Since the distribution of τ is the same as in the normal case, its moments and asymptotic distribution under the null hypothesis are these given by the formulas in Anderson (1984, pp. 384-387), and Nagar and Gupta (1986), for the normal case.

REMARK 8.3.4. Nagao's criterion (see Anderson, 1984, p. 388) is also invariant under G and hence, it has the same distribution as in the normal case. Thus, its asymptotic distribution under the null hypothesis is given by the formulas in Anderson (1984, p. 388).

Hsu (1985b) has proved that the LRT in Theorem 8.3.6 is admissible.

8.3.7. Testing That a Correlation Coefficient is Equal to a Given Number

In Model I (see Section 8.2.1), let $p = 2$. Then Σ and A can be written as

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \text{ and } A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

where $\sigma_1^2 = \sigma_{11}$, $\sigma_2^2 = \sigma_{22}$ and $\rho = \dfrac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}}$. We want to test

$$H_0: \rho = \rho_0 \text{ against } H_1: \rho \neq \rho_0, \tag{8.21}$$

where μ and Σ are unknown and $|\rho_0| < 1$ is given. Problem (8.21) remains invariant under the group G, where G is generated by the linear transformations

i) $g(X) = CX$, with $C = \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix}$ where c_1 and c_2 are positive

scalars and

ii) $g(X) = X + \mathbf{v}\mathbf{e}_n'$, where $\mathbf{v}$ is p-dimensional vector.

THEOREM 8.3.7. *The statistic*

$$r = \frac{a_{12}}{\sqrt{a_{11}a_{22}}} \ ,$$

is maximal invariant of the sufficient statistic under G. *The distribution of* r *is the same as in the normal case. The LRT for problem (8.21) at level* α *has the critical region*

$$r \le r_1(\alpha) \ \text{or} \ r \ge r_2(\alpha),$$

where $r_1(\alpha) = \dfrac{\rho_0 c - (1 - \rho_0^2)\sqrt{1 - c}}{\rho_0^2 c + 1 - \rho_0^2}$, $r_2(\alpha) = \dfrac{\rho_0 c + (1 - \rho_0^2)\sqrt{1 - c}}{\rho_0^2 c + 1 - \rho_0^2}$ *and* c *is chosen such that under* H_0

$$P(r \le r_1(\alpha)) + P(r \ge r_2(\alpha)) = \alpha.$$

The values of $r_1(\alpha)$ *and* $r_2(\alpha)$ *depend on* ρ_0, *but they are the same as in the normal case.*

PROOF: Since G satisfies the conditions of Corollary 8.1.4.2 and in the normal case, r is a maximal invariant under G (see Anderson, 1984, p. 114), it follows that r is also maximal invariant under G in the present case. Moreover, G satisfies the conditions of part (b) of Theorem 8.1.3 and r is invariant under G. Therefore, the distribution of r is the same as in the normal case. Furthermore it follows from Theorem 8.1.1 that the LRT statistic is the same as in the normal case. Thus, the corresponding results of the normal theory can be used (see Anderson, 1984, Section 4.2). ■

REMARK 8.3.5. Since the distribution of r is the same as in the normal case, its asymptotic distribution, as well as the asymptotic distribution of Fisher's z statistic, are the same as those given by the formulas in Anderson (1984, pp. 122-124), and Konishi and Gupta (1989). Therefore, the tests based on Fisher's z statistic can also be used here.

REMARK 8.3.6. It is known (e.g., see Anderson, 1984, p. 114) that in the normal case, for the problem

$$H_0\colon\ \rho = \rho_0 \ \text{ against } \ H_1\colon\ \rho > \rho_0$$

the level α test, whose critical region has the form

$$r > r_3(\alpha)$$

is UMPI. Then, it follows from Theorem 8.1.5 that the same test is also UMPI in the present case.

8.3.8. Testing That a Partial Correlation Coefficient is Equal to a Given Number

In Model I (see Section 8.2.1) partition Σ and A as

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \text{ and } A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

where Σ_{11} and A_{11} are 2×2. Assume $p \ge 3$, and define

$$\Sigma_{11\cdot 2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \text{ and } A_{11\cdot 2} = A_{11} - A_{12}A_{22}^{-1}A_{21}.$$

We use the notation

$$\Sigma_{11\cdot 2} = \begin{pmatrix} \sigma_{11\cdot 3,\ldots,p} & \sigma_{12\cdot 3,\ldots,p} \\ \sigma_{21\cdot 3,\ldots,p} & \sigma_{22\cdot 3,\ldots,p} \end{pmatrix} \text{ and } A_{11\cdot 2} = \begin{pmatrix} a_{11\cdot 3,\ldots,p} & a_{12\cdot 3,\ldots,p} \\ a_{21\cdot 3,\ldots,p} & a_{22\cdot 3,\ldots,p} \end{pmatrix}$$

Then, $\rho_{12\cdot 3,\ldots,p} = \dfrac{\sigma_{12\cdot 3,\ldots,p}}{\sqrt{\sigma_{11\cdot 3,\ldots,p}\sigma_{22\cdot 3,\ldots,p}}}$ is called the partial correlation between the first two variables, having the $(p - 2)$ variables fixed (see Anderson, 1984, p. 37). We want to test

$$H_0\colon \rho_{12\cdot 3,\ldots,p} = \rho \text{ against } H_1\colon \rho_{12\cdot 3,\ldots,p} \neq \rho_0, \tag{8.22}$$

where μ and Σ are unknown and $|\rho_0| < 1$ is given.

Problem (8.22) remains invariant under the group G, where G is generated by the linear transformations

i) $g(X) = CX$ with $C = \begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix}$, where $C_1 = \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix}$, c_1, c_2 positive scalars, C_2 is any $2 \times (p - 2)$ matrix, $C_3 = O$ is $(p - 2) \times 2$ matrix, and $C_4 = I_{p-2}$.

ii) $g(X) = CX$ with $C = \begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix}$, where $C_1 = I_2$, C_2 is $2 \times (p - 2)$ matrix, $C_2 = O$, $C_3 = O$ is $(p - 2) \times 2$ matrix, and C_4 is a $(p - 2) \times (p - 2)$ nonsingular matrix.

iii) $g(X) = X + \mathbf{v}\mathbf{e}_n'$, where $\mathbf{v}$ is p-dimensional vector.

THEOREM 8.3.8. *The statistic*

$$r_{12\cdot 3,\ldots,p} = \frac{a_{12\cdot 3,\ldots,p}}{\sqrt{a_{11\cdot 3,\ldots,p}\, a_{22\cdot 3,\ldots,p}}}$$

is a maximal invariant of the sufficient statistic under G. *The distribution of* $r_{12\cdot 3,\ldots,p}$ *is the same as in the normal case and it is the same as the distribution of* r *in Section 8.3.7, where we replace* n *by* $n - p + 2$.

PROOF: Since G satisfies the conditions of Corollary 8.1.4.2 and in the normal case, $r_{12\cdot 3,\ldots,p}$ is a maximal invariant under G (see Anderson, 1984, p. 152), it follows that r is also a maximal invariant in the present case. Moreover, G satisfies the conditions of part (b) of Theorem 8.1.3 and $r_{12\cdot 3,\ldots,p}$ is invariant under G. Therefore, the distribution of $r_{12\cdot 3,\ldots,p}$ is the same as in the normal case. Furthermore, in the normal case, $r_{12\cdot 3,\ldots,p}$ has the same distribution as r when n is replaced by $n - p + 2$ (see Anderson, 1984, p. 133). ■

Because of the relation between the distribution of $r_{12\cdot3,\ldots,p}$ and r, all statistical methods mentioned in Section 8.3.7 can also be used here. So, for example, Fisher's z test can be applied here. Also, the test based on $r_{12\cdot3,\ldots,p}$ for the problem

$$H_0\colon \rho_{12\cdot3,\ldots,p} = \rho_0 \text{ against } H_1\colon \rho_{12\cdot3,\ldots,p} > \rho_0$$

is UMPI.

8.3.9. Testing That a Multiple Correlation Coefficient is Equal to Zero

In Model I (see Section 8.2.1), let us partition Σ and A as

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_1' \\ \sigma_1 & \Sigma_{22} \end{pmatrix} \text{ and } A = \begin{pmatrix} a_{11} & a_1' \\ a_1 & A_{22} \end{pmatrix},$$

where σ_{11} and a_{11} are one-dimensional. Then, $\bar{\rho}_{1\cdot2,\ldots,p} = \sqrt{\dfrac{\sigma_1'\Sigma_{22}\sigma_1}{\sigma_{11}}}$ is called the multiple correlation between the first variable and the other $(p - 1)$ variables (see Anderson, 1984, p. 40). We want to test

$$H_0\colon \bar{\rho}_{1\cdot2,\ldots,p} = 0 \text{ against } H_1\colon \bar{\rho}_{1\cdot2,\ldots,p} \neq 0, \tag{8.23}$$

where μ and Σ are unknown. Problem (8.23) remains invariant under the group G, where G is generated by the linear transformations

i) $g(X) = CX$ with $C = \begin{pmatrix} c_1 & \mathbf{o}' \\ \mathbf{o} & C_2 \end{pmatrix}$, where c_1 is nonnegative scalar and C_2 is a $(p - 2) \times (p - 2)$ nonsingular matrix.

ii) $g(X) = X + \mathbf{v}\mathbf{e}_n'$, where $\mathbf{v}$ is p-dimensional vector.

THEOREM 8.3.9. *The statistic*

$$\bar{r}_{1\cdot 2,\ldots,p} = \sqrt{\frac{a_1' A_{22}^{-1} a_1}{a_{11}}}$$

is maximal invariant of the sufficient statistic under G. *The distribution of* r *is the same as in the normal case. If* H_0 *holds, then* $\frac{n-p}{p-1} \frac{\bar{r}^2_{1\cdot 2,\ldots,p}}{1-\bar{r}^2_{1\cdot 2,\ldots,p}} \sim F_{p-1,n-p}$. *The LRT for problem (8.23) at level* α *has the critical region*

$$\frac{n-p}{p-1} \frac{\bar{r}^2_{1\cdot 2,\ldots,p}}{1-\bar{r}^2_{1\cdot 2,\ldots,p}} \geq F_{p-1,n-p}(\alpha),$$

where $F_{p-1,n-p}(\alpha)$ *denotes the 100α% point of the* $F_{p,n-p}$ *distribution. The LRT is UMPI.*

PROOF: Since G satisfies the conditions of Corollary 8.1.4.2 and in the normal case, $\bar{r}_{1\cdot 2,\ldots,p}$ is a maximal invariant under G (see Anderson, 1984, p. 148), it follows that $\bar{r}_{1\cdot 2,\ldots,p}$ is also maximal invariant under G in the present case. Moreover, G satisfies the conditions of part (b) of Theorem 8.1.3 and $\bar{r}_{1\cdot 2,\ldots,p}$ is invariant under G. Therefore, the distribution of $\bar{r}_{1\cdot 2,\ldots,p}$ is the same as in the normal case. Furthermore, from Theorem 8.1.1 it follows that the LRT statistic is the same as in the normal case. Thus, the corresponding results of the normal theory can be used (see Anderson, 1984, Section 4.4). It follows from Anderson (1984, p. 148), that in the normal case, the LRT is UMPI. Therefore, by Theorem 8.1.5, the LRT is also UMPI here. ■

REMARK 8.3.7. Since the distribution of $\bar{r}_{1\cdot 2,\ldots,p}$ is the same as in the normal case, its moments and distribution under the nonnull hypothesis are those given by the formulas in Anderson (1984, pp. 138-148), for the normal case.

8.3.10. Testing Equality of Means

In Model II (see Section 8.2.2), let $\Sigma_1 = \Sigma_2 = \dots = \Sigma_q = \Sigma$. We want to test

$$H_0: \mu_1 = \mu_2 = \dots = \mu_q \text{ against}$$

$$H_1: \text{there exist } 1 \le j < k \le q, \text{ such that } \mu_j \ne \mu_k, \tag{8.24}$$

where μ_i, $i = 1,2,\dots,q$ and Σ are unknown. Problem (8.24) remains invariant under the group G, where G is generated by the linear transformations

i) $g(X) = (I_n \otimes C)X$, where C is $p \times p$ nonsingular matrix,

ii) $g(X) = X - \mathbf{e}_n \otimes \mathbf{v}$, where $\mathbf{v}$ is p-dimensional vector.

THEOREM 8.3.10. *The LRT statistic for problem (8.24) is*

$$\tau = \frac{|A|}{|B|}.$$

The critical region at level α *is*

$$\tau \le U_{p,q-1,n}(\alpha),$$

where $U_{p,q-1,n}(\alpha)$ *denotes the* $100\alpha\%$ *point of the* $U_{p,q-1,n}$ *distribution. If* H_0 *holds, then* $\tau \approx U_{p,q-1,n}$. τ *is an invariant of the sufficient statistic under* G.

PROOF: From Theorem 8.1.1 it follows that the LRT statistic is the same as in the normal case. It is easy to see that τ is invariant under G. Moreover, G satisfies the conditions of part (a) of Theorem 8.1.3. Therefore, the null distribution of τ is the same as in the normal case. Hence, the corresponding results of the normal theory can be used here, see Anderson (1984, Section 8.8), Pillai and Gupta (1969), Gupta (1971, 1975), Gupta et al. (1975), and Gupta and Javier (1986). ■

The nonnull distribution of τ depends on ψ. Nevertheless, Quan (1987) has proved that if h is decreasing then the LRT is unbiased.

8.3.11. Testing Equality of Covariance Matrices

In Model II (see Section 8.2.2), we want to test

$$H_0: \Sigma_1 = \Sigma_2 = \ldots = \Sigma_q \text{ against}$$

$$H_1: \text{ there exist } 1 \leq j < k < q, \text{ such that } \Sigma_j \neq \Sigma_k, \tag{8.25}$$

where μ_i and Σ_i, $i = 1,2,\ldots,q$ are unknown. Problem (8.25) remains invariant under the group G, where G is generated by the linear transformations

i) $g(X) = (I_n \otimes C)X$, where C is $p \times p$ nonsingular matrix,

ii) $g(X) = X - \begin{pmatrix} \mathbf{e}_{n_1} \otimes \mathbf{v}_1 \\ \mathbf{e}_{n_2} \otimes \mathbf{v}_2 \\ \vdots \\ \mathbf{e}_{n_q} \otimes \mathbf{v}_q \end{pmatrix}$, where $\mathbf{v}_i$ is p-dimensional vector,

$i = 1,2,\ldots,q$.

THEOREM 8.3.11. *The LRT statistic for problem (8.25) is*

$$\tau = \frac{\prod_{i=1}^{q} |A_i|^{\frac{n_i}{2}}}{|A|^{\frac{n}{2}}} \frac{\prod_{i=1}^{q} n_i^{\frac{pn_i}{2}}}{n^{\frac{pn}{2}}}.$$

The critical region at level α *is*

$$\tau \leq \tau(\varepsilon),$$

where $\tau(\alpha)$ *is the same as in the normal case. The distribution of* τ *is the same as in the normal case.* τ *is an invariant of the sufficient statistic under* G.

PROOF: From Theorem 8.1.1 it follows that the LRT statistic is the same as in the normal case. It is easy to see that τ is invariant under G. Moreover, G satisfies the conditions of part (b) of Theorem 8.1.3. Therefore, the distribution of τ is the same as in the normal case. Hence, the corresponding results of the normal theory can be used here, see Anderson (1984, Section 10.2), and Gupta and Tang (1984). ■

REMARK 8.3.8. Bartlett's modified LRT statistic

$$\tau_1 = \frac{\prod_{i=1}^{q} |A_i|^{\frac{n_i-1}{2}}}{|A|^{\frac{n-q}{2}}}$$

is also invariant under G. So from Theorem 8.1.3 it follows that the distribution of τ_1 is the same as in the normal case. Quan (1987) showed that the α level test with critical region

$$\tau_1 \leq \tau_1(\varepsilon)$$

is unbiased if h is decreasing.

Nagao's test statistic

$$\tau_2 = (n-1)^2 \sum_{i=1}^{q} \frac{\operatorname{tr}(A_i A^{-1} - I_p)^2}{n_i - 1}$$

is also invariant under G and has the same distribution as in the normal case.

For further details on Bartlett's modified LRT statistic and Nagao's test statistic, see Anderson (1984, pp. 406-408).

8.3.12. Testing Equality of Means and Covariance Matrices

In Model II (see Section 8.2.2), we want to test

$$H_0:\ \mu_1 = \mu_2 = \ldots = \mu_q \text{ and } \Sigma_1 = \Sigma_2 = \ldots = \Sigma_q \text{ against} \tag{8.26}$$

$$H_1:\ \text{there exist } 1 \le j < k \le q, \text{ such that } \mu_j \ne \mu_k \text{ or } \Sigma_j \ne \Sigma_k,$$

where μ_i and Σ_i, $i = 1,2,\ldots,q$ are unknown. Problem (8.26) remains invariant under group G, where G is generated by the linear transformations

i) $g(X) = (I_n \otimes C)X$, where C is $p \times p$ nonsingular matrix,

ii) $g(X) = X - \mathbf{e}_n \otimes v$, where v is p-dimensional vector.

THEOREM 8.3.12. *The LRT statistic for problem (8.26) is*

$$\tau = \frac{\prod_{i=1}^{q} |A_i|^{\frac{n_i}{2}}}{|B|^{\frac{n}{2}}}.$$

The critical region at level α *is*

$$\tau \le \tau(\varepsilon)$$

where $\tau(\alpha)$ *is the same as in the normal case. The null distribution of* τ *is the same as in the normal case.* τ *is an invariant of the sufficient statistic under* G.

PROOF: From 8.1.1 it follows that the LRT statistic is the same as in the normal case. It is easy to see that τ is invariant under G. Moreover, G satisfies the conditions of part (a) of Theorem 8.1.3. Therefore, the distribution of τ is the same as in the normal case. Hence the corresponding results of the normal theory can be used here (see Anderson, 1984, Section 10.3). ■

The nonnull distribution of τ depends on ψ. Nevertheless, Quan and Fang (1987) have proved that the LRT is unbiased if h is decreasing.

REMARK 8.3.9. Since the null distribution of τ is the same as in the normal case, its moments, distribution, and asymptotic distribution under the null hypothesis are those given by the formulas in Anderson (1984, Sections 10.4-5) for the normal case.

REMARK 8.3.10. Bartlett's modified LRT statistic

$$\tau_1 = \frac{\prod_{i=1}^{q} |A_i|^{\frac{n_i-1}{2}}}{|B|^{\frac{n-q}{2}}}$$

is also invariant under G. So from Theorem 8.1.3 it follows that its null distribution is the same as in the normal case (see Anderson, 1984, pp. 406-407).

CHAPTER 9

LINEAR MODELS

9.1. ESTIMATION OF THE PARAMETERS IN THE MULTIVARIATE LINEAR REGRESSION MODEL

Let $x_1,x_2,\dots,x_n$ be p-dimensional vectors, such that $x_i \sim E_p(Bz_i,\Sigma,\psi)$, where z_i is a q-dimensional known vector, $i = 1,\dots,n$, and B is a $p \times q$ unknown matrix. Moreover, assume that x_i, $i = 1,\dots,n$ are uncorrelated and their joint distribution is elliptically contoured and absolutely continuous. This model can be expressed as

$$X \sim E_{p,n}(BZ,\Sigma \otimes I_n,\psi), \tag{9.1}$$

where $X = (x_1,x_2,\dots,x_n)$; $Z = (z_1,z_2,\dots,z_n)$ is a $q \times n$ known matrix; $B\ (p \times q)$ and $\Sigma\ (p \times p)$ are unknown matrices. Assume $\mathrm{rk}(Z) = q$ and $p + q \le n$. The joint p.d.f. of $x_1,x_2,\dots,x_n$ can be written as

$$f(X) = \frac{1}{|\Sigma|^n} h\left(\sum_{i=1}^{n} (x_i - Bz_i)' \Sigma^{-1}(x_i - Bz_i)\right)$$

$$= \frac{1}{|\Sigma|^n} h(\mathrm{tr}(X - BZ)' \Sigma^{-1}(X - BZ)). \tag{9.2}$$

Assume $\ell(z) = z^{\frac{pn}{2}} h(z)$, $z \ge 0$ has a finite maximum at $z = z_h > 0$. First, we find the MLE's of B and Σ.

THEOREM 9.1.1. *The MLE's of* B *and* Σ *for the model (9.1) are given by*

$$\hat{B} = XZ'(ZZ')^{-1} \tag{9.3}$$

and

$$\hat{\Sigma} = \frac{p}{z_h} X(I_n - Z'(ZZ')^{-1}Z)X'. \tag{9.4}$$

PROOF: From Theorem 1.3.3 it follows that $rk(ZZ') = q$ and since ZZ' is $q \times q$, it is nonsingular.

Note that B and BZ determine each other uniquely. Indeed, from B we get BZ by postmultiplying B by Z, and from BZ we can obtain B by postmultiplying BZ by $Z'(ZZ')^{-1}$. Hence finding the MLE of B is equivalent to finding the MLE of BZ.

If $X \sim N_{p,n}(BZ, \Sigma \otimes I_n)$, then the MLE's of B and Σ are (see Anderson, 1984, Section 8.2),

$$B^* = XZ'(ZZ')^{-1}$$

and

$$\Sigma^* = \frac{1}{n}(X - B^*Z)(X - B^*Z)'.$$

We can rewrite Σ^* as

$$\Sigma^* = \frac{1}{n} X(I_n - Z'(ZZ')^{-1}Z)(I_n - Z'(ZZ')^{-1}Z)X'$$

$$= \frac{1}{n} X(I_n - Z'(ZZ')^{-1}Z)X'.$$

Then, by using Theorem 7.1.2, we obtain the MLE's of B and Σ as

$$\hat{B} = XZ'(ZZ')^{-1}$$

and

$$\hat{\Sigma} = \frac{pn}{z_h}\frac{1}{n} X(I_n - Z'(ZZ')^{-1}Z)X'$$
$$= \frac{p}{z_h} X(I_n - Z'(ZZ')^{-1}Z)X'. \blacksquare$$

The distributions of $\hat{B}$ and $\hat{\Sigma}$ can also be obtained, and are given in the following theorem.

THEOREM 9.1.2. *The distributions of the MLE's of* B *and* $\frac{z_h}{p}\Sigma$ *for the model (9.1) are given by*

$$\hat{B} \sim E_{p,q}(B,\Sigma \otimes (ZZ')^{-1},\psi) \tag{9.5}$$

and

$$\frac{z_h}{p}\hat{\Sigma} \sim G_{p,1}\left(\Sigma, \frac{n-q}{2}, \psi\right). \tag{9.6}$$

PROOF: From (9.1) and (9.3) we obtain

$$\hat{B} \sim E_{p,q}(BZZ'(ZZ')^{-1}, \Sigma \otimes (ZZ')^{-1} ZZ'(ZZ')^{-1},\psi)$$
$$= E_{p,q}(B,\Sigma \otimes (ZZ')^{-1},\psi).$$

Then from (9.4) we get

$$\frac{z_h}{p}\hat{\Sigma} = X(I_n - Z'(ZZ')^{-1}Z)X'$$
$$= X(I_n - Z'(ZZ')^{-1}Z)(X(I_n - Z'(ZZ')^{-1}Z))'. \tag{9.7}$$

From (9.1) we get

$$X(I_n - Z'(ZZ')^{-1}Z) \sim E_{p,n}(BZ(I_n - Z'(ZZ')^{-1}Z),\Sigma \otimes (I_n - Z'(ZZ')^{-1}Z)$$
$$(I_n - Z'(ZZ')^{-1}Z),\psi)$$

$$= E_{p,n}(O,\Sigma \otimes (I_n - Z'(ZZ')^{-1}Z),\psi). \tag{9.8}$$

Now define the $p \times n$ random matrix, Y, by

$$Y \sim E_{p,n}(O,\Sigma \otimes I_n,\psi). \tag{9.9}$$

Then,

$$\begin{aligned} Y(I_n - Z'(ZZ')^{-1}Z)Y' &= Y(I_n - Z'(ZZ')^{-1}Z)(I_n - Z'(ZZ')^{-1}Z)Y' \\ &= (Y(I_n - Z'(ZZ')^{-1}Z))(Y(I_n - Z'(ZZ')^{-1}Z))' \end{aligned} \tag{9.10}$$

and since

$$Y(I_n - Z'(ZZ')^{-1}Z) \sim E_{p,n}(O,\Sigma \otimes (I_n - Z'(ZZ')^{-1}Z),\psi), \tag{9.11}$$

from (9.7), (9.8), (9.10), and (9.11) we get

$$\frac{z_n}{p}\hat{\Sigma} \approx Y(I_n - Z'(ZZ')^{-1}Z)Y'. \tag{9.12}$$

However, the matrix $(I_n - Z'(ZZ')^{-1}Z)$ is idempotent of rank $n - q$. Hence, using Theorem 5.1.7, we get

$$Y(I_n - Z'(ZZ')^{-1}Z)Y' \sim G_{p,1}\left(\Sigma, \frac{n-q}{2}, \psi\right). \blacksquare$$

Next, we find the unbiased estimators of B and Σ.

THEOREM 9.1.3. *The statistics* $\hat{B}$ *and* $\hat{\Sigma}_U = \frac{-1}{2\psi'(0)(n-q)} X(I_n - Z'(ZZ')^{-1}Z)X'$ *for model (9.1) are unbiased for* B *and* Σ, *if the second order moment of* X *exists.*

PROOF: From (9.5) it follows that $\mathcal{E}(\hat{B}) = B$. On the other hand, from Theorem 3.2.8, we get

$$\mathcal{E}(\hat{\Sigma}_U) = \frac{-1}{2\psi'(0)(n-q)} \mathcal{E}(X(I_n - Z'(ZZ')^{-1}Z)X')$$

$$= \frac{-1}{2\psi'(0)(n-q)} (-2\psi'(0))\ \Sigma\ \mathrm{tr}(I_n - Z'(ZZ')^{-1}Z)$$
$$+ BZ(I_n - Z'(ZZ')^{-1}Z)Z'B$$

$$= \frac{1}{n-q} \Sigma(n-q)$$

$$= \Sigma. \blacksquare$$

The question of sufficiency is studied in the following theorem. First, we prove a lemma.

LEMMA 9.1.1. *Let* X *be* $p \times n$, B *be* $p \times q$, *and* Z *be* $q \times n$ *matrices with* $\mathrm{rk}(Z) = q$. *Define* $\hat{B} = XZ'(ZZ')^{-1}$, *then*

$$(X - BZ)(X - BZ)' = (X - \hat{B}Z)(X - \hat{B}Z)' + (\hat{B} - B)ZZ'(\hat{B} - B)'. \tag{9.13}$$

PROOF: We can write

$$(X - BZ)(X - BZ)' = (X - \hat{B}Z + (\hat{B} - B)Z)(X - \hat{B}Z + (\hat{B} - B)Z)'$$

$$= (X - \hat{B}Z)(X - \hat{B}Z)' + (\hat{B} - B)ZZ'(\hat{B} - B)'$$
$$+ (X - \hat{B}Z)Z'(\hat{B} - B)' + (\hat{B} - B)Z(X - \hat{B}Z)'. \tag{9.14}$$

However,

$$(X - \hat{B}Z)Z' = (X - XZ'(ZZ')^{-1}Z)Z' = XZ' - XZ' = O$$

so the last two terms in (9.14) vanish. Thus, from (9.14), we get (9.13). $\blacksquare$

THEOREM 9.1.4. *In model (9.1),* $\hat{B}$ *and* $\hat{\Sigma}$ *are sufficient for* B *and* Σ.

PROOF: From (9.2) and Lemma 9.1.1, we get

$$f(X) = \frac{1}{|\Sigma|^n} h(\mathrm{tr}(\Sigma^{-1}(X - BZ)(X - BZ)'))$$
$$= \frac{1}{|\Sigma|^n} h(\mathrm{tr}(\Sigma^{-1}[(X - \hat{B}Z)(X - \hat{B}Z)' + (\hat{B} - B)ZZ'(\hat{B} - B)'])) .$$

Now,

$$(X - \hat{B}Z)(X - \hat{B}Z)' = (X - XZ'(ZZ')^{-1}Z)(X - XZ'(ZZ')^{-1}Z)'$$
$$= X(I_n - Z'(ZZ')^{-1}Z)(I_n - Z'(ZZ')^{-1}Z)' X'$$
$$= X(I_n - Z'(ZZ')^{-1}Z)X'$$
$$= \frac{z_h}{p} \hat{\Sigma}.$$

Hence,

$$f(X) = \frac{1}{|\Sigma|^n} h\left(\mathrm{tr}\left(\Sigma^{-1}\left[\frac{z_h}{p} \hat{\Sigma} + (\hat{B} - B)ZZ'(\hat{B} - B)'\right]\right)\right),$$

which proves the theorem. ■

Next, we focus on some optimality properties of the MLE's for Model (9.1). Here, we assume that X has a finite second order moment. Let the $p \times 1$ vector $\mathbf{a}$ and the $q \times 1$ vector $\mathbf{c}$ be given and assume that we want to estimate $\xi = \mathbf{a}'B\mathbf{c}$. We are interested in a linear unbiased estimator of ξ, that is an estimator which can be written in the form $\xi^* = \mathbf{v}' \operatorname{vec}(X')$, where $\mathbf{v}$ is a pn-dimensional vector and $\mathcal{E}(\xi^*) = \xi$.

The MLE of ξ is $\hat{\xi} = \mathbf{a}'\hat{B}\mathbf{c}$. This is a linear estimator of ξ since

$$\mathbf{a}'\hat{B}\mathbf{c} = \mathbf{a}'XZ'(ZZ')^{-1}\mathbf{c}$$
$$= \operatorname{vec}(\mathbf{a}'XZ'(ZZ')^{-1}\mathbf{c})$$

$$= (\mathbf{a}' \otimes (\mathbf{c}'(ZZ')^{-1}Z))\mathrm{vec}(X').$$

Moreover, $\hat{\xi}$ is unbiased for ξ since

$$\mathcal{E}(\hat{\xi}) = \mathbf{a}'\mathcal{E}(\hat{B})\mathbf{c} = \mathbf{a}'B\mathbf{c} = \xi.$$

The next result, also called multivariate Gauss-Markov theorem, shows that $\mathrm{Var}(\hat{\xi}) \leq \mathrm{Var}(\xi^*)$ for any linear unbiased estimator ξ^*.

THEOREM 9.1.5. (Gauss-Markov Theorem) *Let* X *be a* $p \times n$ *random matrix, with* $\mathcal{E}(X) = BZ$ *and* $\mathrm{Cov}(X) = \Sigma \otimes I_n$, *where* $B(p \times q)$, *and* $\Sigma(p \times p)$, *are unknown matrices and* $Z(q \times n)$ *is a known matrix. Assume* $\mathrm{rk}(Z) = q$ *and* $\mathbf{a}(p \times 1)$ *and* $\mathbf{c}(q \times 1)$ *are known vectors. Let* $\xi = \mathbf{a}'B\mathbf{c}$, *and define* $\hat{\xi} = \mathbf{a}'\hat{B}\mathbf{c}$. *If* $\xi^* = \mathbf{v}'\ \mathrm{vec}(X')$ *is any linear unbiased estimator of* ξ, *then* $\mathrm{Var}(\hat{\xi}) \leq \mathrm{Var}(\xi^*)$.

PROOF: See Timm (1975, p. 187). ■

Theorem 9.1.5 does not require that X have a m.e.c. distribution. Now, we show that if the distribution of X is elliptically contoured, we can get a stronger result. In order to do this, we need the following lemma due to Stoyan (1983).

LEMMA 9.1.2. *Let* x *and* y *be two one-dimensional random variables. Then,*

$$\mathcal{E}(f(x)) \leq \mathcal{E}(f(y))$$

holds for all increasing, real function f *if and only if*

$$P(x \leq a) \geq P(y \leq a),$$

for all $a \in \mathbb{R}$.

PROOF: See Stoyan (1983, p. 5). ■

Now, we can derive the result on m.e.c. distributions.

THEOREM 9.1.6. *Assume model (9.1) holds and* X *has a finite second order moment. Let* $\mathbf{a}(p \times 1)$, *and* $\mathbf{c}(q \times 1)$ *be known vectors,* $\xi = \mathbf{a}'\mathbf{B}\mathbf{c}$, *and define* $\hat{\xi} = \mathbf{a}'\hat{\mathbf{B}}\mathbf{c}$. *Assume* $\xi^* = \mathbf{v}' \operatorname{vec}(X')$ *is a linear unbiased estimator of* ξ. *Let* $\ell(z)$ *be a loss function, where* $\ell[0,\infty) \to [0,\infty)$, $\ell(0) = 0$, *and* $\ell(z)$ *is increasing on* $[0,\infty)$. *Then,*

$$\mathcal{E}(\ell(|\hat{\xi} - \xi|)) \le \mathcal{E}(\ell(|\xi^* - \xi|)).$$

That is, $\hat{\xi}$ *is optimal in the class of linear unbiased estimators for the loss function* ℓ.

PROOF: Since $\xi^* = \mathbf{v}'\operatorname{vec}(X')$ and it is unbiased for ξ, that is, $\mathcal{E}(\xi^*) = \xi$, we have $\hat{\xi} \sim E_1(\xi,\sigma^2_{\hat{\xi}},\psi)$. We also have $\xi^* \sim E_1(\xi,\sigma^2_{\xi^*},\psi)$.

Now, from Theorem 9.1.5, it follows that

$$\operatorname{Var}(\hat{\xi}) \le \operatorname{Var}(\xi^*).$$

However, from Theorem 2.4.1 we get

$$\operatorname{Var}(\hat{\xi}) = -2\psi'(0)\sigma^2_{\hat{\xi}} \text{ and } \operatorname{Var}(\xi^*) = -2\psi'(0)\sigma^2_{\xi^*}.$$

Thus,

$$\sigma^2_{\hat{\xi}} \le \sigma^2_{\xi^*}. \tag{9.15}$$

Define a random variable z as $z \sim E_1(0,1,\psi)$. Then, we have $\hat{\xi} - \xi \approx \sigma_{\hat{\xi}} z$ and $\xi^* - \xi \approx \sigma_{\xi^*} z$. Consequently, for every real positive a, we obtain

$$P(|\hat{\xi} - \xi| \le a) = P(\sigma_{\hat{\xi}}|z| \le a)$$

$$= P\left(|z| \le \frac{a}{\sigma_{\hat{\xi}}}\right) \tag{9.16}$$

and

$$P(|\xi^* - \xi| \le a) = P(\sigma_{\xi^*}|z| \le a)$$

$$= P\left(|z| \le \frac{a}{\sigma_{\xi^*}}\right). \tag{9.17}$$

From (9.15), (9.16), and (9.17), it follows that

$$P(|\hat{\xi} - \xi| \le a) \ge P(|\xi^* - \xi| \le a).$$

Then, from Lemma 9.1.2, we obtain

$$\mathcal{E}(\ell(|\hat{\xi} - \xi|)) \le \mathcal{E}(\ell(|\xi^* - \xi|)). \blacksquare$$

Next, we prove an optimality property of $\hat{\Sigma}$. First, we need some concepts and results from the theory of majorization. They are taken from Marshall and Olkin (1979). Assume x is an n-dimensional vector $x' = (x_1, x_2, \ldots, x_n)$. Then $x_{[1]} \ge x_{[2]} \ge \ldots \ge x_{[n]}$ denote the components of x in decreasing order.

Now, let x and y be two n-dimensional vectors. Then, we say that y *majorizes* x, and denote this by $x \prec y$, if $\sum_{i=1}^{j} x_{[i]} \le \sum_{i=1}^{j} y_{[i]}$, $j = 1, \ldots, n-1$ and $\sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]}$.

Let ϕ be a real function, defined on $\mathbb{R}^n$. Then, we say that ϕ is *Schur-convex*, if from $x \prec y$, $x, y \in \mathbb{R}^n$, it follows that $\phi(x) \le \phi(y)$.

LEMMA 9.1.3. *Let* $x_1, x_2, \ldots, x_n$ *be exchangeable, random variables and define*

$\mathbf{x} = (x_1,x_2,...,x_n)'$. *Assume* λ *is a real function defined on* $\mathbb{R}^n \times \mathbb{R}^n$ *and it satisfies the following conditions:*

i) $\lambda(\mathbf{z},\mathbf{a})$ *is convex in* $\mathbf{a} \in \mathbb{R}^n$, *if* $\mathbf{z} \in \mathbb{R}^n$ *is fixed,*

ii) $\lambda((z_{\pi(1)},z_{\pi(2)},...,z_{\pi(n)}),(a_{\pi(1)},a_{\pi(2)},...,a_{\pi(n)})) = \lambda(\mathbf{z},\mathbf{a})$ *for all permutations* π *of the first* n *positive integers, and*

iii) $\lambda(\mathbf{z},\mathbf{a})$ *is Borel measurable in* $\mathbf{z}$, *if* $\mathbf{a}$ *is fixed.*

Then, $\phi(\mathbf{a}) = \mathcal{E}(\lambda(\mathbf{x},\mathbf{a}))$ *is symmetric and convex in* $\mathbf{a}$.

PROOF: See Marshall and Olkin (1979, pp. 286-287). ■

LEMMA 9.1.4. *Let* ϕ *be a real function defined on* $\mathbb{R}^n$. *If* $\phi(\mathbf{a})$ *is symmetric and convex in* $\mathbf{a} \in \mathbb{R}^n$, *then* $\phi(\mathbf{a})$ *is Schur-convex.*

PROOF: See Marshall and Olkin (1979, pp. 67-68). ■

Lemmas 9.1.3 and 9.1.4, together with the definition of a Schur-convex function, imply the following theorem.

THEOREM 9.1.7. *Let* $x_1,x_2,...,x_n$ *be exchangeable random variables and define* $\mathbf{x} = (x_1,x_2,...,x_n)'$. *Let* $\mathbf{a}_1$ *and* $\mathbf{a}_2$ *be n-dimensional vectors such that* $\mathbf{a}_1 \prec \mathbf{a}_2$. *Assume* λ *is a real function defined on* $\mathbb{R}^n \times \mathbb{R}^n$ *and it satisfies the conditions (i), (ii) and (iii) of Lemma 9.1.3. Then* $\mathcal{E}(\lambda(\mathbf{x},\mathbf{a}_1)) \leq \mathcal{E}(\lambda(\mathbf{x},\mathbf{a}_2))$.

PROOF: From Lemma 9.1.3, it follows that $\phi(\mathbf{a}) = \mathcal{E}(\lambda(\mathbf{x},\mathbf{a}))$ is symmetric and convex in $\mathbf{a}$. Then, from Lemma 9.1.4 we get that $\phi(\mathbf{a})$ is Schur-convex, and this means, by definition, that $\mathcal{E}(\lambda(\mathbf{x},\mathbf{a}_1)) \leq \mathcal{E}(\lambda(\mathbf{x},\mathbf{a}_2))$ if $\mathbf{a}_1 \prec \mathbf{a}_2$. ■

The following theorem proves another result about an estimator of Σ.

THEOREM 9.1.8. *Assume model (9.1) holds,* X *has a finite second order moment, and* ψ *is known. Let* $\hat{\Sigma}_U = \frac{-1}{2\psi'(0)(n-q)} X(I_n - Z'(ZZ')^{-1}Z)X'$. *Assume* Σ^* *is an unbiased estimator of* Σ *that has the form* $\Sigma^* = XCX'$, *where* C *is a positive semidefinite* $n \times n$ *matrix depending on* X. *Let* $\ell(z)$ *be a loss function, where* ℓ: $[0,\infty) \to [0,\infty)$, $\ell(0) = 0$, *and* $\ell(z)$ *is convex on* $[0,\infty)$. *Then,*

$\mathcal{E}(\ell(\mathrm{tr}(\hat{\Sigma}_U\Sigma^{-1})) \leq \mathcal{E}(\ell(\Sigma^*\Sigma^{-1}))$.

PROOF: Let Σ^* be an unbiased estimator of Σ, which can be written as $\Sigma^* = XCX'$. From Theorem 3.2.8, we get

$$\mathcal{E}(\Sigma^*) = \mathcal{E}(XCX') = -2\psi'(0)\Sigma\ \mathrm{tr}(C') + BZCZ'B'.$$

So, in order for Σ^* to be an unbiased estimator of Σ, we must have $\mathrm{tr}(C) = \dfrac{-1}{2\psi'(0)}$ and $ZCZ' = O$. Since C is positive semidefinite, we can write $C = HDH'$, where H is orthogonal, and D is diagonal with nonnegative elements. Therefore, $ZCZ' = O$ implies $ZHDH'Z' = O$, which can be rewritten as $(ZHD^{\frac{1}{2}})(ZHD^{\frac{1}{2}})' = O$. Hence,

$$ZHD^{\frac{1}{2}} = O,$$

$$ZC = ZHD^{\frac{1}{2}}D^{\frac{1}{2}}H' = O,$$

$$\Sigma^* = XCX' = (X - BZ)\,C(X - BZ)',$$

and

$$\begin{aligned}\mathrm{tr}(\Sigma^*\Sigma^{-1}) &= \mathrm{tr}(\Sigma^{-\frac{1}{2}}\Sigma^*\Sigma^{-\frac{1}{2}})\\ &= \mathrm{tr}(\Sigma^{-\frac{1}{2}}(X - BZ)HDH'(X - BZ)'\Sigma^{-\frac{1}{2}}).\end{aligned}$$

Now define $L = \Sigma^{-\frac{1}{2}}(X - BZ)H$, where $L \sim E_{p,n}(O, I_p \otimes I_n, \psi)$, and let $L \approx rU$ be the stochastic representation of L. Then,

$$\begin{aligned}\mathrm{tr}(\Sigma^*\Sigma^{-1}) &= \mathrm{tr}((rU)D(rU)')\\ &= r^2\mathrm{tr}(UDU').\end{aligned}$$

Since $ZC = O$ and $\mathrm{rk}(Z) = q$, q diagonal elements of D are zero, and since

$\mathrm{tr}(C) = \dfrac{-1}{2\psi'(0)}$, the sum of the others is $\dfrac{-1}{2\psi'(0)}$. Define $d_i = d_{ii}$, $i = 1,\dots,n$, and $\mathbf{d} = (d_1, d_2, \dots, d_n)'$. Let the vector $\mathbf{d}$, corresponding to $\hat{\Sigma}_U$ be denoted by $\hat{\mathbf{d}}$. Since $(I_n - Z'(ZZ')^{-1}Z)$ is an idempotent matrix with rank $n - q$, we see that $n - q$ elements of $\hat{\mathbf{d}}$ are equal to $\dfrac{-1}{2\psi'(0)(n - q)}$, and the rest are zeros. Since, $d_{[1]} \geq d_{[2]} \geq \dots \geq d_{[n-q]}$, we get $\dfrac{1}{n - q} \sum_{i=1}^{n-q} d_{[i]} \leq \dfrac{1}{j} \sum_{i=1}^{j} d_{[i]}$, $j = 1,\dots,n-q-1$, hence

$$\sum_{i=1}^{j} \hat{d}_{[i]} = \left(\frac{j}{n - q}\right)\left(\frac{-1}{2\psi'(0)}\right)$$

$$= \frac{j}{n - q} \sum_{i=1}^{n-q} d_{[i]}$$

$$\leq \sum_{i=1}^{j} d_{[i]}, \quad j = 1,\dots,n - q - 1$$

and $\sum_{i=1}^{j} \hat{d}_{[i]} = \dfrac{-1}{2\psi'(0)} = \sum_{i=1}^{j} d_{[i]}$, $j = n-q,\dots,n$. Therefore,

$$\hat{\mathbf{d}} \prec \mathbf{d}.$$

Let $\mathbf{u}_i$, $i = 1,\dots,n$ denote the i^{th} column of the matrix U. Then,

$$r^2\mathrm{tr}(UDU') = r^2\mathrm{tr}\left(\sum_{i=1}^{n} d_i \mathbf{u}_i \mathbf{u}_i'\right)$$

$$= \sum_{i=1}^{n} d_i r^2 \mathbf{u}_i' \mathbf{u}_i$$

$$= \sum_{i=1}^{n} d_i w_i$$

$$= \mathbf{w}'\mathbf{d},$$

where $w_i = r^2\mathbf{u}_i'\mathbf{u}_i$, $i = 1,\dots,n$, and $\mathbf{w} = (w_1,w_2,\dots,w_n)'$. Consequently, $w_i \geq 0$ and $w_1,w_2,\dots,w_n$ are exchangeable random variables.

Define the real function λ on $\mathbb{R}^n \times \mathbb{R}^n$ as follows: $\lambda(\mathbf{z},\mathbf{t}) = \ell(\mathbf{z}'\mathbf{t})$, where $\mathbf{z},\mathbf{t} \in \mathbb{R}^n$. Then, λ satisfies the conditions of Theorem 9.1.7, $\hat{\mathbf{d}} \prec \mathbf{d}$ and $\mathbf{w}$ has exchangeable components. Thus, we get

$$\mathcal{E}(\ell(\mathbf{w}'\hat{\mathbf{d}})) \leq \mathcal{E}(\ell(\mathbf{w}'\mathbf{d})).$$

However, $\mathbf{w}'\mathbf{d} = r^2\mathrm{tr}(UDU') = \mathrm{tr}(\Sigma^*\Sigma^{-1})$, so we get

$$\mathcal{E}(\ell(\mathrm{tr}(\hat{\Sigma}_U\Sigma^{-1}))) \leq \mathcal{E}(\ell(\mathrm{tr}(\Sigma^*\Sigma^{-1}))). \blacksquare$$

COROLLARY **9.1.8.1.** *Under the conditions of Theorem 9.1.8,*

$$\mathcal{E}(\mathrm{tr}(\hat{\Sigma}_U\Sigma^{-1}))^2 \leq \mathcal{E}(\mathrm{tr}(\Sigma^*\Sigma^{-1}))^2.$$

Theorem 9.1.6 and Theorem 9.1.8 were derived by Kuritsyn (1986) for vector variate elliptically contoured distribution. The results here are the extensions of the results of that paper to the case of matrix variate elliptically contoured distribution.

9.2. HYPOTHESIS TESTING IN THE MULTIVARIATE LINEAR REGRESSION MODEL

In this section, once again, we focus on model (9.1). We use the notations of Section 9.1. The results are taken from Hsu (1985b).

Let the matrix B be partitioned as $B = (B_1,B_2)$, where B_1 is $p \times q_1$

$(1 \le q_1 < q)$, and partition Z, as $Z = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$, where Z_1 is $q_1 \times n$. Let $q_2 = q - q_1$. Define $A = ZZ'$ and partition A as $\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, where A_{11} is $q_1 \times q_1$. Then, $A_{ij} = Z_i Z_j'$, $i = 1,2$, $j = 1,2$. Also, define $A_{11 \cdot 2} = A_{11} - A_{12}(A_{22})^{-1}A_{21}$. We want to test the hypothesis

$$H_0\text{: } B_1 = B_1^* \text{ against } H_1\text{: } B_1 \neq B_1^*, \tag{9.18}$$

where B and Σ are unknown and B_1^* is a $p \times q_1$ given matrix. Note that problem (9.18) is equivalent to testing

$$H_0\text{: } B_1 = O \text{ against } B_1 \neq O. \tag{9.19}$$

Indeed, if $B_1^* \neq O$, then define $X^* = X - B_1^* Z_1$. Then, we get

$$X^* \sim E_{p,n}(BZ - B_1^* Z_1, \Sigma \otimes I_n, \psi)$$

$$= E_{p,n}((B_1 - B_1^*, B_2)Z, \Sigma \otimes I_n, \psi)$$

and $B_1 = B_1^*$ is equivalent to $B_1 - B_1^* = O$.

Problem (9.19) remains invariant under the group G, where G is generated by the linear transformations

i) $g(X) = CX$, where C is $p \times p$ nonsingular matrix and

ii) $g(X) = X + LZ_2$, where L is $p \times q_2$ matrix.

Now, we derive the likelihood ratio test for the problem (9.18).

THEOREM 9.2.1. *The LRT statistic for problem (9.18) is*

$$\tau = \frac{|X(I_n - Z'(ZZ')^{-1}Z)X'|}{|(X - B_1^*Z_1)(I_n - Z_2'(Z_2Z_2')^{-1}Z_2)(X-B_1^*Z_1)'|}.$$

The critical region at level α *is*

$$\tau \leq U_{p,q_1,n-q}(\alpha),$$

where $U_{p,q_1,n-q}(\alpha)$ *denotes the 100α% point of the* $U_{p,q_1,n-q}$ *distribution. If* H_0 *holds, then* $\tau \sim U_{p,q_1,n-q}$. *Moreover, if* $B_1^* = O$ *then* τ *is an invariant of the sufficient statistic under* G.

PROOF: From Theorem 8.1.1, it follows that the LRT statistic is the same as in the normal case. Since problems (9.18) and (9.19) are equivalent, we can focus on (9.19). The statistic τ is invariant under G. Moreover, G satisfies the conditions of part (a) of Theorem 8.1.3. Therefore, the null distribution of τ is the same as in the normal case. Thus, the corresponding results of the normal theory can be used here (see Anderson, 1984, Section 8.4.1). ■

For more on the null distribution of the test statistic and the asymptotic null distribution, see Anderson (1984, Sections 8.4-8.5), Pillai and Gupta (1969), Gupta (1971), Gupta and Tang (1984, 1988), and Tang and Gupta (1984, 1986, 1987).

Next, we focus on the invariance properties of problem (9.18). The results are based on Anderson (1984). Define $Z_1^* = Z_1 - Z_1Z_2'(Z_2Z_2')^{-1}Z_2$ and $B_2^* = B_2 + B_1Z_1Z_2'(Z_2Z_2')^{-1}$. Then, $BZ = B_1Z_1 + B_2Z_2 = B_1Z_1^* + B_2^*Z_2$,

$$\begin{aligned} Z_1^*Z_1^{*\prime} &= (Z_1 - Z_1Z_2'(Z_2Z_2')^{-1}Z_2)(Z_1' - Z_2'(Z_2Z_2')^{-1}Z_2Z_1') \\ &= Z_1Z_1' - Z_1Z_2'(Z_2Z_2')^{-1}Z_2Z_1' - Z_1Z_2'(Z_2Z_2')^{-1}Z_2Z_1' \\ &\quad + Z_1Z_2'(Z_2Z_2')^{-1}Z_2Z_2'(Z_2Z_2')^{-1}Z_2Z_1' \end{aligned}$$

$$= Z_1Z_1' - Z_1Z_2'(Z_2Z_2')^{-1}Z_2Z_1',$$

and

$$Z_1^*Z_2' = (Z_1 - Z_1Z_2'(Z_2Z_2')^{-1}Z_2)Z_2'$$

$$= Z_1Z_2' - Z_1Z_2'(Z_2Z_2')^{-1}Z_2Z_2'$$

$$= O.$$

Thus, (9.1) can be written in the following equivalent form:

$$X \sim E_{p,n}(B_1Z_1^* + B_2^*Z_2, \Sigma \otimes I_n, \psi), \tag{9.20}$$

where $Z_1^*Z_1^{*\prime} = A_{11\cdot 2}$ and $Z_1^*Z_2' = O$. We want to test

$$H_0\colon\ B_1 = O \ \text{ against } \ H_1\colon\ B_1 \neq O. \tag{9.21}$$

Problem (9.21) remains invariant under group G, where G is generated by

i) $g(Z_1^*) = KZ_1^*$, where K is $q \times q$ nonsingular matrix,

and by the transformations

ii) $g(X) = CX$, where C is $p \times p$ nonsingular matrix, and (9.22)

iii) $g(X) = X + LZ_2$, where L is $p \times q_2$ matrix. (9.23)

Then, we have the following theorem.

THEOREM 9.2.2. *The maximal invariant of* $A_{11\cdot 2}$, *and the sufficient statistic* $\hat{B}$ *and* $\hat{\Sigma}$ *under* G *is the set of roots of*

$$|H - \ell S| = 0, \tag{9.24}$$

where $H = \hat{B}_1 A_{11\cdot 2} \hat{B}_1'$ *and* $S = \frac{z_h}{p} \hat{\Sigma}$. *Here* $\hat{B}_1$ *denotes the* $p \times q_1$ *matrix in the partitioning of* $\hat{B}$ *into* $\hat{B} = (\hat{B}_1, \hat{B}_2)$. *Moreover, if* H_0 *holds in (9.21), then the distribution of the roots of (9.24) are the same as in the normal case.*

PROOF: In Anderson (1984, Section 8.6.1), it is shown that the roots of (9.24) form a maximal invariant under the given conditions. Since the subgroup of G, which is generated by the transformations (9.22) and (9.23), satisfies the conditions of part (a) of Theorem 8.1.3, the null distribution of the roots of (9.24) is the same as in the normal case. ■

It is easy to see that the LRT statistic, τ, is a function of the roots of (9.24): $\tau = |I + HS^{-1}|$. Other test statistics, which are also functions of the roots of (9.24) are the Lawley-Hotelling's trace criterion: $\mathrm{tr}(HS^{-1})$; the Bartlett-Nanda-Pillai's trace criterion: $\mathrm{tr}\, H(S + H)^{-1}$; and the Roy's largest (smallest) root criterion, that is, the largest (smallest) characteristic root of HS^{-1}. Then, they have the same null distribution as in the normal case. For a further discussion of these test statistics, see Anderson (1984, Section 8.6), Pillai and Gupta (1969), and Gupta (1971). These invariant statistics were also studied by Hsu (1985b) for the case of m.e.c. distributions and it was also shown that the LRT, the Lawley-Hotelling's trace test, the Bartlett-Nanda-Pillai's trace test, and the Roy's largest root test, are all admissible.

REMARK 9.2.1. Since the multivariate analysis of variance (MANOVA) problems can be formulated in terms of the regression model (9.1), and most hypotheses in MANOVA can be expressed as (9.19), the results of Theorem 9.2.1 can be used here. As a consequence, the LRT statistics are the same as those developed in the normal theory and their null distributions and critical regions are also the same as in the normal case. For the treatment of the MANOVA problems in the normal case, see Anderson (1984, Section 8.9).

9.3. INFERENCE IN THE RANDOM EFFECTS MODEL

In Section 9.1, the p-dimensional vectors, $x_1,x_2,\ldots,x_n$ have the property that $x_i \sim E_p(Bz_i,\Sigma,\psi)$, where B is a $p \times q$ matrix and z_i is a q-dimensional vector, $i = 1,2,\ldots,n$. This can also be expressed as $x_i = Bz_i + v_i$, where $v_i \sim E_p(o,\Sigma,\psi)$, $i = 1,2,\ldots,n$. The vectors z_i, $i = 1,2,\ldots,n$ are assumed to be known. On the other hand, the matrix B is unknown, but it is also constant. The random vectors v_i, $i = 1,2,\ldots,n$ are called the error terms. Let $V = (v_1,v_2,\ldots,v_n)$. Then, the model (9.1) can be expressed as

$$X = BZ + V, \text{ where } V \sim E_{p,n}(O,\Sigma \otimes I_n,\psi).$$

We get a different model if we assume that the vectors z_i, $i = 1,2,\ldots,n$ are also random. Define $y_i = \begin{pmatrix} z_i \\ v_i \end{pmatrix}$, and assume that $y_i \sim E_{p+q}\left(\begin{pmatrix} m_i \\ o \end{pmatrix}, \begin{pmatrix} \Sigma_1 & O \\ O & \Sigma_2 \end{pmatrix}, \psi\right)$, where m_i is a q-dimensional known vector, $i = 1,2,\ldots,n$, and Σ_1, $q \times q$, Σ_2, $(p - q) \times (p - q)$, are unknown matrices. In this case, we suppose that B and m_i are known. Moreover, let y_i, $i = 1,2,\ldots,n$ be uncorrelated and assume that their joint distribution is elliptically contoured. Then, this model can be expressed as

$$X = BZ + V, \text{ where } \begin{pmatrix} Z \\ V \end{pmatrix} \sim E_{p+q,n}\left(\begin{pmatrix} M \\ O \end{pmatrix}, \begin{pmatrix} \Sigma_1 & O \\ O & \Sigma_2 \end{pmatrix} \otimes I_n, \psi\right), \qquad (9.25)$$

where $q \leq p$, $X = (x_1,x_2,\ldots,x_n)$ is $p \times n$, Z is $q \times n$, and V is $p \times n$. Assume that the $p \times q$ matrix B and the $q \times n$ matrix M are known, but the $q \times q$ matrix Σ_1 and the $(p - q) \times (p - q)$ matrix Σ_2 are unknown. Also assume that $\mathrm{rk}(B) = q$, and the random matrix $Y = \begin{pmatrix} Z \\ V \end{pmatrix}$ has finite second order moment.

We want to find the optimal mean-square estimator of Z given X. This is equivalent to finding $\mathcal{E}(Z \mid X)$. We need the following result.

LEMMA 9.3.1. *Let* $X \sim E_{p,n}(M,\Sigma \otimes \Phi,\psi)$, *with stochastic representation*

$X \approx rA_0UB_0'$. *Let* F *be the distribution function of* r. *Define* $Y = AXB$, *with* $A(q \times p)$, $B(n \times m)$ *matrices,* $rk(A) = q$, *and* $rk(B) = m$. *Then,*

a) *if* $\mathcal{E}(X \mid Y)$ *exists, we have*

$$\mathcal{E}(X \mid Y) = M + \Sigma A'(A\Sigma A')^{-1}(Y - AMB)(B'\Phi B)^{-1}B'\Phi.$$

b) *if* $m = n$, $K = \Phi_n = I_n$ *and* $Cov(X \mid Y)$ *exists, we have*

$$Cov(X \mid Y) = k(\Sigma - \Sigma A'(A\Sigma A')^{-1}A\Sigma) \otimes I_n,$$

where $k = -2\psi_q'(X_2)$ *and* $\psi_q(X_2)$ *is defined by (2.30), (2.31), and (2.32) with* $q(X_2) = tr((Y - AM)' (A\Sigma A')^{-1} (Y - AM))$. *Moreover, if the distribution of* X *is absolutely continuous and the p.d.f. of* $Y = AX$ *is*

$$f(Y) = \frac{1}{|A\Sigma A'|^{\frac{n}{2}}} h_1(tr(Y - AM)'(A\Sigma A')^{-1}(Y - AM)), \tag{9.26}$$

then

$$k = \frac{\int_r^{\infty} h_1(t)dt}{2h_1(r)},$$

where $r = tr((Y - AM)'(A\Sigma A')^{-1}(Y - AM))$.

PROOF: If $q = p$, and $n = m$, the theorem is obvious. So, assume $qm < pn$.

Step 1. Assume $n = 1$, $\mathbf{x} \sim E_p(\mathbf{o}, I_p, \psi)$, and $B = 1$. Using Theorem 1.3.9, we can write $A = PDQ$, where P is a $q \times q$ nonsingular matrix, Q is a $p \times p$ orthogonal matrix, and D is a $q \times p$ matrix with $D = (I_q, O)$. Then,

$$\mathcal{E}(\mathbf{x} \mid \mathbf{y}) = \mathcal{E}(\mathbf{x} \mid A\mathbf{x})$$

$$= \mathcal{E}(\mathbf{x} \mid PDQ\mathbf{x})$$

$$= Q'\mathcal{E}(Qx \mid PDQx)$$

$$= Q'\mathcal{E}(Qx \mid DQx). \tag{9.27}$$

Let $\mathbf{z} = Q\mathbf{x}$. Then $\mathbf{z} \sim E_p(\mathbf{o}, I_p, \psi)$ and $DQ\mathbf{x} = \mathbf{z}_1$, where $\mathbf{z} = \begin{pmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{pmatrix}$, $\mathbf{z}_1$ is q-dimensional vector. From Theorem 2.6.5, it follows that $\mathcal{E}(\mathbf{z}_2 \mid \mathbf{z}_1) = \mathbf{o}$ and $Cov(\mathbf{z}_2 \mid \mathbf{z}_1) = kI_{p-q}$ where $k = -2\psi'_{q(x_2)}$ and $\psi_{q(x_2)}$ is defined by (2.30), (2.31), and (2.32) with $q(x_2) = \mathbf{z}_1' \mathbf{z}_1$. On the other hand, $\mathcal{E}(\mathbf{z}_1 \mid \mathbf{z}_1) = \mathbf{z}_1$. Therefore,

$$\mathcal{E}(\mathbf{z} \mid \mathbf{z}_1) = D'D\mathbf{z}. \tag{9.28}$$

From (9.27) and (9.28), it follows that

$$\mathcal{E}(\mathbf{x} \mid \mathbf{y}) = Q'D'DQ\mathbf{x}$$

$$= Q'D'P'P'^{-1}P^{-1}PDQ\mathbf{x}$$

$$= A'(PP')^{-1}A\mathbf{x}$$

$$= A'(PDQQ'D'P')^{-1}A\mathbf{x}$$

$$= A'(AA')^{-1}A\mathbf{x}$$

$$= A'(AA')^{-1}\mathbf{y}$$

and

$$Cov(\mathbf{x} \mid \mathbf{y}) = \mathcal{E}(\mathbf{x}\mathbf{x}' \mid \mathbf{y}) - \mathcal{E}(\mathbf{x} \mid \mathbf{y})(\mathcal{E}(\mathbf{x} \mid \mathbf{y}))'$$

$$= \mathcal{E}(Q'\mathbf{z}\mathbf{z}'Q \mid PD\mathbf{z}) - \mathcal{E}(Q'\mathbf{z} \mid PD\mathbf{z})(E(Q'\mathbf{z} \mid PD\mathbf{z}))'$$

$$= Q'[\mathcal{E}(\mathbf{z}\mathbf{z}' \mid D\mathbf{z}) - E(\mathbf{z} \mid D\mathbf{z})(E(\mathbf{z} \mid D\mathbf{z}))']Q. \tag{9.29}$$

Now,

$$\mathcal{E}(\mathbf{z} \mid \mathbf{Dz}) = E(\mathbf{z} \mid \mathbf{z}_1) = \mathbf{D'Dz}. \tag{9.30}$$

$$\mathcal{E}(\mathbf{zz'} \mid \mathbf{Dz}) = \mathcal{E}\left(\begin{pmatrix} \mathbf{z}_1\mathbf{z}_1' & \mathbf{z}_1\mathbf{z}_2' \\ \mathbf{z}_2\mathbf{z}_1' & \mathbf{z}_2\mathbf{z}_2' \end{pmatrix} \mid \mathbf{z}_1\right) = \begin{pmatrix} \mathcal{E}(\mathbf{z}_1\mathbf{z}_1' \mid \mathbf{z}_1) & \mathcal{E}(\mathbf{z}_1\mathbf{z}_2' \mid \mathbf{z}_1) \\ \mathcal{E}(\mathbf{z}_2\mathbf{z}_1' \mid \mathbf{z}_1) & \mathcal{E}(\mathbf{z}_2\mathbf{z}_2' \mid \mathbf{z}_1) \end{pmatrix},$$

$$\mathcal{E}(\mathbf{z}_1\mathbf{z}_1' \mid \mathbf{z}_1) = \mathbf{z}_1\mathbf{z}_1' = \mathbf{Dzz'D'},$$

$$\mathcal{E}(\mathbf{z}_1\mathbf{z}_2' \mid \mathbf{z}_1) = \mathbf{z}_1\mathcal{E}(\mathbf{z}_2' \mid \mathbf{z}_1) = \mathbf{o},$$

$$\mathcal{E}(\mathbf{z}_2\mathbf{z}_1' \mid \mathbf{z}_1) = \mathcal{E}(\mathbf{z}_2 \mid \mathbf{z}_1)\mathbf{z}_1' = \mathbf{o}, \text{ and}$$

$$\mathcal{E}(\mathbf{z}_2\mathbf{z}_2' \mid \mathbf{z}_1) = \mathcal{E}(\mathbf{z}_2\mathbf{z}_2' \mid \mathbf{z}_1) - \mathcal{E}(\mathbf{z}_2 \mid \mathbf{z}_1)\mathcal{E}(\mathbf{z}_2 \mid \mathbf{z}_1)' = \mathrm{Cov}(\mathbf{z}_2,\mathbf{z}_1) = k\mathbf{I}_{p-q}.$$

However,

$$\begin{aligned} \mathbf{z}_1'\mathbf{z}_1 &= \mathbf{x'Q'D'DQx} \\ &= \mathbf{x'A'(AA')^{-1}Ax} \\ &= \mathbf{y'(AA')^{-1}y}. \end{aligned}$$

So, $k = -2\psi'_{q(x_2)}$ where $\psi_{q(x_2)}$ is defined by (2.30), (2.31), and (2.32) with $q(\mathbf{x}_2) = \mathbf{y'(AA')^{-1}y}$. Hence,

$$\mathcal{E}(\mathbf{zz'} \mid \mathbf{Dz}) = \mathbf{D'Dzz'D'D} + k(\mathbf{I}_p - \mathbf{DD'}). \tag{9.31}$$

If the distribution of $\mathbf{x}$ is absolutely continuous and the p.d.f. of $\mathbf{y} = \mathbf{Hx}$ is given by (9.26), then the p.d.f. of $\mathbf{z}_1 = \mathbf{DQx}$ is $h_1(\mathbf{z}_1'\mathbf{z}_1)$. Hence $k = \dfrac{\int_r^\infty h_1(t)dt}{2h_1(r)}$, with $r = \mathbf{z}_1'\mathbf{z}_1 = \mathbf{y'(AA')^{-1}y}$.

It follows from (9.29), (9.30), and (9.31) that

$$\begin{aligned}
\mathrm{Cov}(\mathbf{x} \mid \mathbf{y}) &= Q'(D'D\mathbf{zz}'D'D + k(I_p - DD') - D'D\mathbf{zz}'D'D)Q \\
&= k(I_p - Q'D'DQ) \\
&= k(I_p - A'(AA')^{-1}A).
\end{aligned}$$

Step 2. Let $X \sim E_{p,n}(M, \Sigma \otimes \Phi, \psi)$. We have

$$\begin{aligned}
(\mathrm{vec}(X') \mid \mathrm{vec}(Y')) &= ((\mathrm{vec}(X - M)' + \mathrm{vec}(M')) \mid \mathrm{vec}(AXB)') \\
&= (\mathrm{vec}(X - M)' \mid \mathrm{vec}(A(X - M)B)') + \mathrm{vec}(M') \\
&= (((\Sigma^{\frac{1}{2}} \otimes \Phi^{\frac{1}{2}})\mathrm{vec}(\Sigma^{-\frac{1}{2}}(X - M)\Phi^{-\frac{1}{2}})') \mid (((A\Sigma^{\frac{1}{2}}) \otimes (B'\Phi^{\frac{1}{2}})) \\
&\quad \mathrm{vec}(\Sigma^{\frac{1}{2}}(X - M)\Phi^{-\frac{1}{2}})')) + \mathrm{vec}(M') \\
&= (\Sigma^{\frac{1}{2}} \otimes \Phi^{\frac{1}{2}})((\mathrm{vec}(\Sigma^{-\frac{1}{2}}(X - M)\Phi^{-\frac{1}{2}})') \mid (((A\Sigma^{\frac{1}{2}}) \otimes (B'\Phi^{\frac{1}{2}})) \\
&\quad \mathrm{vec}(\Sigma^{-\frac{1}{2}}(X - M)\Phi^{-\frac{1}{2}})')) + \mathrm{vec}(M'). \qquad (9.32)
\end{aligned}$$

Now, $\mathrm{vec}(\Sigma^{-\frac{1}{2}}(X - M)\Phi^{-\frac{1}{2}})' \sim E_{pn}(\mathbf{o}, I_{pn}, \psi)$ and using Step 1, we get

$$\begin{aligned}
&\mathcal{E}((\mathrm{vec}(\Sigma^{-\frac{1}{2}}(X - M)\Phi^{-\frac{1}{2}})') \mid (((A\Sigma^{\frac{1}{2}}) \otimes (B'\Phi^{\frac{1}{2}}))\mathrm{vec}(\Sigma^{-\frac{1}{2}}(X - M)\Phi^{-\frac{1}{2}})')) \\
&= ((\Sigma^{\frac{1}{2}}A') \otimes (\Phi^{\frac{1}{2}}B))((A\Sigma^{\frac{1}{2}}\Sigma^{\frac{1}{2}}A') \otimes (B'\Phi^{\frac{1}{2}}\Phi^{\frac{1}{2}}B))^{-1}((A\Sigma^{\frac{1}{2}}) \otimes (B'\Phi^{\frac{1}{2}})) \\
&\quad \mathrm{vec}(\Sigma^{-\frac{1}{2}}(X - M)\Phi^{-\frac{1}{2}})' \\
&= (\Sigma^{\frac{1}{2}}A'(A\Sigma A')^{-1}A\Sigma^{\frac{1}{2}}) \otimes (\Phi^{\frac{1}{2}}B(B'\Phi B)^{-1}B'\Phi^{\frac{1}{2}})\mathrm{vec}(\Sigma^{-\frac{1}{2}}(X - M)\Phi^{-\frac{1}{2}})'
\end{aligned}$$

$$= \mathrm{vec}(\Sigma^{\frac{1}{2}}A'(A\Sigma A')^{-1}A(X - M)B(B'\Phi B)^{-1}B'\Phi^{\frac{1}{2}})'. \tag{9.33}$$

From (9.32) and (9.33), we get

$$\mathcal{E}(X \mid Y) = M + \Sigma A'(A\Sigma A')^{-1}A(X - M)B(B'\Phi B)^{-1}B'\Phi$$

$$= M + \Sigma A'(A\Sigma A')^{-1}(Y - AMB)(B'\Phi B)^{-1}B'\Phi.$$

If $m = n$ and $B = \Phi = I_n$, then from Step 1, it follows that

$$\mathrm{Cov}((\mathrm{vec}(\Sigma^{-\frac{1}{2}}(X - M))') \mid (((A\Sigma^{\frac{1}{2}}) \otimes I_n)\mathrm{vec}(\Sigma^{-\frac{1}{2}}(X - M))'))$$

$$= k(I_{pn} - (\Sigma^{\frac{1}{2}}A'(A\Sigma A')^{-1}A\Sigma^{\frac{1}{2}}) \otimes I_n)$$

$$= k(I_p - \Sigma^{\frac{1}{2}}A'(A\Sigma A')^{-1}A\Sigma^{\frac{1}{2}}) \otimes I_n. \tag{9.34}$$

It follows from (9.32) and (9.34), that

$$\mathrm{Cov}(X \mid Y) = k(\Sigma - \Sigma A'(A\Sigma A')^{-1}A\Sigma) \otimes I_n,$$

where $k = -2\psi'_{q(x_2)}$ and $\psi_{q(x_2)}$ is defined by (2.30), (2.31), and (2.32) with $q(x_2) = \mathrm{tr}((Y - AM)' (A\Sigma A')^{-1} (Y - AM))$.

If the distribution of X is absolutely continuous, then $k = \dfrac{\int_r^\infty h_1(t)dt}{2h_1(r)}$, with $r = \mathrm{tr}((Y - AM)'(A\Sigma A')^{-1}(Y - AM))$. ■

Now, we find the optimal mean square estimator of Z given X.

THEOREM 9.3.1. *Assume that in model (9.25),* $\mathrm{rk}(B) = q$ *and the random matrix* $Y = \begin{pmatrix} Z \\ V \end{pmatrix}$ *has a finite second order moment. Let* rA_0UB_0' *be the stochastic* *representation of* Y *and* F *be the distribution function of* r. *Then, the optimal mean-square estimator of* Z *given* X *is*

$$\hat{Z} = \mathcal{E}(Z|X) = M + \Sigma_1 B'(B\Sigma_1 B' + \Sigma_2)^{-1}(X - BM).$$

Furthermore,

$$\mathrm{Cov}(Z|X) = k(\Sigma_1 - \Sigma_1 B'(B\Sigma_1 B' + \Sigma_2)^{-1}B\Sigma_1) \otimes I_n,$$

where $k = -2\psi'_{q(x_2)}$ and $\psi_{q(x_2)}$ is defined by (2.30), (2.31), and (2.32) with $q(x_2) = \mathrm{tr}((X - BM)'(B\Sigma_1 B' + \Sigma_2)^{-1}(X - BM)$. *If* Y *is absolutely continuous and the p.d.f. of* X *is*

$$f(X) = \frac{1}{|B\Sigma_1 B' + \Sigma_2|^{\frac{n}{2}}} h(\mathrm{tr}((X - BM)'(B\Sigma_1 B' + \Sigma_2)^{-1}(X - BM))$$

then

$$k = \frac{\int_r^\infty h(t)dt}{2h(r)},$$

where $r = \mathrm{tr}((X - BM)'(B\Sigma_1 B' + \Sigma_2)^{-1}(X - BM))$.

PROOF: We have $X = (B, I_p)\begin{pmatrix} Z \\ V \end{pmatrix}$. Hence,

$$X \sim E_{p,n}(BM, (B\Sigma_1 B' + \Sigma_2) \otimes I_n, \psi).$$

Using Lemma 9.3.1, we get

$$\mathcal{E}\left(\begin{pmatrix} Z \\ V \end{pmatrix} | X\right) = \begin{pmatrix} M \\ O \end{pmatrix} + \begin{pmatrix} \Sigma_1 & O \\ O & \Sigma_2 \end{pmatrix}\begin{pmatrix} B' \\ I_p \end{pmatrix}\left[(B, I_p)\begin{pmatrix} \Sigma_1 & O \\ O & \Sigma_2 \end{pmatrix}\begin{pmatrix} B' \\ I_p \end{pmatrix}\right]^{-1}$$

$$\cdot \left[(X - (B, I_p)) \begin{pmatrix} M \\ O \end{pmatrix} \right]$$

$$= \begin{pmatrix} M \\ O \end{pmatrix} + \begin{pmatrix} \Sigma_1 B' \\ \Sigma_2 \end{pmatrix} (B\Sigma_1 B' + \Sigma_2)^{-1}(X - BM)$$

$$= \begin{pmatrix} M + \Sigma_1 B'(B\Sigma_1 B' + \Sigma_2)^{-1}(X - BM) \\ \Sigma_2 (B\Sigma_1 B' + \Sigma_2)^{-1}(X - BM) \end{pmatrix}.$$

Thus,

$$\hat{Z} = \mathcal{E}(Z \mid X) = M + \Sigma_1 B'(B\Sigma_1 B' + \Sigma_2)^{-1} (X - BM).$$

We also get

$$\mathrm{Cov}\left(\begin{pmatrix} Z \\ V \end{pmatrix} \mid X \right) = k\left(\begin{pmatrix} \Sigma_1 & O \\ O & \Sigma_2 \end{pmatrix} - \begin{pmatrix} \Sigma_1 B' \\ \Sigma_2 \end{pmatrix} (B\Sigma_1 B' + \Sigma_2)^{-1} (B\Sigma_1 \; \Sigma_2) \right) \otimes I_n.$$

Therefore,

$$\mathrm{Cov}(Z \mid X) = k(\Sigma_1 - \Sigma_1 B'(B\Sigma_1 B' + \Sigma_2)^{-1} B\Sigma_1) \otimes I_n,$$

where $k = -2\psi'_{q(x_2)}$ and $\psi_{q(x_2)}$ is defined by (2.30), (2.31), and (2.32) with $q(x_2) = \mathrm{tr}((X - BM)'(B\Sigma_1 B' + \Sigma_2)^{-1}(X - BM))$.

If $\begin{pmatrix} Z \\ V \end{pmatrix}$ is absolutely continuous, then

$$k = \frac{\int_r^\infty h_1(t)dt}{2h_1(r)}$$

with $r = \mathrm{tr}((X - BM)'(B\Sigma_1 B' + \Sigma_2)^{-1}(X - BM))$. ■

The results of this section were derived by Chu (1973) for the vector variate case.

BIBLIOGRAPHY

Abramowitz, M. and Stegun, I. A. (1965). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover Publications, Inc., New York.

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis,* 2nd Edition. Wiley, New York.

Anderson, T. W. and Fang, K. T. (1982a). Inference in multivariate elliptically contoured distributions based on maximum likelihood. Technical Report 1. Department of Statistics, Stanford University, California. Reprinted in Fang and Anderson (1990).

Anderson, T. W. and Fang, K. T. (1982b). On the theory of multivariate elliptically contoured distributions and their applications. Technical Report 54. Department of Statistics, Stanford University, California. Reprinted in Fang and Anderson (1990).

Anderson, T. W. and Fang, K. T. (1987). Cochran's theorem for elliptically contoured distributions. *Sankhyā* , Ser. A **49**, 305-315.

Anderson, T. W., Fang, K. T., and Hsu, H. (1986). Maximum likelihood estimates and likelihood-ratio criteria for multivariate elliptically contoured distributions. *The Canadian Journal of Statistics* **14**(1), 55-59.

Bentler, P. M. and Berkane, M. (1985). Developments in the elliptical theory generalization of normal multivariate analysis. *American Statistical Association, Proceedings of Social Statistics Section,* 291-295.

Berkane, M. and Bentler, P. M. (1986a). Moments of elliptically distributed random variates. *Statistics and Probability Letters* **4**, 333-335.

Berkane, M. and Bentler, P. M. (1986b). Characterizing parameters of elliptical distributions. *American Statistical Association, Proceedings of Social Statistics Section,* 278-279.

Billingsley, P. (1979). *Probability and Measure.* Wiley, New York.

Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Pschology* **37**, 62-83.

Cacoullos, T. and Koutras, M. (1984). Quadratic forms in spherical random variables: generalized noncentral χ^2 distribution. *Naval Research Logistics Quarterly* **41**, 447-461.

Cacoullos, T. and Koutras, M. (1985). Minimum-distance discrimination for spherical distributions. *Statistical Theory and Data Analysis.* Elsevier Science Publishers B.V., North Holland, 91-102.

Cambanis, S., Huang, S., and Simons, G. (1981). On the theory of elliptically contoured distributions. *Journal of Multivariate Analysis* **11**, 368-385.

Cellier, D., Fourdrinier, D., and Robert, C. (1989). Robust shrinkage estimators of the location parameter for elliptically symmetric distributions. *Journal of Multivariate Analysis* **29**, 39-52.

Chmielewski, M. A. (1980). Invariant scale matrix hypothesis tests under elliptical symmetry. *Journal of Multivariate Analysis* **10**, 343-350.

Chmielewski, M. A. (1981). Elliptically symmetric distributions: a review and bibliography. *International Statistical Review* **49**, 67-74.

Chu, K. C. (1973). Estimation and decision for linear systems with elliptical random processes. *IEEE Transactions on Automatic Control* **18**, 499-505.

Cléroux, R. and Ducharme, G. R. (1989). Vector correlation for elliptical distributions. *Communications in Statistics—Theory and Methods* **18(4)**, 1441-1454.

Dawid, A. P. (1977). Spherical matrix distributions and a multivariate model. *Journal of the Royal Statistical Society Ser. B* **39**, 254-261.

Dickey, J. M. (1967). Matricvariate generalizations of the multivariate t distribution and the inverted multivariate t distribution. *Annals of Mathematical Statistics* **38**, 511-518.

Eaton, M. L. (1972). *Multivariate Statistical Analysis.* Institut for Mathematisk Statistik, Kobenhavns Universitet, Kobenhavn.

Fan, J. and Fang, K. T. (1985a). Minimax estimators and Stein's two-stage estimators of location parameters. *Chinese Journal of Applied Probability and Statistics* **2**, 103-114. Reprinted in Fang and Anderson (1990).

Fan, J. and Fang, K. T. (1985b). Inadmissibility of sample mean and sample regression coefficients for elliptically contoured distributions. *Northeastern Mathematical Journal* **1**, 68-81. Reprinted in Fang and Anderson (1990).

Fang, K. T. and Anderson, T. W. (1990). *Statistical Inference in Elliptically Contoured and Related Distributions.* Allerton Press Inc., New York.

Fang, K. T. and Chen, H. F. (1984). Relationships among classes of spherical matrix distributions. *Acta Mathematicae Sinica* (English Series) **1**(2), 138-148. Reprinted in Fang and Anderson (1990).

Fang, K. T., Kotz, S., and Ng, K. W. (1990). *Symmetric Multivariate and Related Distributions.* Chapman and Hall, London, New York.

Fang, K. T. and Wu, Y. (1984). Distribution of quadratic forms and Cochran's theorem. *Mathematics in Economics* **1**, 29-48.

Fang, K. T. and Zhang, Y. T. (1990). *Generalized Multivariate Analysis.* Springer-Verlag, New York.

Feller, W. (1957). *An Introduction to Probabilit Theoryy and Its Applications,* vol. I, 2nd Edition. Wiley, New York.

Graybill, F. A. (1969). *Introduction to Matrices with Applications to Statistics.* Wadsworth Publishing Company, Belmont, California.

Grübel, R. and Rocke, D. M. (1989). On the cumulants of affine equivariant estimators in elliptical families. Department of Mathematics, Imperial College, London. Unpublished.

Gupta, A. K. (1971). Noncentral distribution of Wilks' statistic in MANOVA. *Annals of Mathematical Statistics,* **42**, 1254-1261.

Gupta, A. K. (1975). On a stochastic inequality for the Wilks' statistic. *Annals of the Institute of Statistical Mathematics,* **27**, 341-348.

Gupta, A. K. (1977). On the sphericity test criterion in the multivariate Gaussian distribution. *Australian Journal of Statistics,* **19**, 202-205.

Gupta, A. K., Chattopadhyay, A. K., and Krishnaiah, P. R. (1975). Asymptotic distributions of the determinants of some random matrices. *Communications in Statistics,* **4**, 33-47.

Gupta, A. K. and Javier, W. R. (1986). Nonnull distribution of the determinant of B-statistic in multivariate analysis. *South African Journal of Statistics,* **20**, 87-102.

Gupta, A. K. and Nagar, D. K. (1987). Likelihood ratio test for multisample sphericity. In "*Advances in Multivariate Statistical Analysis*" (A. K. Gupta, editor), Reidel Publishing Company, Dordretch, 111-139.

Gupta, A. K. and Nagar, D. K. (1988). Asymptotic expansion of the nonnull distribution of likelihood ratio statistic for testing multisample sphericity. *Communications in Statistics—Theory and Methods*, **17**, 3145-3156.

Gupta, A. K. and Tang, J. (1984). Distribution of likelihood ratio statistic for tetsing equality of covariance matrices of multivariate Gaussian models. *Biometrika*, **71**, 555-559.

Gupta, A. K. and Tang, J. (1988). On a general distribution theory for a class of likelihood ratio criteria. *Australian Journal of Statistics*, **30**, 359-366.

Gupta, A. K. and Varga T. (1990). Characterization of joint density by conditional densities. *Communications In Statistics—Theory and Methods*, **19**, 4643-4652.

Gupta, A. K. and Varga, T. (1991a). Characterization of matrix variate normality through conditional distributions. Department of Mathematics and Statistics, Bowling Green State University, Technical Report No. 91-04.

Gupta, A. K. and Varga, T. (1991b). Rank of a quadratic form in an elliptically contoured matrix random variable. *Statistics and Probability Letters*, **12**, 131-134.

Gupta, A. K. and Varga, T. (1991c). A new class of matrix variate elliptically contoured distributions. Department of Mathematics and Statistics, Bowling Green State University, Technical Report No. 91-05.

Gupta, A. K. and Varga, T. (1992a). Characterization of matrix variate normal distributions. *Journal of Multivariate Analysis*, **41**, 80-88..

Gupta, A. K. and Varga, T. (1992b). Moment and other expected values for matrix variate elliptically contoured distributions. Department of Mathematics and Statistics, Bowling Green State University, Technical Report No. 92-02.

Gupta, A. K. and Varga, T. (1992c). Normal mixture representation of matrix variate elliptically contoured distributions. Department of Mathematics and Statistics, Bowling Green State University, Technical Report No. 92-03.

Gupta, A. K. and Varga, T. (1992d). Some inference problems for matrix variate elliptically contoured distributions. Department of Mathematics and Statistics, Bowling Green State University, Technical Report No. 92-05.

Gupta, A. K. and Varga, T. (1992e). Characterization of matrix variate elliptically contoured distributions. Department of Mathematics and Statistics, Bowling Green State University, Technical Report No. 92-06.

Gupta, A. K. and Varga, T. (1992f). Some applications of the stochastic representation of elliptically contoured distribution. Department of Mathematics and Statistics, Bowling Green State University, Technical Report No. 92-10.

Hayakawa, T. (1987). Normalizing and variance stabilizing transformations of multivariate statistics under an elliptical population. *Annals of the Institute of Statistical Mathematics* **39A**, 299-306.

Hocking, R. R. (1985). *The Analysis of Linear Models.* Brooks and Cole Publishing Company, Belmont, California.

Hsu, H. (1985a). Generalized T^2-test for multivariate elliptically contoured distributions. Technical Report 14. Department of Statistics, Stanford University, California. Reprinted in Fang and Anderson (1990).

Hsu, H. (1985b). Invariant tests for multivariate elliptically contoured distributions. Technical Report 14. Department of Statistics, Stanford University, California. Reprinted in Fang and Anderson (1990).

Jajuga, K. (1987). Elliptically symmetric distributions and their application to classification and regression. *Proceedings of the Second International Tampere Conference in Statistics.* Department of Mathematical Sciences, University of Tampere, 491-498.

Javier, W. R. and Gupta, A. K. (1985a). On matrix variate t-distribution. *Communications in Statistics—Theory and Methods* **14**(6), 1413-1425.

Javier, W. R. and Gupta, A. K. (1985b). On generalized matrix variate beta distributions. *Statistics* **16**, 549-558.

Johnson, M. E. (1987). *Multivariate Statistical Simulation.* Wiley, New York.

Johnson, N. L. and Kotz, S. (1972). *Distributions in Statistics: Continuous Multivariate Distributions.* Wiley, New York.

Kariya, T. (1981). A robustness property of Hotelling's T^2-test. *The Annals of Statistics* **9**(1), 211-214.

Karlin, S. (1956). Decision theory for Polya-type distributions. Case of two actions, I. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability,* **1**, 115-129. University of California Press, Berkeley.

Kelker, D. (1970). Distribution theory of spherical distributions and a location-scale parameter generalization. *Sankhyā, Ser. A* **32**, 419-430.

Khatri, C. G. (1980). A note on an inequality for elliptically contoured distributions. *Gujarat Statistical Review* **7**, 17-21.

Khatri, C. G. (1987). Quadratic forms and null robustness for elliptical distributions. *Proceedings of the Second International Tampere Conference in Statistics.* Department of Mathematical Sciences, University of Tampere, 177-203.

Khatri, C. G. (1988). Some asymptotic inferential problems connected with elliptical distributions. *Journal of Multivariate Analysis* **27**, 319-333.

Khatri, C. G. and Mukerjee, R. (1987). Characterization of normality within the class of elliptical contoured distributions. *Statistics and Probability Letters* **5**, 187-190.

Khattree, R. and Peddada, S. D. (1987). A short note on Pitman nearness for elliptically symmetric estimators. *Journal of Statistical Planning and Inference* **16**, 257-260.

Kingman, J. F. C. (1972). On random sequences with spherical symmetry. *Biometrika* **59**, 494-494.

Konishi, S. and Gupta, A. K. (1989). Testing the equality of several intraclass correation coefficients. *Journal of Statistical Planning and Inference,* **19**, 93-105.

Krishnaiah, P. R. and Lin, J. (1986). Complex elliptically symmetric distributions. *Communications in Statistics—Theory and Methods* **15**(12), 3693-3718.

Kuritsyn, Y. G. (1986). On the least-squares method for elliptically contoured distributions. *Theory of Probability and its Applications* **31**, 738-740.

Lehmann, E. L. (1959). *Testing Statistical Hypothesis.* Wiley, New York.

Lukacs, E. (1956). Characterizations of populations by properties of suitable statistics. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability,* **2**, University of California Press, Los Angeles and Berkeley.

Magnus, J. R. and Neudecker, H. (1988). *Matrix Differential Calculus with Applications in Statistics and Econometrics.* Wiley, New York.

Marcinkiewicz, J. (1938). Sur les fonctions indépendantes III. *Fundamenta Mathematicae* **31**, 86-102.

Marshall, A. and Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications,* Academic Press, New York.

Mitchell, A. F. S. (1989). The information matrix, skewness tensor and L-connections for the general multivariate elliptic distributions. *Annals of the Institute of Statistical Mathematics* **41**, 289-304.

Mitchell, A. F. S. and Krzanowski, W. J. (1985). The Mahalanobis distance and elliptic distributions. *Biometrika* **72**(2), 464-467.

Mood, A. M., Graybill, F. A., and Boes, D. C. (1974). *Introduction to the Theory of Statistics.* McGraw-Hill, New York.

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

Nagao, H. (1973). On some test criteria for covariance matrix. *Annals of Statistics,* **1**, 700-709.

Nagar, D. K. and Gupta, A. K. (1986). Nonnull distribution of LRC for testing H_0: $\mu = 0, \Sigma = \sigma^2 i$ in multivariate normal distribution. *Statistica,* **46**, 291-296.

Nel, H. M. (1977). On distributions and moments associated with matrix normal distributions. Technical Report 24, Department of Mathematical Statistics, University of the Orange Free State.

Okamoto, M. (1973). Distinctness of the eigenvalues of a quadratic form in a multivariate sample. *The Annals of Statistics* **1**, 763-765.

Olkin, I. and Rubin, H. (1964). Multivariate beta distributions and independence properties of the Wishart distribution. *Annals of Mathematical Statistics* **35**, 261-269.

Pillai, K. C. S. and Gupta, A. K. (1969). Exact distribution of Wilks' likelihood ratio criterion. *Biometrika,* **56**, 109-118.

Press, S. J. (1972). *Applied Multivariate Analysis.* Holt, Rinehart, and Winston Incorporated, New York.

Quan, H. (1987). Some optimal properties of testing hypotheses of elliptically contoured distributions. *Acta Mathematicae Applicatae Sinica* **3**:1, 1-14. Reprinted in Fang and Anderson (1990).

Quan, H. and Fang, K. T. (1987). Unbiasedness of the parameter tests for generalized multivariate distributions. *Acta Mathematicae Applicatae Sinica* **10**, 215-234. Reprinted in Fang and Anderson (1990).

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd Edition. Wiley, New York.

Rao, C. R. and Mitra, S. K. (1971). *Generalized Inverse of Matrices and its Applications*. Wiley, New York.

Richards, D. (1984). Hyperspherical models, fractional derivatives and exponential distributions on matrix spaces. *Sankhyā*, Ser. A **46**, 155-165.

Sampson, A. R. (1983). Positive dependence properties of elliptically symmetric distributions. *Journal of Multivariate Analysis* **13**, 375-381.

Schoenberg, I. J. (1938). Metric spaces and completely monotone functions. *Annals of Mathematics* **39**, 811-841.

Siotani, M., Hayakawa, T. and Fujikoshi, Y. (1985). *Modern Multivariate Statistical Analysis: A Graduate Course and Handbook*. American Sciences Press, Inc., Columbus, Ohio.

Smith, M. D. (1989). On the expectation of a ratio of quadratic forms in normal variables. *Journal of Multivariate Analysis* **31**, 244-257.

Stoyan, D. (1983). *Comparison Methods for Queues and Other Stochastic Models*. Wiley, New York.

Sutradhar, B. C. and Ali, M. M. (1989). A generalization of the Wishart distribution for the elliptical model and its moments for the multivariate t model. *Journal of Multivariate Analysis* **29**, 155-162.

Tang, J. and Gupta, A. K. (1984). On the distribution of the product of independent beta random variables. *Statistics and Probability Letters*, **2**, 165-168.

Tang, J. and Gupta, A. K. (1986). Exact distribution of certain general test statistics in multivariate analysis. *Australian Journal of Statistics*, **28**, 107-114.

Tang, J. and Gupta, A. K. (1987). On the type-B integral equation and the distribution of Wilks' statistic for testing independence of several groups of variables. *Statistics,* **18,** 379-387.

Timm, N. H. (1975). *Multivariate Analysis with Applications in Education and Psychology.* Wadsworth Publishing Company, Inc., Belmont, California.

AUTHOR INDEX

SUBJECT INDEX