Winter 1994

Invariance Properties of Statistical Tests for Dependent Observations

Akhil K. Vaish
Old Dominion University

INVARIANCE PROPERTIES OF STATISTICAL TESTS FOR DEPENDENT OBSERVATIONS

by

Akhil K. Vaish

M.Sc. May 1989, Indian Institute of Technology
Kanpur, India

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirement for the Degree of

DOCTOR OF PHILOSOPHY

COMPUTATIONAL AND APPLIED MATHEMATICS

OLD DOMINION UNIVERSITY
December, 1994

Approved by:

Dr. N. R. Chaganty (Director)

Dr. R. C. Dahiya

Dr. D. N. Naik

Dr. E. P. Markowski

ABSTRACT

INVARIANCE PROPERTIES OF STATISTICAL TESTS

FOR DEPENDENT OBSERVATIONS

Akhil K. Vaish

Old Dominion University, 1994

Director: Dr. N. R. Chaganty

In this dissertation we assume that the observations are from normal populations but are correlated and study the problem of characterizing the class of covariance structures such that the distributions of the popular test statistics remain invariant, that is, they remain the same except for a constant factor. We first obtain some simple extensions and variations of the well known Cauchy-Schwarz inequality. Incidentally, several inequalities that are useful in the detection of outliers can be deduced from our results.

Our main result is a characterization of the class of all nonnegative definite solutions $\mathbf{W}$ to the matrix equation $\mathbf{A}\,\mathbf{W}\,\mathbf{A} = \mathbf{B}$, where $\mathbf{A}$ is a symmetric and $\mathbf{B}$ is a nonnegative definite matrix. We illustrate the proof of this characterization by considering a special case where $\mathbf{A} = \mathbf{B} = \mathbf{A}^* = \mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}'$, $\mathbf{I}$ is the identity matrix and $\mathbf{e}$ is a vector of ones. We thus have an elegant characterization of the class of all nonnegative definite g-inverses of the centering matrix $\mathbf{A}^*$. Next we present the

statistical applications of our matrix theoretic results. For example, we show that the usual two sample t-statistic has a t-distribution if the observations in one of the samples are positively equicorrelated and those in the other sample are negatively equicorrelated with the same correlation in absolute value. More generally, we have a complete characterization of the class of covariance matrices for which the distributional properties of the quadratic forms in ANOVA problems remain invariant. These results are contained in Chapter 3.

In Chapters 4 and 5, we generalize our results to the multivariate test statistics, first considering a special covariance structure that occurs in repeated measurements and later for an arbitrary covariance structure. These include invariance properties of the distributions of quadratic forms in MANOVA problems and one- and two-sample Hotelling's T^2 statistics. As preliminaries to the multivariate results, we obtain a very general version of the Cochran's theorem concerning the independence and Wishartness of the multivariate quadratic forms.

Dedicated

to

Amma and Papa

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Narasinga Rao Chaganty for sharing his knowledge and enthusiasm with me. I sincerely appreciate his help and the time he spent in guiding me through this dissertation.

I am indebted to Dr. Ram C. Dahiya for all his help. I extend my heartfelt gratitude to Dr. Dayanand N. Naik for his constant encouragement. I am thankful to Dr. Edward P. Markowski for being on my dissertation committee. I am obliged to Mr. M. S. Hegde for painstakingly going through the manuscript and correcting numerous errors. Special thanks to Dr. Surya P. Dhakar for his help and support.

Last but not least, I thank Shri Krishana and Shakuntla Vaish, Alpana, Shampa and Amit for their love and moral support.

Contents

Chapter 1

Introduction

1.1 Summary

The common and popular statistical tests are derived under the assumption that samples are taken independently from one or more normal populations. In the study of robustness of these tests much research has been done on the effect of nonnormality on these tests, see Lehmann (1983) for an exposition. Typically, this part of the robustness literature still makes the assumption of mutual independence of the observations. While the independence assumption may be approximately valid, due to the choice of experimental designs, the case of dependence between the observations is of practical as well as aesthetic interest. One can even argue that in practical applications, observations are frequently not independent and the physical systems responsible for generation of the observations automatically introduce some dependence among the observations. It is then clear that the statistical interest should be to determine if procedures valid under the independence assumption continue to remain valid with only a simple adjustment when independence assumption is violated.

The main goal of this dissertation is to consider the above mentioned issue in the context of several common statistical problems. A plethora of literature exists in this general area; in particular, see Jensen (1989a, b, c), Huber (1981), Albers (1978), Ali (1973), Gastwirth and Rubin (1971), Basu and DasGupta (1991), Tukey (1960) and Walsh (1947). While the emphasis in these articles is on examining the effect of specific dependence structures, our goal is to obtain a simpler characterization of the covariance structures under which the usual procedures remain valid with possibly a scale factor adjustment. The organization of this dissertation is as follows: The main results of this dissertation are in Chapter 3 for the univariate test statistics and in Chapters 4 and 5 for the multivariate test statistics. The results presented in Chapter 2 are of independent interest and they play an important role in the simplification of some of the conditions in the main theorems. We now give a brief summary of each chapter in this dissertation.

In Chapter 2, we first obtain an inequality involving two quadratic forms; $\mathbf{y}'\mathbf{A}\mathbf{y}$ and $\mathbf{y}'\mathbf{B}\mathbf{y}$ where $\mathbf{y} \in \Re^n$, $\mathbf{A}$ is a symmetric matrix and $\mathbf{B}$ is a positive semidefinite matrix. This result, stated as Theorem 2.2.1 in Section 2.2, is a simple extension of the well known Cauchy-Schwarz inequality and is used later in Chapter 3 to simplify some of the conditions in our main theorems. Theorem 2.2.1 can also be viewed as a maximization and minimization problem concerning the ratio, $\mathbf{y}'\mathbf{A}\mathbf{y}/\mathbf{y}'\mathbf{B}\mathbf{y}$ of the two quadratic forms. This alternative statement of Theorem 2.2.1 has a natural generalization which we state as Theorem 2.2.5. The special case of Theorem 2.2.5 where the matrix **B** is assumed to be positive definite is known as the Courant-Fischer theorem in the literature. Another special case of Theorem 2.2.1 or Theorem 2.2.5 where we assume the matrix **A** to be of the form $\mathbf{b}\,\mathbf{b}'$, for some vector $\mathbf{b} \in \Re^n$, has numerous applications in statistical methodology. We consider this special case in Theorem 2.2.3 because of its importance to statistics.

In a recent paper, Olkin (1992) presented an interesting survey of several inequalities that are useful in the detection of outliers in statistical data analysis. In Section 2.3, we deduce several of these inequalities from Theorems 2.2.1 and 2.2.3. Thus our results provide a unified treatment of the inequalities given in Olkin's paper. Also in Section 2.3, as another application of Theorem 2.2.3, we derive Scheffé's S-method of constructing simultaneous confidence intervals for the case where the design matrix is not of full rank and the set of estimable functions are linearly dependent.

Our main purpose in Chapter 3 is to characterize the class of covariance matrices for which the distribution of common univariate test statistics remain invariant, that is, the distributions remain the same except for a scale factor. As an important step in achieving this goal, we first characterize, in Theorem 3.2.2, the class of nonnegative definite (n.n.d.) solutions **W** for the consistent matrix equation

$$\mathbf{A}\,\mathbf{W}\,\mathbf{A} = \mathbf{B} \tag{1.1.1}$$

where **A** is any symmetric matrix and **B** is any n.n.d. matrix. Theorem 3.2.2 is also useful in characterizing n.n.d. g-inverses of n.n.d. matrices. Several authors have obtained n.n.d. solutions $\mathbf{W}$ to the matrix equation (1.1.1), for example, Bhimasankaram and Majumdar (1980) and Khatri and Mitra (1976). However, we take an entirely different approach to the problem and obtain a simpler and minimal representation of the class of all n.n.d. solutions to the matrix equation (1.1.1). Theorem 3.2.2 is crucial to the proof of several results including Theorem 5.2.1. To motivate the proof of our main Theorem 3.2.2, we consider in Theorem 3.2.1 a special case where $\mathbf{A} = \mathbf{B} = \mathbf{A}^*$, the centering matrix defined in Section 3.2. We thus obtain an elegant characterization of the class, $\mathcal{G}_n$, of n.n.d. g-inverses of the centering matrix. Another special case of Theorem 3.2.2 useful in ANOVA problems is given in Theorem 3.2.3.

The statistical applications of our theorems are presented in Section 3.3. In particular, in Theorem 3.3.2, we show that the sample variance for a set of correlated normal observations is distributed as chi-square except for a scale adjustment and is independent of the sample mean if and only if the observations are equicorrelated. We also show that the usual two sample t-statistic has a t-distribution if the observations in one of the samples are positively equicorrelated and those in the other sample are negatively equicorrelated with the same correlation in absolute value. In Theorem 3.3.5, we obtain a complete characterization of the class of covariance matrices under which the distributions of various sums of squares in ANOVA problems are independent and distributed as chi-square except for a scale factor. The set of covariance matrices considered by Bhat (1962) in Theorem 2 and by Scariano and Davenport (1987) in Theorem 1 are subsets of the set of covariance matrices we obtain in Theorem 3.3.5. Thus we provide a complete solution to the problems considered by these authors.

In Chapter 4, we generalize the results of Chapter 3 to multivariate test statistics. Two important matrix operations, the Kronecker product $\otimes$ and the Vec operator *vec*, play an important role in the proofs of our theorems in this chapter. Section 4.2 contains a comprehensive summary of the properties of these two matrix operations. In Sections 4.3 and 4.4, we define the matrix normal distribution and the Wishart distribution. We also derive some important properties of these two multivariate distributions by extensively making use of the matrix operators $\otimes$ and *vec*.

Suppose, we have correlated p-variate normal observations, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. Let $\mathbf{\Sigma}$ be the covariance matrix of the vector $\tilde{\mathbf{x}}' = [\mathbf{x}_1{}'\mathbf{x}_2{}' \ldots \mathbf{x}_n{}']$. In Section 4.6, we assume $\mathbf{\Sigma} = \mathbf{W} \otimes \mathbf{V}$, that is, $cov(\mathbf{x}_i, \mathbf{x}_j) = w_{ij}\,\mathbf{V}$ where $\mathbf{W} = (w_{ij})$ and $\mathbf{V}$ are n.n.d matrices. The covariance structure $\mathbf{W} \otimes \mathbf{V}$ occurs naturally in the analysis of multivariate repeated measurement designs where $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ can be thought of as multivariate measurements taken on the same subject at n different time periods. With this assumption, necessary and sufficient conditions for Wishartness and independence of multivariate quadratic forms and a multivariate generalization of the

Cochran's theorem are derived in Section 4.6. Basu et al. (1974) have shown that the sample mean and the covariance matrix are independent and their distributions remain invariant if the matrix $\mathbf{W}$ has an equicorrelated structure. In Section 4.7, we show that the converse of this result holds as well (see Theorem 4.7.1); and further we characterize the class of matrices $\mathbf{W}$ such that the distributions of the popular test statistics in MANOVA problems and one- and two-sample Hotelling's T^2 statistics remain invariant for a fixed but unknown $\mathbf{V}$.

In Chapter 5, we generalize the results obtained in Chapter 4 for the case where the covariance matrix Σ is arbitrary and not necessarily of the form $\mathbf{W} \otimes \mathbf{V}$. Pavur (1987) derived a class of matrices Σ such that the sample covariance matrix has a Wishart distribution and distributed independently of the sample mean vector. He also gave a characterization of Σ under which quadratic forms in MANOVA problems have Wishart distribution and are mutually independent. Unfortunately, the collection of matrices Σ given by Pavur contains matrices which are not n.n.d., see Remarks 5.2.5 and 5.4.2, therefore the collection is only sufficient but not necessary for the distributions to be invariant. In this chapter, we obtain elegant characterizations of the class of matrices Σ such that the distributions of several test statistics remain invariant. These results are multivariate extensions of the theorems contained in Chapters 3 and 4.

1.2 Notations and Conventions

The following notations and conventions are used throughout this dissertation. We denote a random variable having a chi-square distribution with m degrees of freedom by $\chi^2(m)$, a noncentral chi-square distribution with m degrees of freedom and noncentrality parameter δ by $\chi^2(m;\delta)$, a t-distribution with m degrees of freedom by $t(m)$ and an F-distribution with (m_1, m_2) degrees of freedom by $F(m_1, m_2)$. We write $\mathbf{x} \sim \mathbf{y}$ or $\mathbf{x} = \mathbf{y}$ to mean that both $\mathbf{x}$ and $\mathbf{y}$ have the same probability distribution. By $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ we mean that the $p \times 1$ random vector $\mathbf{x}$ has a p-variate normal distribution with mean vector $\boldsymbol{\mu} = E(\mathbf{x})$ of order $p \times 1$ and covariance matrix $\boldsymbol{\Sigma} = V(\mathbf{x})$ of order $p \times p$ where $E(.)$ denotes expected value and $V(.)$ the covariance matrix of the corresponding random variable or vector.

We denote the column space, null space, rank, trace and transpose of the matrix $\mathbf{A}$ by $\mathcal{M}(\mathbf{A})$, $\mathcal{N}(\mathbf{A})$, $r(\mathbf{A})$, $tr(\mathbf{A})$ and $\mathbf{A}'$, respectively. Also, $\mathbf{A}^+$ and $\mathbf{A}^-$ denote the Moore-Penrose inverse and ordinary g-inverse of $\mathbf{A}$, respectively. We follow the definition given in Rao (1973), Table 1, page 67 concerning nonnegative definiteness of matrices. Also, all n.n.d. matrices are assumed to be symmetric. The vector $\mathbf{e}$ represents a vector of ones of order $n \times 1$ whereas $\mathbf{e}_m$ denotes a vector of ones of order $m \times 1$. Similarly, $\mathbf{I}$ represents the identity matrix of order $n \times n$ whereas $\mathbf{I}_m$ denotes the identity matrix of order $m \times m$. The vector $\mathbf{0}$ represents a vector of zeros and $\mathbf{O}$ represents a matrix of zeros of appropriate orders.

Chapter 2

Inequalities for Positive Semidefinite Matrices and Statistical Applications

2.1 Introduction

The characterization of eigenvalues of a real symmetric matrix as the extreme points of the quadratic form involving a symmetric matrix subject to some constraints has been very useful for several branches of science, including statistics. In its general form it is known as the Courant-Fischer min-max theorem, see Bellman (1970), page 115. This theorem is the basis of several inequalities, principal component analysis and other topics in multivariate statistical analysis. In Theorems 2.2.1 and 2.2.5 of this chapter we extend the Courant-Fischer theorem for the ratio of two quadratic forms involving a symmetric matrix **A** in the numerator and a positive semidefinite matrix **B** in the denominator. Under an additional hypothesis that the column space of **A** is contained in the column space of **B**, Theorem 2.2.1 yields inequalities concerning positive semidefiniteness of the difference of the two matrices **A** and **B**. This result is contained in Theorem 2.2.2. Theorems 2.2.1 and 2.2.2 restricted to the matrix $\mathbf{A} = \mathbf{b}\mathbf{b}'$ where $\mathbf{b}$ is a vector in $\Re^n$, have numerous applications in statistical methodology. We treat this special case in Theorem 2.2.3.

In a recent paper, Olkin (1992) presented an interesting survey of several inequalities that are useful in the detection of outliers in statistical data analysis. We deduce most of those inequalities as an application to our Theorems 2.2.2 and 2.2.3. Thus our theorems provide a unified treatment of those inequalities. As a second application to Theorem 2.2.3, we extend the Scheffé's S-method of constructing simultaneous confidence intervals for the case where the design matrix is not of full rank and the set of estimable functions are linearly dependent. Several applications of Theorems 2.2.2 and 2.2.3 in studying invariance property of some statistical tests will be presented in later chapters. The organization of this chapter is as follows: The main theorems of this chapter are in Section 2.2 while the statistical applications are deferred to Section 2.3.

2.2 Main Results

We start with the following elementary and well known lemma. It plays a very important role in the proofs of the theorems in this chapter. The proof of Lemma 2.2.1 can be found in Mirsky (1955), page 200 and in Marshall and Olkin (1979), page 216.

Lemma 2.2.1 *Let* $\mathbf{C}$ *and* $\mathbf{D}$ *be two matrices of order* $n \times k$ *and* $k \times n$*, respectively. Assume that* $n \geq k$*. Then* $(n-k)$ *eigenvalues of the matrix* $\mathbf{CD}$ *are zeros and the* *remaining k eigenvalues of* **CD**, *some of which may be zeros, coincide with the k eigenvalues of the matrix* **DC**.

We develop some preliminaries before stating the main theorems of this section. Let $\mathbf{A}$ be a symmetric matrix of order $n \times n$ and $\mathbf{B}$ a positive semidefinite matrix of order $n \times n$ of rank equal to k. Let $\mathcal{M}(\mathbf{B})$ denote the column space of $\mathbf{B}$. Let $\mathbf{B} = \mathbf{L}\mathbf{L}'$ be the rank factorization of $\mathbf{B}$ where $\mathbf{L}$ is a matrix of full column rank and of order $n \times k$. Let

$$\begin{aligned} \mathbf{R} &= \mathbf{L}(\mathbf{L}'\mathbf{L})^{-1} \\ \mathbf{A}_1 &= \mathbf{R}'\mathbf{A}\mathbf{R}. \end{aligned} \tag{2.2.1}$$

Note that the Moore-Penrose inverse of $\mathbf{B}$, see Searle (1982), page 220, is given by $\mathbf{B}^+ = \mathbf{R}\mathbf{R}'$. Observe that the column spaces of $\mathbf{B}$, $\mathbf{B}^+$, $\mathbf{R}$ and $\mathbf{L}$ are all equal. We use this observation in the proofs of the results in this section without mentioning it explicitly. Let $\{\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k\}$ be an ordered set of eigenvalues of $\mathbf{A}_1$. Applying Lemma 2.2.1 for $\mathbf{C} = \mathbf{R}$ and $\mathbf{D} = \mathbf{R}'\mathbf{A}$ we can see that the set of eigenvalues of the matrix $\mathbf{B}^+\mathbf{A}$ is given by $\{\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k, 0, \ldots, 0\}$. It is possible that some of the λ_i's may be zeros and also, all the λ_i's may be negative. Therefore λ_1 need not be the largest eigenvalue of $\mathbf{B}^+\mathbf{A}$. Similarly, λ_k need not be the smallest eigenvalue of $\mathbf{B}^+\mathbf{A}$. In fact, the largest eigenvalue of $\mathbf{B}^+\mathbf{A}$ is given by $\max\{0, \lambda_1\}$ and the smallest eigenvalue of $\mathbf{B}^+\mathbf{A}$ is equal to $\min\{0, \lambda_k\}$. We are now ready to state an inequality concerning two quadratic forms.

Theorem 2.2.1 *Let* $\mathbf{A}$ *be a symmetric matrix of order* $n \times n$. *Let* $\mathbf{B}$ *be a positive semidefinite matrix of order* $n \times n$ *and rank equal to* k. *Let* $\{\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k, 0, \ldots, 0\}$ *be the set of* n *eigenvalues of* $\mathbf{B}^+\mathbf{A}$. *Then*

$$\lambda_k\, \mathbf{y}'\mathbf{B}\,\mathbf{y} \leq \mathbf{y}'\mathbf{A}\,\mathbf{y} \leq \lambda_1\, \mathbf{y}'\mathbf{B}\,\mathbf{y} \tag{2.2.2}$$

for all $\mathbf{y} \in \mathcal{M}(\mathbf{B})$. *There exists an eigenvector* $\mathbf{y}_i$ *of* $\mathbf{B}^+\mathbf{A}$ *corresponding to the eigenvalue* λ_i *such that* $\mathbf{y}_i \in \mathcal{M}(\mathbf{B})$ *for* $1 \leq i \leq k$. *Further, equality holds in the first and second inequality of* (2.2.2) *if we choose* $\mathbf{y}$ *equal to* $\mathbf{y}_k$ *and* $\mathbf{y}_1$, *respectively.*

Proof. Let $\mathbf{B} = \mathbf{L}\mathbf{L}'$ be the rank factorization of $\mathbf{B}$. Let $\mathbf{R}$ and $\mathbf{A}_1$ be as defined in (2.2.1). Since λ_1 and λ_k are the largest and the smallest eigenvalues of $\mathbf{A}_1$, by a well known inequality, see (1f.2.1) of Rao (1973), page 62, we have

$$\lambda_k\, \mathbf{v}'\mathbf{v} \leq \mathbf{v}'\mathbf{A}_1\mathbf{v} \leq \lambda_1\, \mathbf{v}'\mathbf{v} \tag{2.2.3}$$

for all $\mathbf{v} \in \Re^k$. Let $\mathbf{y}$ be a vector in $\mathcal{M}(\mathbf{B})$ and $\mathbf{v} = \mathbf{L}'\mathbf{y}$. It is easy to verify that $\mathbf{y} = \mathbf{R}\mathbf{v}$, since $\mathbf{y}$ is also in the column space of $\mathbf{L}$. Thus, we have

$$\begin{aligned} \mathbf{v}'\mathbf{v} &= \mathbf{y}'\mathbf{B}\mathbf{y} \\ \mathbf{v}'\mathbf{A}_1\mathbf{v} &= \mathbf{y}'\mathbf{A}\mathbf{y}. \end{aligned} \tag{2.2.4}$$

The assertion (2.2.2) follows from (2.2.3) and (2.2.4). We now proceed to show that the two inequalities in (2.2.2) become equalities for appropriate choices of $\mathbf{y}$. For $1 \leq i \leq k$, let $\mathbf{v}_i \neq 0$ be an eigenvector of $\mathbf{A}_1$ corresponding to the eigenvalue λ_i and

$\mathbf{y}_i = \mathbf{R}\,\mathbf{v}_i$. Note that $\mathbf{y}_i \neq 0$, since $\mathbf{R}$ is of full column rank and $\mathbf{v}_i \neq 0$. We also have

$$
\begin{aligned}
\mathbf{B}^{+}\mathbf{A}\,\mathbf{y}_i &= \mathbf{R}\mathbf{R}'\mathbf{A}\mathbf{R}\,\mathbf{v}_i \\
&= \mathbf{R}\,\mathbf{A}_1\,\mathbf{v}_i = \lambda_i\,\mathbf{R}\,\mathbf{v}_i \\
&= \lambda_i\,\mathbf{y}_i. \qquad (2.2.5)
\end{aligned}
$$

Thus $\mathbf{y}_i$ is an eigenvector of $\mathbf{B}^{+}\mathbf{A}$ corresponding to the eigenvalue λ_i. Clearly $\mathbf{y}_i \in \mathcal{M}(\mathbf{B})$ since it is in the column space of $\mathbf{R}$. Therefore from (2.2.4) we have $\mathbf{v}_i'\,\mathbf{A}_1\,\mathbf{v}_i = \mathbf{y}_i'\,\mathbf{A}\,\mathbf{y}_i$ and $\mathbf{v}_i'\,\mathbf{v}_i = \mathbf{y}_i'\,\mathbf{B}\,\mathbf{y}_i$. Thus $\mathbf{y}_i'\,\mathbf{A}\,\mathbf{y}_i = \lambda_i\,\mathbf{y}_i'\,\mathbf{B}\,\mathbf{y}_i$ for $1 \leq i \leq k$. Therefore, the first and second inequality in (2.2.2) become equalities if we choose $\mathbf{y}$ equal to $\mathbf{y}_k$ and $\mathbf{y}_1$, respectively. This completes the proof of Theorem 2.2.1.

When $\lambda_1 = \cdots = \lambda_k$, from (2.2.2), we have $\mathbf{y}'\,\mathbf{A}\,\mathbf{y} = \lambda_1\,\mathbf{y}'\,\mathbf{B}\,\mathbf{y}$ for all $\mathbf{y} \in \mathcal{M}(\mathbf{B})$. The following example shows that this need not be true and also (2.2.2) may not hold for vectors $\mathbf{y}$ not in the $\mathcal{M}(\mathbf{B})$.

Example 2.2.1 Let $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 1/4 & 1/4 \\ 1/4 & 1/4 \end{bmatrix}$. Clearly, $\mathbf{B}$ is positive semidefinite matrix of rank $k = 1$. The Moore-Penrose inverse of $\mathbf{B}$ is given by $\mathbf{B}^{+} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. It is easy to verify that the set of eigenvalues of $\mathbf{B}^{+}\mathbf{A}$ is given by $\{1, 0\}$ and therefore λ_1 equals 1. Consider the vector $\mathbf{y}' = (2, 0)$ which is not in the column space of $\mathbf{B}$. A little calculation shows that $\mathbf{y}'\mathbf{B}\,\mathbf{y} = 1$ and $\mathbf{y}'\mathbf{A}\,\mathbf{y} = 4$ and hence $\mathbf{y}'\,\mathbf{A}\,\mathbf{y} > \lambda_1\,\mathbf{y}'\mathbf{B}\,\mathbf{y}$. Similarly, for $\mathbf{y}' = (0, 2)$, we have $\mathbf{y}'\,\mathbf{A}\,\mathbf{y} < \lambda_1\,\mathbf{y}'\mathbf{B}\,\mathbf{y}$. Therefore this example shows that the inequalities in (2.2.2) need not hold for all $\mathbf{y}$.

The following theorem gives sufficient condition for the inequality (2.2.2) to hold for all $\mathbf{y} \in \Re^n$.

Theorem 2.2.2 *Let* $\mathbf{A}$ *be a symmetric matrix of order* $n \times n$. *Let* $\mathbf{B}$ *be a positive semidefinite matrix of order* $n \times n$ *and rank equal to* k. *Let* $\{\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k, 0, \ldots, 0\}$ *be the set of* n *eigenvalues of* $\mathbf{B}^+\mathbf{A}$. *If* $\mathcal{M}(\mathbf{A}) \subseteq \mathcal{M}(\mathbf{B})$ *then the matrices* $\lambda_1 \mathbf{B} - \mathbf{A}$ *and* $\mathbf{A} - \lambda_k \mathbf{B}$ *are positive semidefinite.*

Proof: Fix, $\mathbf{y} \in \Re^n$. Then we can write $\mathbf{y} = \mathbf{y}_b + \mathbf{y}_b^{\perp}$, where $\mathbf{y}_b$ is the projection of $\mathbf{y}$ onto the column space of $\mathbf{B}$ and $\mathbf{y}_b^{\perp} = \mathbf{y} - \mathbf{y}_b$. Note that $\mathbf{B}\mathbf{y}_b^{\perp} = 0$. If $\mathcal{M}(\mathbf{A}) \subseteq \mathcal{M}(\mathbf{B})$, we also have $\mathbf{A}\mathbf{y}_b^{\perp} = 0$. Therefore,

$$\begin{aligned} \mathbf{y}'\mathbf{A}\mathbf{y} &= \mathbf{y}_b'\mathbf{A}\mathbf{y}_b \\ \mathbf{y}'\mathbf{B}\mathbf{y} &= \mathbf{y}_b'\mathbf{B}\mathbf{y}_b. \end{aligned} \tag{2.2.6}$$

Since $\mathbf{y}_b \in \mathcal{M}(\mathbf{B})$ by (2.2.2) of Theorem 2.2.1, we have

$$\lambda_k\, \mathbf{y}_b'\mathbf{B}\mathbf{y}_b \leq \mathbf{y}_b'\mathbf{A}\mathbf{y}_b \leq \lambda_1\, \mathbf{y}_b'\mathbf{B}\mathbf{y}_b. \tag{2.2.7}$$

Combining (2.2.6) and (2.2.7), we get

$$\lambda_k\, \mathbf{y}'\mathbf{B}\mathbf{y} \leq \mathbf{y}'\mathbf{A}\mathbf{y} \leq \lambda_1\, \mathbf{y}'\mathbf{B}\mathbf{y}. \tag{2.2.8}$$

Since $\mathbf{y} \in \Re^n$ is arbitrary, (2.2.8) shows that the matrices $\lambda_1 \mathbf{B} - \mathbf{A}$ and $\mathbf{A} - \lambda_k \mathbf{B}$ are positive semidefinite.

The following lemma shows that for any two symmetric matrices $\mathbf{A}$ and $\mathbf{B}$, if $\mathcal{M}(\mathbf{A}) \subseteq \mathcal{M}(\mathbf{B})$ then the set of eigenvalues of $\mathbf{B}^{-}\mathbf{A}$ is invariant of the choice of the g-inverse $\mathbf{B}^{-}$ of $\mathbf{B}$. Thus we can replace $\mathbf{B}^{+}$ by any g-inverse $\mathbf{B}^{-}$ of $\mathbf{B}$ in the statement of Theorem 2.2.2.

Lemma 2.2.2 *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two symmetric matrices both of order* $n \times n$. *If* $\mathcal{M}(\mathbf{A}) \subseteq \mathcal{M}(\mathbf{B})$ *then the set of eigenvalues of* $\mathbf{B}^{-}\mathbf{A}$ *is invariant of the choice of the g-inverse* $\mathbf{B}^{-}$ *of* $\mathbf{B}$.

Proof. Let $\mathbf{A}$ and $\mathbf{B}$ be two symmetric matrices of order $n \times n$ such that $\mathcal{M}(\mathbf{A}) \subseteq \mathcal{M}(\mathbf{B})$. By spectral decomposition, there exists an orthogonal matrix $\mathbf{P}$ such that

$$\mathbf{A} = \mathbf{P} \begin{bmatrix} \mathbf{\Lambda} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} \mathbf{P}' = \mathbf{P}_1 \mathbf{\Lambda} \mathbf{P}_1'$$

where $\mathbf{\Lambda}$ is a diagonal matrix and $\mathbf{P} = [\mathbf{P}_1 \quad \mathbf{P}_2]$. Since $\mathcal{M}(\mathbf{A}) = \mathcal{M}(\mathbf{P}_1)$ and $\mathcal{M}(\mathbf{A}) \subseteq \mathcal{M}(\mathbf{B})$, we have $\mathcal{M}(\mathbf{P}_1) \subseteq \mathcal{M}(\mathbf{B})$. Hence we can write $\mathbf{P}_1 = \mathbf{B}\,\mathbf{U}$ for some matrix $\mathbf{U}$. Therefore,

$$\mathbf{A} = \mathbf{P}_1 \mathbf{\Lambda} \mathbf{P}_1' = \mathbf{B}\,\mathbf{U}\,\mathbf{\Lambda}\,\mathbf{U}'\mathbf{B} = \mathbf{B}\,\mathbf{V}\,\mathbf{B} \tag{2.2.9}$$

where $\mathbf{V} = \mathbf{U}\,\mathbf{\Lambda}\,\mathbf{U}'$ is a symmetric matrix. Let $\mathbf{B}^{-}$ be a g-inverse of $\mathbf{B}$. Then, if we choose $\mathbf{C} = \mathbf{B}^{-}\,\mathbf{B}\,\mathbf{V}$ and $\mathbf{D} = \mathbf{B}$, by Lemma 2.2.1, we get the same set of eigenvalues for $\mathbf{B}^{-}\mathbf{A}$ and $\mathbf{B}\,\mathbf{V}$. Thus, the eigenvalues of $\mathbf{B}^{-}\,\mathbf{A}$ do not depend on the choice of the g-inverse $\mathbf{B}^{-}$ of $\mathbf{B}$. This completes the proof of the lemma.

The following example shows that the conclusion of Lemma 2.2.2 need not be true if we do not assume that $\mathcal{M}(\mathbf{A})$ is contained in $\mathcal{M}(\mathbf{B})$.

Example 2.2.2 Consider the matrices $\mathbf{A}$ and $\mathbf{B}$ as in Example 2.2.1. It is easy to verify that $\mathcal{M}(\mathbf{A})$ is not contained in $\mathcal{M}(\mathbf{B})$. We have seen in Example 2.2.1 that the set of eigenvalues of $\mathbf{B}^+\mathbf{A}$ is given by $\{1, 0\}$. Consider another g-inverse $\mathbf{B}^- = \begin{bmatrix} 2 & 0 \\ 2 & 0 \end{bmatrix}$ of $\mathbf{B}$. We can easily verify that the set of eigenvalues of $\mathbf{B}^-\mathbf{A}$ is given by $\{2, 0\}$, which is different from the set of eigenvalues of $\mathbf{B}^+\mathbf{A}$. Thus, the conclusion of Lemma 2.2.2 need not be true if $\mathcal{M}(\mathbf{A})$ is not contained in $\mathcal{M}(\mathbf{B})$.

Let $\mathbf{b}$ be a vector in $\Re^n$. Theorems 2.2.1 and 2.2.2 restricted to the matrix $\mathbf{A} = \mathbf{b}\mathbf{b}'$ give rise to several interesting inequalities. We treat this special case in Theorem 2.2.3. In Section 2.3, we use Theorem 2.2.3 to obtain several inequalities that are useful in the detection of outliers in statistical data analysis.

Theorem 2.2.3 *Let* $\mathbf{B}$ *be a positive semidefinite matrix of order* $n \times n$. *Let* $\mathbf{B}^+$ *be the Moore-Penrose inverse of* $\mathbf{B}$. *Let* $\mathcal{M}(\mathbf{B})$ *denotes the column space of* $\mathbf{B}$. *If* $\mathbf{b}$ *is an* $n \times 1$ *vector then*

$$(\mathbf{b}'\mathbf{y})^2 \leq \mathbf{b}'\mathbf{B}^+\mathbf{b}\ \mathbf{y}'\mathbf{B}\mathbf{y} \tag{2.2.10}$$

for all $\mathbf{y} \in \mathcal{M}(\mathbf{B})$. *Moreover, equality holds in* (2.2.10) *if we choose* $\mathbf{y} = \mathbf{B}^+\mathbf{b}$. *Also, if rank of* $\mathbf{B}$ *equals 1 then equality holds in* (2.2.10) *for all* $\mathbf{y} \in \mathcal{M}(\mathbf{B})$. *If* $\mathbf{b} \in \mathcal{M}(\mathbf{B})$ *then* (2.2.10) *holds for all* $\mathbf{y} \in \Re^n$, *equivalently, the matrix* $(\mathbf{b}'\mathbf{B}^+\mathbf{b})\,\mathbf{B} - \mathbf{b}\mathbf{b}'$ *is positive semidefinite.*

Proof. Let $\mathbf{b}$ be an $n \times 1$ vector. Let us choose $\mathbf{A} = \mathbf{b}\,\mathbf{b}'$ in Theorems 2.2.1 and 2.2.2. Letting $\mathbf{C} = \mathbf{B}^{+}\mathbf{b}$ and $\mathbf{D} = \mathbf{b}'$ in Lemma 2.2.1, we can see that the set of eigenvalues of $\mathbf{B}^{+}\mathbf{A}$ is given by $\{\mathbf{b}'\mathbf{B}^{+}\mathbf{b}, 0, \ldots, 0\}$. Let the rank of $\mathbf{B}$ be equal to k. Then, in the notation of Theorem 2.2.1, the eigenvalue λ_1 equals $\mathbf{b}'\mathbf{B}^{+}\mathbf{b}$ and $\lambda_k = \lambda_1$ if $k = 1$ and $\lambda_k = 0$ if $k \geq 2$. Therefore, Theorem 2.2.3 follows from Theorems 2.2.1 and 2.2.2.

The following corollary is an easy consequence of Theorem 2.2.3. We apply this corollary in Section 2.3 to extend Scheffé's S-method of constructing simultaneous confidence intervals when the design matrix is not of full rank and the set of estimable functions are linearly dependent.

Corollary 2.2.1 *Let* $\mathbf{B}$ *be a positive semidefinite matrix of order* $n \times n$ *and* $\mathbf{B}^{-}$ *be a g-inverse of* $\mathbf{B}$. *Let* $\mathbf{b} \in \mathcal{M}(\mathbf{B})$, *then* $\eta\,\mathbf{B} - \mathbf{b}\,\mathbf{b}'$ *is positive semidefinite if and only if* $\eta \geq \mathbf{b}'\,\mathbf{B}^{-}\,\mathbf{b}$.

Proof: Let $\mathbf{b} \in \mathcal{M}(\mathbf{B})$. It is easy to verify that $\mathbf{b}'\,\mathbf{B}^{-}\,\mathbf{b} = \mathbf{b}'\,\mathbf{B}^{+}\,\mathbf{b}$ for any choice of the g-inverse $\mathbf{B}^{-}$ of $\mathbf{B}$. Suppose $\eta \geq \mathbf{b}'\,\mathbf{B}^{+}\,\mathbf{b}$, then from Theorem 2.2.3, we have

$$\mathbf{y}'\,\mathbf{b}\,\mathbf{b}'\,\mathbf{y} \leq (\mathbf{b}'\,\mathbf{B}^{+}\,\mathbf{b})\;\; \mathbf{y}'\,\mathbf{B}\,\mathbf{y} \leq \eta\,\mathbf{y}'\,\mathbf{B}\,\mathbf{y} \qquad (2.2.11)$$

for all $\mathbf{y} \in \Re^{n}$. Therefore, $\eta\,\mathbf{B} - \mathbf{b}\,\mathbf{b}'$ is positive semidefinite. The other implication follows easily, if we choose $\mathbf{y} = \mathbf{B}^{+}\,\mathbf{b}$.

Theorem 2.2.3 essentially asserts that if $\mathbf{B}$ is a positive semidefinite matrix and

$\mathbf{b} \in \Re^n$, then

$$\sup_{\substack{\mathbf{y} \in \mathcal{M}(\mathbf{B}) \\ \mathbf{y} \neq 0}} \frac{(\mathbf{b}'\mathbf{y})^2}{\mathbf{y}'\mathbf{B}\mathbf{y}} = \mathbf{b}'\mathbf{B}^+\mathbf{b}. \tag{2.2.12}$$

Thus we see that Theorem 2.2.3 generalizes the result contained in Appendix A4 of Seber (1977), page 388, where the above equality (2.2.12) was obtained for positive definite matrix **B**. Similarly, if **A** is symmetric and **B** is positive semidefinite then the conclusion of Theorem 2.2.1 can be restated as

$$\sup_{\substack{\mathbf{y} \in \mathcal{M}(\mathbf{B}) \\ \mathbf{y} \neq 0}} \frac{\mathbf{y}'\mathbf{A}\mathbf{y}}{\mathbf{y}'\mathbf{B}\mathbf{y}} = \lambda_1, \qquad \inf_{\substack{\mathbf{y} \in \mathcal{M}(\mathbf{B}) \\ \mathbf{y} \neq 0}} \frac{\mathbf{y}'\mathbf{A}\mathbf{y}}{\mathbf{y}'\mathbf{B}\mathbf{y}} = \lambda_k \tag{2.2.13}$$

where λ_1 and λ_k are the eigenvalues of $\mathbf{B}^+\mathbf{A}$ defined in Theorem 2.2.1. A similar representation is also true for the other eigenvalues λ_p for $2 \leq p \leq (k-1)$ and is given in the following theorem.

Theorem 2.2.4 *Let* **A** *be a symmetric matrix of order* $n \times n$*. Let* **B** *be positive semidefinite matrix of order* $n \times n$ *and rank equal to* k*. Let* $\{\lambda_1 \geq \cdots \geq \lambda_k, 0, \ldots, 0\}$ *be the set of* n *eigenvalues of* $\mathbf{B}^+\mathbf{A}$*. Then there exist eigenvectors* $\{\mathbf{y}_1, \ldots, \mathbf{y}_k\}$ *of* $\mathbf{B}^+\mathbf{A}$ *corresponding to the eigenvalues* $\{\lambda_1, \ldots, \lambda_k\}$ *such that* $\mathbf{y}_i \in \mathcal{M}(\mathbf{B})$ *and* $\mathbf{y}_i'\mathbf{B}\mathbf{y}_j = 0$ *for* $1 \leq i \neq j \leq k$*. Further,*

$$\sup_{\{\mathbf{y} \in \mathcal{B}_p,\ \mathbf{y} \neq 0\}} \frac{\mathbf{y}'\mathbf{A}\mathbf{y}}{\mathbf{y}'\mathbf{B}\mathbf{y}} = \lambda_p \tag{2.2.14}$$

where $\mathcal{B}_p = \{\mathbf{y} \in \mathcal{M}(\mathbf{B}) : \mathbf{y}_i'\mathbf{B}\mathbf{y} = 0,\ 1 \leq i \leq (p-1)\}$ *for* $2 \leq p \leq (k-1)$*.*

Proof. Let $\mathbf{A}$ be symmetric and $\mathbf{B}$ be positive semidefinite matrix of rank equal to k and $\mathbf{B}^+$ denote the Moore-Penrose inverse of $\mathbf{B}$. Let $\mathbf{L}$, $\mathbf{R}$ and $\mathbf{A}_1$ be as defined in (2.2.1). Let $\{\lambda_1 \geq \cdots \geq \lambda_k\}$ be the set of ordered eigenvalues of $\mathbf{A}_1$ and $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be corresponding orthogonal eigenvectors. By Theorem 1 of Bellman (1970), page 113, we have

$$\sup_{\substack{\{\mathbf{v} \in \Re^k \,:\, \mathbf{v}_i' \mathbf{v} = 0\} \\ \mathbf{v} \neq 0 \qquad 1 \leq i \leq (p-1)}} \frac{\mathbf{v}' \mathbf{A}_1 \mathbf{v}}{\mathbf{v}' \mathbf{v}} = \lambda_p \qquad \text{for } 2 \leq p \leq (k-1). \tag{2.2.15}$$

Let $\mathbf{y} = \mathbf{R}\,\mathbf{v}$, then as $\mathbf{v}$ varies in $\Re^k$, the vector $\mathbf{y}$ varies in $\mathcal{M}(\mathbf{B})$ and by (2.2.4), we have $\mathbf{v}'\mathbf{v} = \mathbf{y}'\,\mathbf{B}\,\mathbf{y}$ and $\mathbf{v}'\,\mathbf{A}_1\,\mathbf{v} = \mathbf{y}'\,\mathbf{A}\,\mathbf{y}$. Let us define $\mathbf{y}_i = \mathbf{R}\,\mathbf{v}_i$ for $1 \leq i \leq k$. Then by Theorem 2.2.1, $\mathbf{y}_i$ is the eigenvector of $\mathbf{B}^+\,\mathbf{A}$ corresponding to the eigenvalue λ_i. Further $\mathbf{y}_i'\mathbf{B}\,\mathbf{y}_j = 0$, since $\mathbf{v}_i = \mathbf{L}'\,\mathbf{y}_i$ and $\mathbf{v}_i'\mathbf{v}_j = 0$ for $1 \leq i \neq j \leq k$. The identity (2.2.14) now follows from (2.2.15).

The main drawback of Theorem 2.2.4 lies in the fact that it characterizes the eigenvalues λ_p as functions of the eigenvectors $\{\mathbf{y}_i,\ 1 \leq i \leq k\}$. The following theorem removes this dependence on the eigenvectors and provides a characterization of λ_p in the form of a min-max theorem. If we assume $\mathbf{B}$ to be a positive definite matrix, then the conclusion of Theorem 2.2.1 as restated in (2.2.13) and Theorem 2.2.5 below reduces to the theorem popularly known as the Courant-Fischer theorem.

Theorem 2.2.5 *Let $\mathbf{A}$ be a symmetric matrix of order $n \times n$. Let $\mathbf{B}$ be a positive semidefinite matrix of order $n \times n$ and rank equal to k. Let $\{\lambda_1 \geq \cdots \geq \lambda_k, 0, \ldots, 0\}$*

be the set of n eigenvalues of $\mathbf{B}^+\mathbf{A}$. *Then*

$$\lambda_p = \inf_{\substack{\{\mathbf{x}_i\,:\,\mathbf{x}_i'\mathbf{B}\,\mathbf{x}_i=1\}\\ 1\le i\le(p-1)}} \sup_{\substack{\mathbf{y}\in\mathcal{C}_p\\ \mathbf{y}\neq 0}} \frac{\mathbf{y}'\mathbf{A}\,\mathbf{y}}{\mathbf{y}'\mathbf{B}\,\mathbf{y}}, \tag{2.2.16}$$

$$\lambda_{k-p+1} = \sup_{\substack{\{\mathbf{x}_i\,:\,\mathbf{x}_i'\mathbf{B}\,\mathbf{x}_i=1\}\\ 1\le i\le(p-1)}} \inf_{\substack{\mathbf{y}\in\mathcal{C}_p\\ \mathbf{y}\neq 0}} \frac{\mathbf{y}'\mathbf{A}\,\mathbf{y}}{\mathbf{y}'\mathbf{B}\,\mathbf{y}} \tag{2.2.17}$$

where $\mathcal{C}_p = \{\mathbf{y}\in\mathcal{M}(\mathbf{B}) : \mathbf{x}_i'\mathbf{B}\,\mathbf{y} = 0,\ \ 1\le i\le (p-1)\}$ *for* $2\le p\le (k-1)$.

Proof. The proof of this theorem parallels the proof of Theorem 2.2.4. Let $\mathbf{L}$, $\mathbf{R}$ and $\mathbf{A}_1$ be as defined in (2.2.1). Let $\{\lambda_1\ge\cdots\ge\lambda_k\}$ be the set of ordered eigenvalues of $\mathbf{A}_1$. Theorem 2 of Bellman (1970), page 115, applied to $\mathbf{A}_1$ gives

$$\lambda_p = \inf_{\substack{\{\mathbf{v}_i\,:\,\mathbf{v}_i'\mathbf{v}_i=1\}\\ 1\le i\le(p-1)}} \sup_{\substack{\{\mathbf{v}\,:\,\mathbf{v}_i'\mathbf{v}=0\}\\ \mathbf{v}\neq 0}} \frac{\mathbf{v}'\mathbf{A}_1\,\mathbf{v}}{\mathbf{v}'\mathbf{v}}, \tag{2.2.18}$$

$$\lambda_{k-p+1} = \sup_{\substack{\{\mathbf{v}_i\,:\,\mathbf{v}_i'\mathbf{v}_i=1\}\\ 1\le i\le(p-1)}} \inf_{\substack{\{\mathbf{v}\,:\,\mathbf{v}_i'\mathbf{v}=0\}\\ \mathbf{v}\neq 0}} \frac{\mathbf{v}'\mathbf{A}_1\,\mathbf{v}}{\mathbf{v}'\mathbf{v}} \tag{2.2.19}$$

for $2\le p\le (k-1)$. The equalities (2.2.16) and (2.2.17) follow from (2.2.18) and (2.2.19) if we make the transformation $\mathbf{y} = \mathbf{R}\,\mathbf{v}$ and $\mathbf{x}_i = \mathbf{R}\,\mathbf{v}_i$ for $1\le i\le (k-2)$ and by noting that $\mathbf{v}'\mathbf{A}_1\mathbf{v} = \mathbf{y}'\mathbf{A}\,\mathbf{y}$ and $\mathbf{v}'\mathbf{v} = \mathbf{y}'\mathbf{B}\,\mathbf{y}$. This completes the proof of the theorem.

2.3 Statistical Applications

In this section, we present some applications of the theorems in Section 2.2. Our first application deals with some inequalities that are useful for the detection of outliers in statistical data. As a second application, we extend Scheffé's S-method of constructing simultaneous confidence intervals when the design matrix is not of full rank and the set of given estimable functions are linearly dependent.

Application 2.3.1 In a recent paper, Olkin (1992) considered the following problem which is of great interest in the detection of outliers. Given, the mean and the standard deviation of a finite sample, find the maximum deviation of any particular observation from the sample mean as a multiple of the sample standard deviation. More specifically, let $\{y_1, \ldots, y_n\}$ be a sample of n observations then the problem is to find the minimum value of c such that

$$(y_k - \bar{y})^2 \leq c \sum_{i=1}^{n} (y_i - \bar{y})^2, \quad k = 1, \ldots, n \tag{2.3.1}$$

where $\bar{y} = \sum_{i=1}^{n} y_i / n$ is the sample mean. The above problem and $c = (n-1)/n$ as the best solution was first brought into the limelight by Samuelson (1968). The inequality (2.3.1) with $c = (n-1)/n$ is now popularly known as Samuelson's inequality. Some extensions of Samuelson's inequality can be found in the paper by Wolkowicz and Styan (1979). Recently, Olkin (1992) gave an interesting survey of the known proofs of Samuelson's inequality and raised the question whether there is room for yet another proof. He then provided a new proof with some generalizations. We now show that Samuelson's inequality and several other inequalities in Olkin (1992) follow from our theorems of Section 2.2. Let $\mathbf{B} = \mathbf{I} - \frac{1}{n} \mathbf{e}\,\mathbf{e}'$. Note that $\mathbf{B}$ is symmetric and idempotent matrix. Hence $\mathbf{B}$ is positive semidefinite and $\mathbf{B}^+ = \mathbf{B}$. Fix, $1 \leq k \leq n$. Consider the vector $\mathbf{b}_1$ where the jth component is given by

$$b_{1j} = \begin{cases} 1 - \dfrac{1}{n} & \text{if } j = k \\ -\dfrac{1}{n} & \text{if } j \neq k. \end{cases} \tag{2.3.2}$$

Since $\mathbf{b}_1' \mathbf{e} = 0$, we have $\mathbf{b}_1 \in \mathcal{M}(\mathbf{B})$. Also, $\mathbf{b}_1' \mathbf{B}^+ \mathbf{b}_1 = \mathbf{b}_1' \mathbf{b}_1 = (n-1)/n$. If we choose $\mathbf{b} = \mathbf{b}_1$, by the last assertion of Theorem 2.2.3, we have $((n-1)/n)\,\mathbf{B} - \mathbf{b}\,\mathbf{b}'$ a positive semidefinite matrix. Thus for any $\mathbf{y} \in \Re^n$, we get

$$\mathbf{y}' \mathbf{b}_1 \mathbf{b}_1' \mathbf{y} \leq \frac{(n-1)}{n} \mathbf{y}' \mathbf{B}\, \mathbf{y} \tag{2.3.3}$$

which is equivalent to Samuelson's inequality:

$$(y_k - \overline{y})^2 \leq \frac{(n-1)}{n} \sum_{i=1}^{n} (y_i - \overline{y})^2. \tag{2.3.4}$$

In a similar fashion, we can deduce inequalities (2.3) and (2.4) of Olkin (1992) as a consequence of Theorem 2.2.3 if we choose $\mathbf{b} = \mathbf{b}_2$ and $\mathbf{b} = \mathbf{b}_3$, respectively. Where the jth component of the vectors $\mathbf{b}_2$ and $\mathbf{b}_3$ are respectively given by

$$b_{2j} = \begin{cases} \dfrac{1}{k} - \dfrac{1}{n} & \text{if } 1 \leq j \leq k \\ -\dfrac{1}{n} & \text{if } k+1 \leq j \leq n; \end{cases} \tag{2.3.5}$$

$$b_{3j} = \begin{cases} \dfrac{1}{k} & \text{if } 1 \leq j \leq k \\ -\dfrac{1}{r} & \text{if } k+1 \leq j \leq k+r \\ 0 & \text{if } k+r < j \leq n. \end{cases} \tag{2.3.6}$$

Also, the inequality in Olkin (1992) involving Gini mean difference due to Nair (1956) follows from Theorem 2.2.3, if we choose $\mathbf{b} = \mathbf{b}_4$ where the jth component of $\mathbf{b}_4$ is given by

$$b_{4j} = \begin{cases} \dfrac{2(2i-n-1)}{n(n-1)} & \text{if } 1 \leq j \leq n. \end{cases} \tag{2.3.7}$$

Let us choose the vector $\mathbf{b} = \mathbf{b}_5$ in Theorem 2.2.3 where the jth component of $\mathbf{b}_5$ is given by

$$b_{5j} = \begin{cases} -1 & \text{if } j = 1 \\ 1 & \text{if } j = n \\ 0 & \text{otherwise.} \end{cases} \tag{2.3.8}$$

Clearly $\mathbf{b}_5 \in \mathcal{M}(\mathbf{B})$ and $\mathbf{b}_5' \, \mathbf{B}^+ \, \mathbf{b}_5 = \mathbf{b}_5' \, \mathbf{b}_5 = 2$. Thus, from Theorem 2.2.3, $2\,\mathbf{B} - \mathbf{b}_5 \, \mathbf{b}_5'$ is a positive semidefinite matrix. For a vector $\mathbf{y}' = (y_1, \ldots, y_n)$ let $\tilde{\mathbf{y}}' = (y_{(1)}, \ldots, y_{(n)})$ where $y_{(i)}$'s are the ordered values of the components of $\mathbf{y}$. Since $2\mathbf{B} - \mathbf{b}_5 \, \mathbf{b}_5'$ is positive semidefinite, we have

$$\tilde{\mathbf{y}}' \, \mathbf{b}_5 \, \mathbf{b}_5 \, \tilde{\mathbf{y}} \leq 2 \, \tilde{\mathbf{y}}' \, \mathbf{B} \, \tilde{\mathbf{y}} \tag{2.3.9}$$

which after simplification reduces to an inequality, due to Thomson (1955), given by

$$(y_{(n)} - y_{(1)})^2 \leq 2 \sum_{i=1}^{n} (y_i - \bar{y})^2. \tag{2.3.10}$$

We now show that the multidimensional inequalities contained in Olkin (1992) can also be deduced from our theorems. Let $\mathbf{W}$ be a matrix of order $l \times n$ such that $\mathbf{W}\,\mathbf{e} = 0$ and $\mathbf{W}\,\mathbf{W}' = \mathbf{I}_l$. Let $\mathbf{B} = \mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}'$ and $\mathbf{A} = \mathbf{W}'\mathbf{W}$. Since $\mathbf{B}^+ = \mathbf{B}$ and $\mathbf{A}\,\mathbf{e} = 0$, we have $\mathcal{M}(\mathbf{A}) \subseteq \mathcal{M}(\mathbf{B})$ and $\mathbf{B}^+ \, \mathbf{A} = \mathbf{A}$. Hence, the largest eigenvalue of

B^+A equals the largest eigenvalue of A which in turn equals the largest eigenvalue of WW'. Therefore $\lambda_1 = 1$ for this choice of B and A. Hence by Theorem 2.2.2, $B - A$ is a positive semidefinite matrix. Thus, we get

$$\mathbf{y}'\mathbf{W}'\mathbf{W}\mathbf{y} \leq \mathbf{y}'(\mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}')\mathbf{y} \quad \textit{for all } \mathbf{y} \in R^n. \tag{2.3.11}$$

Therefore for any $m \times n$ matrix $\mathbf{Z}$, the matrix

$$\mathbf{Z}(\mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}')\mathbf{Z}' - \mathbf{Z}\mathbf{W}'\mathbf{W}\mathbf{Z}' \tag{2.3.12}$$

is positive semidefinite. Hence inequality (3.6) in Olkin (1992) holds.

Application 2.3.2 Our second application deals with multiple comparison procedures in linear models. One of the most important problems in multiple comparisons is the problem of construction of simultaneous confidence intervals for a given set of estimable functions. Among the methods available, Scheffé's technique has been the most popular and widely used method for the construction of simultaneous confidence intervals. A very nice description of Scheffé's S-method can be found in Seber (1977), page 128. In many texts the S-method is usually described by assuming that the design matrix is of full rank and the set of given estimable functions are linearly independent. However, this is rarely the case in practice. As an important application of the results of Section 2.2, we show that Scheffé's S-method can be extended to the case where the design matrix is not of full rank and the set of estimable functions are linearly dependent.

Consider the linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ where $\mathbf{y}$ is an $n \times 1$ vector of observations, $\boldsymbol{\beta}$ is a $p \times 1$ vector of parameters, $\mathbf{X}$ is a design matrix of order $n \times p$ and $\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of random errors. Let us assume that $\boldsymbol{\varepsilon}$ is distributed as multivariate normal with mean $\mathbf{0}$ and covariance matrix $\sigma^2\mathbf{I}$. Assume that the rank of $\mathbf{X}$ is r, where $r < p$. Consider s estimable functions $\mathbf{K}'\boldsymbol{\beta}$ where $\mathbf{K}_{p\times s}$ is a matrix of rank $q < s$. It is well known that the condition of estimability is equivalent to $\mathcal{M}(\mathbf{K}) \subseteq \mathcal{M}(\mathbf{X}'\mathbf{X})$. Let $\mathbf{G}$ be a g-inverse of $\mathbf{X}'\mathbf{X}$ and $(n-r)\,\hat{\sigma^2} = \mathbf{y}'\,(\mathbf{I} - \mathbf{X}\,\mathbf{G}\,\mathbf{X}')\,\mathbf{y}$. From Theorem 4.6 of Seber (1977), it follows that the statistic

$$F = \frac{(\mathbf{K}'\hat{\boldsymbol{\beta}} - \mathbf{K}'\boldsymbol{\beta})'\,(\mathbf{K}'\,\mathbf{G}\,\mathbf{K})^-\,(\mathbf{K}'\hat{\boldsymbol{\beta}} - \mathbf{K}'\boldsymbol{\beta})/q}{\hat{\sigma^2}} \qquad (2.3.13)$$

has an F-distribution with q and $(n-r)$ degrees of freedom, where $\hat{\boldsymbol{\beta}}$ is any solution to the equation $\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \mathbf{X}'\mathbf{y}$. Let $F^{\alpha}_{q,n-r}$ be the $100(1-\alpha)$ percentile of the F-distribution with q and $(n-r)$ degrees of freedom, then from (2.3.13), we have

$$\begin{aligned} 1-\alpha &= Pr\left(F \leq F^{\alpha}_{q,n-r}\right) \\ &= Pr\left((\mathbf{K}'\hat{\boldsymbol{\beta}} - \mathbf{K}'\boldsymbol{\beta})'\,(\mathbf{K}'\mathbf{G}\mathbf{K})^-\,(\mathbf{K}'\hat{\boldsymbol{\beta}} - \mathbf{K}'\boldsymbol{\beta}) \leq q\,\hat{\sigma^2}\,F^{\alpha}_{q,n-r}\right) \\ &= Pr\left(\,\mathbf{b}'\mathbf{B}^-\mathbf{b} \;\leq\; \eta\,\right) \qquad (2.3.14) \end{aligned}$$

where $\eta = q\hat{\sigma^2}\,F^{\alpha}_{q,n-r}$, $\mathbf{B} = \mathbf{K}'\,\mathbf{G}\,\mathbf{K}$ and $\mathbf{b} = \mathbf{K}'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$. Note that $\mathbf{b} \in \mathcal{M}(\mathbf{B})$ since $\mathbf{b} \in \mathcal{M}(\mathbf{K}')$; and from Lemma 2.3.1, we have $\mathcal{M}(\mathbf{K}') = \mathcal{M}(\mathbf{K}'\,\mathbf{G}\,\mathbf{K})$. Since $\mathbf{B}$ is a positive semidefinite matrix and $\mathbf{b} \in \mathcal{M}(\mathbf{B})$ by Corollary 2.2.1, we have that (2.3.14) is equivalent to

$$\begin{aligned} 1-\alpha &= Pr(\ \mathbf{h}'\mathbf{b}\mathbf{b}'\mathbf{h} \ \leq \ \eta\, \mathbf{h}'\mathbf{B}\,\mathbf{h} \quad \text{for all } \mathbf{h}\ ) \\ &= Pr\left(|\mathbf{h}'(\mathbf{K}'\hat{\boldsymbol{\beta}} - \mathbf{K}'\boldsymbol{\beta})| \ \leq \ \sqrt{\eta\, \mathbf{h}'\mathbf{B}\,\mathbf{h}} \quad \text{for all } \mathbf{h} \right). \qquad (2.3.15) \end{aligned}$$

Therefore, we have simultaneous confidence intervals for any linear function $\mathbf{h}'(\mathbf{K}'\boldsymbol{\beta})$ of the estimable functions $\mathbf{K}'\boldsymbol{\beta}$, namely,

$$\mathbf{h}'(\mathbf{K}'\hat{\boldsymbol{\beta}}) \pm (qF^{\alpha}_{q,n-r})^{1/2}\, \hat{\sigma}\, \sqrt{\mathbf{h}'(\mathbf{K}'\mathbf{G}\mathbf{K})\mathbf{h}} \qquad (2.3.16)$$

such that the overall probability for the whole class of such intervals is equal to $(1-\alpha)$.

The following lemma was used in Application 2.3.2.

Lemma 2.3.1 *Let* $\mathbf{K}$ *and* $\mathbf{X}$ *be as given in Application* 2.3.2. *Suppose that* $\mathcal{M}(\mathbf{K}) \subseteq \mathcal{M}(\mathbf{X}'\mathbf{X})$. *Let* $\mathbf{G}$ *be a g-inverse of* $\mathbf{X}'\mathbf{X}$ *then* $\mathcal{M}(\mathbf{K}') = \mathcal{M}(\mathbf{K}'\mathbf{G}\mathbf{K})$.

Proof: Clearly, $\mathcal{M}(\mathbf{K}'\mathbf{G}\mathbf{K}) \subseteq \mathcal{M}(\mathbf{K}')$. Let $\mathbf{G}$ be a g-inverse of $\mathbf{X}'\mathbf{X}$. Since $\mathcal{M}(\mathbf{K}) \subseteq \mathcal{M}(\mathbf{X}'\mathbf{X})$, we can write $\mathbf{K} = (\mathbf{X}'\mathbf{X})\mathbf{D}$ for some matrix $\mathbf{D}$. Therefore, the rank of $\mathbf{K}'\mathbf{G}\mathbf{K}$ is the same as that of $\mathbf{D}'(\mathbf{X}'\mathbf{X})\mathbf{D}$ which in turn equals the rank of $\mathbf{D}'\mathbf{X}'$. Thus, we have

$$\text{rank of } (\mathbf{K}'\mathbf{G}\mathbf{K}) \ = \ \text{rank of } (\mathbf{D}'\mathbf{X}')$$

$$\geq \quad \text{rank of } (\mathbf{K}'). \tag{2.3.17}$$

Since the other inequality always holds, we have that rank of $\mathbf{K}'\,\mathbf{G}\,\mathbf{K}$ equals rank of $\mathbf{K}'$. This completes the proof of the lemma.

Chapter 3

Invariance Properties of Certain Univariate Test Statistics

3.1 Introduction

The basic results of this dissertation are contained in this chapter. The organization of this chapter is as follows: In Section 3.2, we obtain an elegant characterization of the class of all n.n.d. matrices $\mathbf{W}$ satisfying the matrix equation $\mathbf{A}\mathbf{W}\mathbf{A} = \mathbf{B}$ where $\mathbf{A}$ is any symmetric matrix and $\mathbf{B}$ is an n.n.d. matrix. This result is contained in Theorem 3.2.2. Special cases of this theorem which are of statistical importance are presented in Theorems 3.2.1 and 3.2.3. The above mentioned matrix equation when restricted to $\mathbf{B} = \mathbf{A}$ occurs frequently in the study of the chi-squaredness of quadratic forms. In this case, the class of all n.n.d. $\mathbf{W}$'s is simply the class of all n.n.d. g-inverses of the matrix $\mathbf{A}$.

The statistical applications of our matrix theoretic results are presented in Section 3.3. For example, we show that the two sample t-statistic has a t-distribution if the observations in one of the samples are positively equicorrelated and those in the other sample are negatively equicorrelated with the same correlation in absolute value. Other applications include characterization of covariance matrices such that the independence and chi-squaredness of the quadratic forms occurring in ANOVA problems are preserved. We present a complete solution to some of the problems considered by Smith and Lewis (1980), Pavur and Lewis (1983), Bhat (1962) in Theorem 2 and by Scariano and Davenport (1987) in Theorem 1.

3.2 Results in Linear Algebra

In this section, we obtain a characterization of the class of all n.n.d. solutions to a general matrix equation useful in statistics. To motivate the proof of our main Theorem 3.2.2, we first consider the following special case in which we characterize the class of all n.n.d. g-inverses of the centering matrix, $\mathbf{A}^* = \mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}'$. The centering matrix is of great importance in statistics because it is used in the representation of the sample variance as a quadratic form.

Theorem 3.2.1 *The class of all n.n.d. g-inverses, $\mathcal{G}_n$, of the centering matrix $\mathbf{A}^*$ is given by*

$$\mathbf{W} = \mathbf{A}^* + \frac{1}{n}(\mathbf{e}\,\mathbf{a}' + \mathbf{a}\,\mathbf{e}') - \frac{\overline{a}}{n}\mathbf{e}\,\mathbf{e}' \tag{3.2.1}$$

where $\mathbf{a}' = (a_1, \ldots, a_n)$ is such that

$$\frac{1}{n}\sum_{i=1}^{n}(a_i - \overline{a})^2 \leq \overline{a} \tag{3.2.2}$$

and $\overline{a} = (\mathbf{a}'\,\mathbf{e})/n$ is the mean of the components of the vector $\mathbf{a}$.

Proof: From the result 1b.5.2 of Rao (1973), it is easy to verify that a general solution to the equation $\mathbf{A}^* \mathbf{W} \mathbf{A}^* = \mathbf{A}^*$ is given by

$$\mathbf{W} = \mathbf{A}^* + \frac{1}{n}(\mathbf{e}\,\mathbf{a}' + \mathbf{a}\,\mathbf{e}') - \frac{\overline{a}}{n}\mathbf{e}\,\mathbf{e}' \tag{3.2.3}$$

where $\mathbf{a} \in \Re^n$. Now, $\mathbf{W}$ is n.n.d. if and only if $\mathbf{x}'\mathbf{W}\mathbf{x} \geq 0$ for all $\mathbf{x} \in \Re^n$. Let $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ be the orthogonal decomposition of $\mathbf{x}$ where $\mathbf{x}_1 \in \mathcal{M}(\mathbf{A}^*)$ and $\mathbf{x}_2 \in \mathcal{M}(\mathbf{e})$. Since $\mathbf{x}_2 = d\,\mathbf{e}$ for some d, it is easy to verify that $\mathbf{W}$ is n.n.d. if and only if

$$n\overline{a}\,d^2 + 2d\,\mathbf{a}'\mathbf{x}_1 + \mathbf{x}_1'\,\mathbf{A}^*\mathbf{x}_1 \geq 0 \tag{3.2.4}$$

for all d and for all $\mathbf{x}_1 \in \mathcal{M}(\mathbf{A}^*)$. Since the left hand side of (3.2.4) is a quadratic equation in d, the inequality (3.2.4) holds if and only if

$$(\mathbf{a}'\mathbf{x}_1)^2 - (\mathbf{x}_1'\,\mathbf{A}^*\mathbf{x}_1)(n\overline{a}) \leq 0 \tag{3.2.5}$$

for all $\mathbf{x}_1 \in \mathcal{M}(\mathbf{A}^*)$ and $\overline{a} \geq 0$. Since $\mathbf{A}^*$ is the Moore-Penrose inverse of itself, it follows from Theorem 2.2.3 that the inequality (3.2.5) holds if and only if

$$(\mathbf{a}'\mathbf{A}^*\mathbf{a}) \leq n\overline{a} \tag{3.2.6}$$

which is equivalent to the condition (3.2.2). This completes the proof of the theorem.

Remark 3.2.1 The representation of the class $\mathcal{G}_n$, by matrices $\mathbf{W}$ given by (3.2.1) is minimal in the sense that each set of $\{a_1, \ldots, a_n\}$ satisfying (3.2.2) uniquely determines an n.n.d. g-inverse of the centering matrix $\mathbf{A}^*$. It is easy to show that $(n-2)$

of the eigenvalues of $\mathbf{W}$ given by (3.2.1) are ones and the other two eigenvalues are the solutions of the following quadratic equation

$$\lambda^2 - \lambda(1+\bar{a}) + (\bar{a} - \frac{1}{n}\sum_{i=1}^{n}(a_i - \bar{a})^2) = 0. \tag{3.2.7}$$

We can show that the two roots of the above equation (3.2.7) are nonnegative if and only if condition (3.2.2) is satisfied. Thus, we have an alternate proof of Theorem 3.2.1.

Remark 3.2.2 A different characterization of the class $\mathcal{G}_n$ was obtained by Jensen (1989c), however our characterization is elegant in the sense that it provides an easy method of generating a $\mathbf{W} \in \mathcal{G}_n$. All we need to do is to choose a set $\{a_1, \ldots, a_n\}$ of n numbers and if the inequality (3.2.2) is not satisfied by the a_i's then translate them by an appropriate constant. Recall, translation of a set of numbers changes the mean but not the variance.

Remark 3.2.3 Note that for $d > 0$, the class of all n.n.d. matrices $\mathbf{W}$ satisfying the equation $\mathbf{A}^* \mathbf{W} \mathbf{A}^* = d\,\mathbf{A}^*$ is simply given by $\mathcal{G}_{d,n} = \{d\mathbf{W} : \mathbf{W} \in \mathcal{G}_n\}$.

We now state a lemma that enables us to obtain a generalization of Theorem 3.2.1. The following lemma is a multivariate analogue of the simple problem of finding the restrictions on the coefficients of a quadratic equation $f(x) = ax^2 + 2bx + c$ such that $f(x) \geq 0$ for all x. It plays a crucial role in the proof of the main Theorem 3.2.2.

This lemma is also useful in quadratic programming problems where the objective is to minimize the multivariate quadratic loss function.

Lemma 3.2.1 *Let* $\mathbf{A}$ *be a symmetric matrix of order* $n \times n$ *and* $\mathbf{b}$ *be a vector in* $\Re^n$. *Let* c *be a real number. In order that*

$$f(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x} + 2\mathbf{x}'\mathbf{b} + c \geq 0 \quad \textit{for all} \quad \mathbf{x} \in \Re^n, \tag{3.2.8}$$

it is necessary and sufficient that (1) $\mathbf{b} \in \mathcal{M}(\mathbf{A})$, (2) $\mathbf{A}$ *is an n.n.d. matrix and* (3) $c - \mathbf{b}'\mathbf{A}^{-}\mathbf{b} \geq 0$. *In particular, if* $c = 0$ *then* (3.2.8) *holds if and only if* $\mathbf{b} = \mathbf{0}$.

Proof: We prove the lemma by first showing that (3.2.8) implies (1) and (2). Next we assume (1), (2) and show that (3.2.8) holds if and only if (3) holds. Let us write $\mathbf{b} = \mathbf{b}_1 + \mathbf{b}_2$ where $\mathbf{b}_1 \in \mathcal{M}(\mathbf{A})$ and $\mathbf{b}_2 \in \mathcal{N}(\mathbf{A})$. Choosing $\mathbf{x} = \alpha\,\mathbf{b}_2$ where α is a real number, we can see that (3.2.8) implies

$$2\alpha\,\mathbf{b}_2'\mathbf{b}_2 + c \geq 0 \quad \text{for all} \quad \alpha \in \Re, \tag{3.2.9}$$

which is true if and only if $\mathbf{b}_2 = \mathbf{0}$ and $c \geq 0$. Hence $\mathbf{b} = \mathbf{b}_1 \in \mathcal{M}(\mathbf{A})$. Now suppose that $\mathbf{A}$ is not an n.n.d. matrix. Hence there exists an eigenvalue λ of $\mathbf{A}$ such that $\lambda < 0$. Let $\mathbf{u}$ be the normalized eigenvector of $\mathbf{A}$ corresponding to λ. Choosing $\mathbf{x} = \beta\,\mathbf{u}$ where β is a real number, we can see that (3.2.8) implies

$$\lambda\beta^2 + 2\beta\,\mathbf{b}'\mathbf{u} + c \geq 0 \quad \text{for all} \quad \beta \in \Re, \tag{3.2.10}$$

which is a contradiction since $\lambda < 0$. Thus (3.2.8) implies that $\mathbf{A}$ is n.n.d. Let us now assume that (1) and (2) hold. Taking derivative of the function $f(\mathbf{x})$ in (3.2.8) with respect to $\mathbf{x}$ and equating to zero, we get

$$\mathbf{A}\,\mathbf{x} = -\mathbf{b} \tag{3.2.11}$$

which is a consistent equation since $\mathbf{b} \in \mathcal{M}(\mathbf{A})$. Therefore, $\mathbf{x}_0 = -\mathbf{A}^-\,\mathbf{b}$ is a point of minimum for the function $f(\mathbf{x})$ since $\mathbf{A}$ is an n.n.d. matrix. Thus

$$\min_{\mathbf{x}} f(\mathbf{x}) = f(\mathbf{x}_0) = c - \mathbf{b}'\,\mathbf{A}^-\,\mathbf{b}; \tag{3.2.12}$$

hence (3.2.8) holds if and only if $c - \mathbf{b}'\,\mathbf{A}^-\,\mathbf{b} \geq 0$. If $c = 0$, then (3.2.8) holds if and only if $\mathbf{b} \in \mathcal{M}(\mathbf{A})$ and $\mathbf{b}'\,\mathbf{A}^-\,\mathbf{b} = 0$. It is easy to see that these two conditions are equivalent to $\mathbf{b} = \mathbf{0}$. This completes the proof of the lemma.

We now state the main theorem of this section.

Theorem 3.2.2 *Let* $\mathbf{A}$ *be a symmetric matrix of order* $n \times n$ *and* $\mathbf{B}$ *be an n.n.d. matrix of order* $n \times n$ *such that*

$$\mathbf{A}\,\mathbf{W}\,\mathbf{A} = \mathbf{B} \tag{3.2.13}$$

is a consistent equation. Let $\mathbf{J} = (\mathbf{I} - \mathbf{A}^+\,\mathbf{A})$. *Then the class of all n.n.d.* $\mathbf{W}$*'s satisfying* (3.2.13) *is given by*

$$\mathbf{W} = \mathbf{A}^+\,\mathbf{B}\,\mathbf{A}^+ + \mathbf{J}\,\mathbf{C} + \mathbf{C}\,\mathbf{J} - \mathbf{J}\,\mathbf{C}\,\mathbf{J} \tag{3.2.14}$$

where $\mathbf{C}$ *is a symmetric matrix such that* $\mathbf{J}\,\mathbf{C}\,\mathbf{J}$ *is nonnull and it satisfies the following two conditions:*

(a) $\mathcal{M}(\mathbf{A\,C\,J}) \subseteq \mathcal{M}(\mathbf{B})$

(b) $\mathbf{D} \stackrel{\text{def}}{=} \mathbf{J\,C\,J} - \mathbf{J\,C\,A\,B^{-}\,A\,C\,J}$ *is an n.n.d. matrix.*

If $\mathbf{J\,C\,J}$ *is a null matrix, then* $\mathbf{W}$ *given by* (3.2.14) *is an n.n.d. solution for* (3.2.13) *if and only if* $\mathbf{C\,J} = \mathbf{O}$.

Proof: Let $\mathbf{A}$ be a symmetric matrix and $\mathbf{B}$ be an n.n.d. matrix. We first note that by Theorem 2.3.2 of Rao and Mitra (1971), the equation (3.2.13) is consistent if and only if $\mathbf{A\,A^{-}\,B\,A^{-}\,A} = \mathbf{B}$ for any g-inverse $\mathbf{A}^{-}$ of $\mathbf{A}$, in which case the general solution is given by (3.2.14) where $\mathbf{C}$ is an arbitrary matrix. Therefore, our problem reduces to characterizing the class of all $\mathbf{C}$'s such that $\mathbf{W}$ given by (3.2.14) is symmetric and n.n.d. matrix. Without loss of generality we can take $\mathbf{C}$ to be symmetric if not, we can replace $\mathbf{C}$ by $\mathbf{C}^* = (\mathbf{C} + \mathbf{C}')/2$. Note that $\mathbf{C}$ and $\mathbf{C}^*$ generate the same $\mathbf{W}$. We can rewrite the matrix $\mathbf{W}$ in (3.2.14) as

$$\begin{aligned} \mathbf{W} &= \mathbf{A^{+}\,B\,A^{+}} + \mathbf{J\,C} + \mathbf{C\,J} - \mathbf{J\,C\,J} \\ &= \mathbf{A^{+}\,B\,A^{+}} + \mathbf{J\,C\,A^{+}A} + \mathbf{A^{+}\,A\,C\,J} + \mathbf{J\,C\,J}. \end{aligned} \tag{3.2.15}$$

Since $\mathbf{A}$ is symmetric we have $\mathbf{A^{+}\,A} = \mathbf{A\,A^{+}}$ and $\mathcal{M}(\mathbf{A^{+}\,A}) = \mathcal{M}(\mathbf{A})$. Let $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ be the orthogonal decomposition of $\mathbf{x}$ where $\mathbf{x}_1 \in \mathcal{M}(\mathbf{A^{+}\,A})$ and $\mathbf{x}_2 \in \mathcal{M}(\mathbf{J})$. It is easy to see that

$$\mathbf{x}'\,\mathbf{W}\,\mathbf{x} \geq 0, \quad \forall \mathbf{x} \in \Re^n \tag{3.2.16}$$

if and only if

$$\mathbf{x}_1' \mathbf{A}^+ \mathbf{B} \mathbf{A}^+ \mathbf{x}_1 + 2\mathbf{x}_1' \mathbf{C} \mathbf{x}_2 + \mathbf{x}_2' \mathbf{C} \mathbf{x}_2 \geq 0, \ \forall \mathbf{x}_1 \in \mathcal{M}(\mathbf{A}) \text{ and } \forall \mathbf{x}_2 \in \mathcal{M}(\mathbf{J}). \tag{3.2.17}$$

Since (3.2.13) is a consistent equation we have $\mathbf{A}\,\mathbf{A}^+\,\mathbf{B}\,\mathbf{A}^+\,\mathbf{A} = \mathbf{B}$. Therefore, from (3.2.17) $\mathbf{W}$ is n.n.d. if and only if

$$\mathbf{v}'\mathbf{B}\,\mathbf{v} + 2\,\mathbf{v}'\mathbf{A}\,\mathbf{C}\,\mathbf{J}\,\mathbf{w} + \mathbf{w}'\mathbf{J}\,\mathbf{C}\,\mathbf{J}\,\mathbf{w} \geq 0, \quad \forall \mathbf{v} \in \Re^n \text{ and } \forall \mathbf{w} \in \Re^n. \tag{3.2.18}$$

If $\mathbf{J}\,\mathbf{C}\,\mathbf{J}$ is a nonnull matrix then by Lemma 3.2.1 it follows that (3.2.18) holds if and only if the following two conditions are satisfied:

(1) $\mathbf{A}\,\mathbf{C}\,\mathbf{J}\,\mathbf{w} \in \mathcal{M}(\mathbf{B}), \quad \forall \ \mathbf{w} \in \Re^n$

(2) $\mathbf{w}'\mathbf{J}\,\mathbf{C}\,\mathbf{J}\,\mathbf{w} - \mathbf{w}'\mathbf{J}\,\mathbf{C}\,\mathbf{A}\,\mathbf{B}^-\mathbf{A}\,\mathbf{C}\,\mathbf{J}\,\mathbf{w} \geq 0, \quad \forall \ \mathbf{w} \in \Re^n.$

It is easy to see that (1) is equivalent to (a) and condition (2) is equivalent to (b). If $\mathbf{J}\,\mathbf{C}\,\mathbf{J}$ is a null matrix then by (3.2.18) and Lemma 3.2.1 $\mathbf{W}$ is n.n.d. if and only if $\mathbf{A}\,\mathbf{C}\,\mathbf{J}\,\mathbf{w} = \mathbf{0}$ for all $\mathbf{w} \in \Re^n$ or equivalently, $\mathbf{A}\,\mathbf{C}\,\mathbf{J}$ is a null matrix. Hence $\mathbf{C}\,\mathbf{J} = \mathbf{C}\,\mathbf{J} - \mathbf{J}\,\mathbf{C}\,\mathbf{J} = (\mathbf{I} - \mathbf{J})\,\mathbf{C}\,\mathbf{J} = \mathbf{A}^+\,\mathbf{A}\mathbf{C}\,\mathbf{J} = \mathbf{O}$, since we have assumed that $\mathbf{J}\,\mathbf{C}\,\mathbf{J}$ is a null matrix. This proves the last assertion of the theorem.

Remark 3.2.4 If $\mathbf{J}\,\mathbf{C}\,\mathbf{J}$ is a nonnull matrix, it also follows from Lemma 3.2.1 that (3.2.18) holds if and only if the following three conditions are satisfied:

(c) $\mathcal{M}(\mathbf{J}\,\mathbf{C}\,\mathbf{A}) \subseteq \mathcal{M}(\mathbf{J}\,\mathbf{C}\,\mathbf{J})$

(d) $\mathbf{J\,C\,J}$ is an n.n.d. matrix

(e) $\mathbf{B} - \mathbf{A\,C\,J}\,(\mathbf{J\,C\,J})^{-}\,\mathbf{J\,C\,A}$ is an n.n.d. matrix.

Therefore conditions (a) and (b) are equivalent to conditions (c), (d) and (e).

Remark 3.2.5 Let $\mathbf{C}^* = ((\mathbf{J\,C\,J})^{+}\,\mathbf{J\,C\,A\,B}^{-}\,\mathbf{A\,C\,J})$ and $\lambda_1(\mathbf{C}^*)$ denotes the maximum eigenvalue of $\mathbf{C}^*$. If (c) and (d) hold then by Theorem 2.2.2, (b) is true if and only if

(f) $\lambda_1(\mathbf{C}^*) \leq 1$.

Therefore, (a) and (b) are equivalent to (a), (c), (d) and (f). We can thus replace conditions (a) and (b) in the statement of Theorem 3.2.2 either with (c), (d) and (e) or with (a), (c), (d) and (f).

Remark 3.2.6 The representation of the class of n.n.d. solutions $\mathbf{W}$ in Theorem 3.2.2 is minimal in the sense that two symmetric matrices $\mathbf{C}$ and $\mathbf{C}'$ generate the same $\mathbf{W}$ if and only if $\mathbf{J\,C} = \mathbf{J\,C}'$. Also, from the proof of the above theorem it follows that the class of all positive definite (positive semidefinite) $\mathbf{W}$'s satisfying (3.2.13) is obtained by choosing $\mathbf{C}$ such that $\mathbf{D}$ is positive definite (positive semidefinite) matrix. In the case where $\mathbf{J\,C} = \mathbf{O}$, the only n.n.d. solution $\mathbf{W} = \mathbf{A}^{+}\,\mathbf{B}\,\mathbf{A}^{+}$ is positive definite or positive semidefinite according as $\mathbf{B}$ is positive definite or positive semidefinite matrix.

An important application of Theorem 3.2.2 is the following characterization of the class of all n.n.d. g-inverses of an n.n.d. matrix $\mathbf{A}$. Note that condition (a) of Theorem 3.2.2 is trivially satisfied if $\mathbf{B} = \mathbf{A}$.

Example 3.2.1 Let $\mathbf{A}$ be an n.n.d. matrix of order $n \times n$ and $\mathbf{J} = (\mathbf{I} - \mathbf{A}^{+}\mathbf{A})$. Then the class of all n.n.d. g-inverses of the matrix $\mathbf{A}$ is given by

$$\mathbf{W} = \mathbf{A}^{+} + \mathbf{J}\,\mathbf{C} + \mathbf{C}\,\mathbf{J} - \mathbf{J}\,\mathbf{C}\,\mathbf{J} \tag{3.2.19}$$

where $\mathbf{C}$ is a symmetric matrix such that $\mathbf{J}\,\mathbf{C}\,\mathbf{J}$ is nonnull and

$$\mathbf{D} = \mathbf{J}\,\mathbf{C}\,\mathbf{J} - \mathbf{J}\,\mathbf{C}\,\mathbf{A}\,\mathbf{C}\,\mathbf{J} \tag{3.2.20}$$

is an n.n.d. matrix.

Let us now look at the case where $\mathbf{A}$ and $\mathbf{B}$ can be expressed as linear combinations of k orthogonal and idempotent matrices. We first state a lemma which is needed in the proofs of Theorems 3.2.3 and 3.3.4.

Lemma 3.2.2 *Let $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$ be symmetric and idempotent matrices of order $n \times n$ such that $\mathbf{A}_i\mathbf{A}_j = \mathbf{O}$ for all $i \neq j$. Let $\mathbf{A} = \sum_{i=1}^{k} \mathbf{A}_i$ and $\mathbf{B} = \sum_{i=1}^{k} c_i\mathbf{A}_i$ where $c_i > 0$ for $1 \leq i \leq k$. Then $\mathbf{A}\,\mathbf{W}\,\mathbf{A} = \mathbf{B}$ if and only if*

$$\mathbf{A}_i\,\mathbf{W}\,\mathbf{A}_j = \begin{cases} c_i\,\mathbf{A}_i & \text{if } i = j \\ \mathbf{O} & \text{if } i \neq j. \end{cases} \tag{3.2.21}$$

Further, $\mathcal{M}(\mathbf{A}) = \mathcal{M}(\mathbf{B})$.

Proof: The proof of the first assertion is easy. We proceed to show $\mathcal{M}(\mathbf{A}) = \mathcal{M}(\mathbf{B})$. Since $\mathbf{A}\,\mathbf{B} = \mathbf{B}$, we have $\mathcal{M}(\mathbf{B}) \subseteq \mathcal{M}(\mathbf{A})$. Suppose $\mathbf{x} = \mathbf{A}\,\mathbf{v}$ and let $\mathbf{v}^* = \sum_{i=1}^{k} \frac{1}{c_i} \mathbf{A}_i\,\mathbf{v}$. It is easy to verify that $\mathbf{x} = \mathbf{B}\,\mathbf{v}^*$ and hence $\mathcal{M}(\mathbf{A}) \subseteq \mathcal{M}(\mathbf{B})$. This completes the proof of the lemma.

The following special case of Theorem 3.2.2 is useful to study the invariance properties of common statistical tests for dependent observations.

Theorem 3.2.3 *Let* $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$ *and* $\mathbf{B}$ *be as in Lemma* 3.2.2. *Suppose that* $\sum_{i=1}^{k} \mathbf{A}_i = \mathbf{A}^*$ *where* $\mathbf{A}^*$ *is the centering matrix. Then the class of all n.n.d. matrix solutions for the equation*

$$\mathbf{A}^*\,\mathbf{W}\,\mathbf{A}^* = \mathbf{B} \tag{3.2.22}$$

is given by

$$\mathbf{W} = \mathbf{B} + \frac{1}{n}(\mathbf{e}\,\mathbf{a}' + \mathbf{a}\,\mathbf{e}') - \frac{\bar{a}}{n}\mathbf{e}\,\mathbf{e}' \tag{3.2.23}$$

where $\mathbf{a}$ *is an arbitrary vector satisfying*

$$\frac{1}{n}\sum_{i=1}^{k} \frac{\mathbf{a}'\,\mathbf{A}_i\,\mathbf{a}}{c_i} \leq \bar{a}. \tag{3.2.24}$$

Proof: We prove the theorem by simply verifying the conditions of Theorem 3.2.2. It is easy to check that $\mathbf{A}^*\,\mathbf{B}\,\mathbf{A}^* = \mathbf{B}$ and hence the equation (3.2.22) is consistent. Since $\mathbf{A}^*$ is idempotent, we have $(\mathbf{A}^*)^+ = \mathbf{A}^*$ and $\mathbf{J} = (\mathbf{I} - \mathbf{A}^*) = \frac{1}{n}\mathbf{e}\,\mathbf{e}'$. Let $\mathbf{C}$ be a symmetric matrix and $\mathbf{a}' = \mathbf{e}'\,\mathbf{C}$. Note that $\mathbf{J}\,\mathbf{C}\,\mathbf{J} = \frac{\mathbf{a}'\,\mathbf{e}}{n^2}\,\mathbf{e}\,\mathbf{e}' = \frac{\overline{a}}{n}\mathbf{e}\,\mathbf{e}'$. The general solution to the equation (3.2.22) is given by

$$\begin{aligned} \mathbf{W} &= \mathbf{B} + \mathbf{J}\,\mathbf{C} + \mathbf{C}\,\mathbf{J} - \mathbf{J}\,\mathbf{C}\,\mathbf{J} \\ &= \mathbf{B} + \frac{1}{n}(\mathbf{e}\,\mathbf{a}' + \mathbf{a}\,\mathbf{e}') - \frac{\overline{a}}{n}\mathbf{e}\,\mathbf{e}'. \end{aligned} \tag{3.2.25}$$

It remains to show that $\mathbf{W}$ given by (3.2.25) is n.n.d. if and only if the vector $\mathbf{a}$ satisfies inequality (3.2.24). By Lemma 3.2.2, we have $\mathcal{M}(\mathbf{A}^*) = \mathcal{M}(\mathbf{B})$ and therefore condition (a) of Theorem 3.2.2 is trivially satisfied. Let $\mathbf{D}$ be as defined in Theorem 3.2.2. It is easy to verify that $\mathbf{B}^- = \sum_{i=1}^{k} \frac{1}{c_i}\mathbf{A}_i$, $\mathbf{A}^*\,\mathbf{B}^-\,\mathbf{A}^* = \mathbf{B}^-$ and

$$\begin{aligned} \mathbf{D} &= \mathbf{J}\,\mathbf{C}\,\mathbf{J} - \mathbf{J}\,\mathbf{C}\,\mathbf{B}^-\,\mathbf{C}\,\mathbf{J} \\ &= \frac{\mathbf{a}'\,\mathbf{e}}{n^2}\mathbf{e}\,\mathbf{e}' - \frac{\mathbf{a}'\,\mathbf{B}^-\,\mathbf{a}}{n^2}\;\mathbf{e}\,\mathbf{e}.' \end{aligned} \tag{3.2.26}$$

Since $\mathbf{e}\,\mathbf{e}'$ is an n.n.d. matrix, from (3.2.26) we can see that $\mathbf{D}$ is n.n.d. if and only if inequality (3.2.24) is satisfied. This completes the proof of the theorem.

3.3 Statistical Applications

In this section, we present some statistical applications of the theorems of Section 3.2. These applications are concerned with the problem of characterizing the class of all covariance matrices such that the distributions of common test statistics remain invariant, that is, the distributions are preserved except for a scale factor. We begin with the following lemma regarding the distribution of quadratic forms in normal variates with n.n.d. covariance matrix.

Lemma 3.3.1 *Let* $\mathbf{z} \sim N_n(\boldsymbol{\mu}, \mathbf{W})$ *where* $\mathbf{W}$ *is an n.n.d. matrix. Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two symmetric matrices of order* $n \times n$ *and* $\mathbf{a}$ *be a vector in* $\Re^n$. *Then*

(1) $\mathbf{z}'\mathbf{A}\,\mathbf{z} \sim \chi^2(r(\mathbf{A}); \boldsymbol{\mu}'\mathbf{A}\,\boldsymbol{\mu})$ *if and only if* $\mathbf{A}\mathbf{W}\mathbf{A} = \mathbf{A}$.

(2) *If* $\mathbf{A}$ *and* $\mathbf{B}$ *are n.n.d. matrices then* $\mathbf{z}'\mathbf{A}\,\mathbf{z}$ *and* $\mathbf{z}'\mathbf{B}\,\mathbf{z}$ *are independent if and only if* $\mathbf{A}\mathbf{W}\mathbf{B} = \mathbf{O}$.

(3) $\mathbf{z}'\mathbf{a}$ *and* $\mathbf{z}'\mathbf{B}\,\mathbf{z}$ *are independent if and only if* $\mathbf{B}\mathbf{W}\mathbf{a} = \mathbf{0}$ *where* $\mathbf{B}$ *is an n.n.d. matrix.*

Proof: On using Corollary 2s.1 of Searle (1971) and Lemma 4.6.1 for $p = 1$, we get (1). Result (2) follows from Theorem 4s of Searle (1971) and Lemma 4.6.2 with $p = 1$. Finally, Result (3) follows from (2).

The following examples are simple consequences of Lemma 3.3.1 and the results of the previous section.

Example 3.3.1 Let $\mathbf{z} \sim N_n(\boldsymbol{\mu}, \mathbf{W})$ where $\mathbf{W}$ is an n.n.d. matrix. Let $\mathbf{A}$ be an n.n.d. matrix of order $n \times n$. By Lemma 3.3.1 (1), we have

$$\mathbf{z}'\mathbf{A}\,\mathbf{z} \sim \chi^2(r(\mathbf{A}); \boldsymbol{\mu}'\mathbf{A}\,\boldsymbol{\mu}) \tag{3.3.1}$$

if and only if $\mathbf{A}\,\mathbf{W}\,\mathbf{A} = \mathbf{A}$. Therefore, for a given n.n.d. matrix $\mathbf{A}$, the class of all n.n.d. $\mathbf{W}$'s for which (3.3.1) holds is given by Example 3.2.1.

Example 3.3.2 In Example 3.3.1, $\mathbf{z}'\,\mathbf{A}\,\mathbf{z} \sim d\,\chi^2(r(\mathbf{A}); \delta)$ if and only if $\mathbf{A}\,\mathbf{W}\,\mathbf{A} = d\,\mathbf{A}$ where $d > 0$ and $\delta = \frac{1}{d}\,\boldsymbol{\mu}'\,\mathbf{A}\,\boldsymbol{\mu}$. Thus, we can use Theorem 3.2.2 to obtain a complete characterization of the covariance matrices $\mathbf{W}$ such that $\mathbf{y}'\,\mathbf{A}\,\mathbf{y} \sim d\,\chi^2(r(\mathbf{A}); \delta)$.

Example 3.3.3 Let $\mathbf{y} \sim N_n(\mu\,\mathbf{e}, \mathbf{W})$ where μ is a constant and $\mathbf{W}$ is an n.n.d. matrix. Let s^2 be the sample variance of the vector $\mathbf{y}$. Then, for any $d > 0$, we have $(n-1)\,s^2 \sim d\,\chi^2(n-1)$ if and only if $\mathbf{W} \in \mathcal{G}_{d,n}$.

Theorem 3.3.1 *Let $\mathbf{y}_1 \sim N_{n_1}(\mu_1\,\mathbf{e}_{n_1}, \mathbf{W}_1)$ and $\mathbf{y}_2 \sim N_{n_2}(\mu_2\,\mathbf{e}_{n_2}, \mathbf{W}_2)$ where μ_1, μ_2 are constants and $\mathbf{W}_1$, $\mathbf{W}_2$ are n.n.d. matrices. Assume that $\mathbf{y}_1$ is independent of $\mathbf{y}_2$. Let s_1^2 and s_2^2 be the sample variances of the vectors $\mathbf{y}_1$ and $\mathbf{y}_2$, respectively. Then for $c > 0$, we have s_1^2/s_2^2 is distributed as $c\,F(n_1-1,\, n_2-1)$ if and only if $\mathbf{W}_1 \in \mathcal{G}_{cd,n_1}$ and $\mathbf{W}_2 \in \mathcal{G}_{d,n_2}$ for some constant $d > 0$.*

Proof: Since s_1^2 and s_2^2 are independent, it follows from a result of Baldessari (1965) that s_1^2/s_2^2 is distributed as $c\,F(n_1-1,\, n_2-1)$ if and only if $(n_1-1)\,s_1^2 \sim cd\,\chi^2(n_1-1)$ and $(n_2-1)\,s_2^2 \sim d\,\chi^2(n_2-1)$. The theorem now follows from Example 3.3.3.

We now characterize the class of all covariance matrices such that the sample mean and variance are independent for a normal sample of dependent observations. Theorem 3.3.2 shows that the sample variance is distributed as chi-square except for a scale adjustment and is independent of the sample mean if and only if the observations are equicorrelated, that is, the correlation is same between each pair of observations.

Theorem 3.3.2 *Let* $\mathbf{y} \sim N_n(\mu\,\mathbf{e},\, \mathbf{W})$ *where* μ *is a constant and* $\mathbf{W}$ *is an n.n.d. matrix. Let* $\bar{y}$ *and* s^2 *be the mean and variance of the vector* $\mathbf{y}$*. Then* $(n-1)\,s^2 \sim d\,\chi^2(n-1)$ *and* $\bar{y}$ *is independent of* s^2 *if and only if* $\mathbf{W} = d\,(\mathbf{I} - \frac{(1-c)}{n}\,\mathbf{e}\,\mathbf{e}')$ *for some* $c \geq 0$ *and* $d > 0$.

Proof: For any $d > 0$, by Lemma 3.3.3, we have $(n-1)\,s^2 \sim d\,\chi^2(n-1)$ if and only if $\mathbf{W} \in \mathcal{G}_{d,n}$. From Lemma 3.3.1 (3), $\bar{y}$ and s^2 are independent if and only if

$$(\mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}')\,\mathbf{W}\,\mathbf{e} = \mathbf{0}. \tag{3.3.2}$$

It is easy to check that $\mathbf{W} \in \mathcal{G}_{d,n}$ and satisfies (3.3.2) if and only if

$$(\mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}')\,\mathbf{a} = \mathbf{0}. \tag{3.3.3}$$

Now, (3.3.3) holds if and only if $\mathbf{a} = c\,\mathbf{e}$ where $c = \bar{a} \geq 0$. Thus, $(n-1)\,s^2 \sim d\,\chi^2(n-1)$ and s^2 is independent of $\bar{y}$ if and only if

$$\mathbf{W} \;=\; d\left((\mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}') + \frac{1}{n}(\mathbf{e}\,\mathbf{a}' + \mathbf{a}\,\mathbf{e}') - \frac{\bar{a}}{n}\,\mathbf{e}\,\mathbf{e}'\right)$$

$$= d\left((\mathbf{I}-\frac{1}{n}\mathbf{e}\,\mathbf{e}') + \frac{2c}{n}\mathbf{e}\,\mathbf{e}' - \frac{c}{n}\mathbf{e}\,\mathbf{e}'\right)$$

$$= d\left(\mathbf{I} - \frac{(1-c)}{n}\mathbf{e}\,\mathbf{e}'\right) \tag{3.3.4}$$

where $c \geq 0$. This completes the proof of the theorem.

Our next result is concerned with an invariance property of the two sample t-test. Theorem 3.3.3 below shows that the commonly used two sample t-statistic has a t-distribution if one of the samples is positively equicorrelated and the other is negatively equicorrelated such that the correlation is same in absolute value in both the samples.

Theorem 3.3.3 *Let* $\mathbf{y}_1 \sim N_{n_1}(\mu_1\,\mathbf{e}_{n_1}, \mathbf{W}_1)$ *and* $\mathbf{y}_2 \sim N_{n_2}(\mu_2\,\mathbf{e}_{n_2}, \mathbf{W}_2)$ *where* μ_1, μ_2 *are constants and* $\mathbf{W}_1$, $\mathbf{W}_2$ *are n.n.d. matrices. Suppose that* $\mathbf{y}_1$ *and* $\mathbf{y}_2$ *are independently distributed. Let* $\overline{y}_1$, s_1^2 *and* $\overline{y}_2$, s_2^2 *be the mean and variance of the two vectors* $\mathbf{y}_1$ *and* $\mathbf{y}_2$, *respectively. Let* $s_p^2 = [(n_1-1)\,s_1^2 + (n_2-1)\,s_2^2]/(n_1+n_2-2)$ *be the pooled sample variance. Then*

$$\frac{(\overline{y}_1 - \overline{y}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \sim t(n_1 + n_2 - 2) \tag{3.3.5}$$

if

$$\mathbf{W}_1 = d\,(\mathbf{I}_{n_1} + \beta\,\mathbf{e}_{n_1}\,\mathbf{e}_{n_1}') \text{ and } \mathbf{W}_2 = d\,(\mathbf{I}_{n_2} - \beta\,\mathbf{e}_{n_2}\,\mathbf{e}_{n_2}') \tag{3.3.6}$$

for some constants β and d such that $d > 0$ and $-1/n_1 < \beta < 1/n_2$.

Proof: It follows from the proof of Theorem 3.3.2 that for any $d > 0$, $(n_i - 1)s_i^2$ is distributed as $d\chi^2(n_i - 1)$ and s_i^2 is independent of $\overline{y}_i$ if and only if

$$\mathbf{W}_i = d\left(\mathbf{I}_{n_i} - \frac{(1-c_i)}{n_i}\,\mathbf{e}_{n_i}\mathbf{e}_{n_i}'\right) \tag{3.3.7}$$

where $c_i \geq 0$ for $i = 1, 2$. Thus for $c_1, c_2 > 0$, we have

$$\frac{(\overline{y}_1 - \overline{y}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\dfrac{c_1}{n_1} + \dfrac{c_2}{n_2}}} \sim t(n_1 + n_2 - 2) \tag{3.3.8}$$

if $\mathbf{W}_i$'s are given by (3.3.7) for $i = 1, 2$. Now for (3.3.5) to hold we require

$$\frac{c_1}{n_1} + \frac{c_2}{n_2} = \frac{1}{n_1} + \frac{1}{n_2} \tag{3.3.9}$$

or equivalently,

$$\frac{-(1-c_1)}{n_1} = \frac{(1-c_2)}{n_2} = \beta \quad \text{(say)}. \tag{3.3.10}$$

Since $c_1, c_2 > 0$, we have $-1/n_1 < \beta < 1/n_2$. The theorem now follows from (3.3.7), (3.3.8) and (3.3.10).

We need the following version of Cochran's theorem for the distribution of quadratic forms in normal variates with n.n.d. covariance matrix. Theorem 3.3.4 is useful to obtain the invariance properties of the quadratic forms in ANOVA models.

Theorem 3.3.4 *Let $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$ be symmetric and idempotent matrices of order $n \times n$ such that $\mathbf{A}_i \mathbf{A}_j = \mathbf{O}$ for all $i \neq j$. Let $\mathbf{A} = \sum_{i=1}^{k} \mathbf{A}_i$ and $\mathbf{B} = \sum_{i=1}^{k} c_i \mathbf{A}_i$ where $c_i > 0$ for $1 \leq i \leq k$. Let $\mathbf{y} \sim N_n(\boldsymbol{\mu}, \mathbf{W})$ where $\mathbf{W}$ is an n.n.d. matrix. Let $Q_i = \mathbf{y}' \mathbf{A}_i \mathbf{y}$ for $1 \leq i \leq k$. Then the quadratic forms Q_i's are pairwise independent and distributed as $c_i \chi^2(r(\mathbf{A}_i); \delta_i)$ for $1 \leq i \leq k$ if and only if $\mathbf{A} \mathbf{W} \mathbf{A} = \mathbf{B}$ where $\delta_i = \frac{1}{c_i} \boldsymbol{\mu}' \mathbf{A}_i \boldsymbol{\mu}$.*

Proof: From Example 3.3.2 and Lemma 3.3.1 (2), Q_i's are pairwise independent and distributed as $c_i \chi^2(r(\mathbf{A}_i); \delta_i)$ for $1 \leq i \leq k$ if and only if

$$\mathbf{A}_i \mathbf{W} \mathbf{A}_j = \begin{cases} c_i \mathbf{A}_i & \text{if } i = j \\ \mathbf{O} & \text{if } i \neq j. \end{cases} \tag{3.3.11}$$

The theorem now follows from Lemma 3.2.2.

In the next theorem, we obtain invariance property of the distributions of quadratic forms in the ANOVA table.

Theorem 3.3.5 *Let $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$ be symmetric and idempotent matrices of order $n \times n$ such that $\mathbf{A}_i \mathbf{A}_j = \mathbf{O}$ for all $i \neq j$. Let $\sum_{i=1}^{k} \mathbf{A}_i = \mathbf{A}^*$ and $\mathbf{B} = \sum_{i=1}^{k} c_i \mathbf{A}_i$ where $\mathbf{A}^*$ is the centering matrix and $c_i > 0$ for $1 \leq i \leq k$. Let $\mathbf{y} \sim N_n(\boldsymbol{\mu}, \mathbf{W})$ where $\mathbf{W}$ is an n.n.d. matrix. Let $Q_i = \mathbf{y}' \mathbf{A}_i \mathbf{y}$ for $1 \leq i \leq k$. Then the quadratic forms Q_i's are pairwise independent and distributed as $c_i \chi^2(r(\mathbf{A}_i); \delta_i)$ for $1 \leq i \leq k$ if and only if $\mathbf{W}$ is of the form* (3.2.23) *where $\mathbf{a}$ is an arbitrary vector satisfying* (3.2.24) *and $\delta_i = \frac{1}{c_i} \boldsymbol{\mu}' \mathbf{A}_i \boldsymbol{\mu}$.*

Proof: From Theorem 3.3.4, we get that Q_i's are pairwise independent and distributed as $c_i\,\chi^2(r(\mathbf{A}_i);\delta_i)$ for $1 \leq i \leq k$ if and only if $\mathbf{A}^*\,\mathbf{W}\,\mathbf{A}^* = \mathbf{B}$. The desired result follows from Theorem 3.2.3.

Remark 3.3.1 In the above theorem, if $\boldsymbol{\mu} = \mu\,\mathbf{e}$ where μ is some constant then Q_i's are pairwise independent and distributed as $c_i\,\chi^2(r(\mathbf{A}_i))$ for $1 \leq i \leq k$ if and only if $\mathbf{A}^*\,\mathbf{W}\,\mathbf{A}^* = \mathbf{B}$.

As another simple application of the results of Section 3.2, we get the following characterization of the covariance matrices such that the null distribution of the quadratic forms in one way ANOVA remains invariant.

Theorem 3.3.6 *Consider the one way ANOVA model*

$$y_{ij} = \mu_i + \varepsilon_{ij}, \qquad j = 1, \ldots, n_i \;\; and \;\; i = 1, \ldots, g. \tag{3.3.12}$$

Let $\varepsilon' = (\varepsilon_{11}, \ldots, \varepsilon_{1n_1}, \ldots, \varepsilon_{g1}, \ldots, \varepsilon_{gn_g})$ *and* $n = \sum_{i=1}^{g} n_i$. *Assume that* $\varepsilon \sim N_n(\mathbf{0}, \mathbf{W})$ *where* $\mathbf{W}$ *is an n.n.d. matrix. Let* $\overline{y}_{i\cdot} = \sum_{j=1}^{n_i} y_{ij}/n_i$ *and* $\overline{y}_{\cdot\cdot} = \sum_{i=1}^{g}\sum_{j=1}^{n_i} y_{ij}/n$. *Let* $SSR = \sum_{i=1}^{g} n_i\,(\overline{y}_{i\cdot} - \overline{y}_{\cdot\cdot})^2$ *and* $SSE = \sum_{i=1}^{g}\sum_{j=1}^{n_i} (y_{ij} - \overline{y}_{i\cdot})^2$ *be the treatment and the error sum of squares, respectively. Then, under the hypothesis,* $\mu_i = \mu$ *for* $1 \leq i \leq g$ *the following hold*

(1) $SSR \sim d\,\chi^2(g-1)$

(2) $SSE \sim d\chi^2(n-g)$

(3) *SSR is independent of SSE*

if and only if $\mathbf{W} \in \mathcal{G}_{d,n}$ *for some constant* $d > 0$.

Proof: Let $\mathbf{y}' = (y_{11}, \ldots, y_{1n_1}, \ldots, y_{g1}, \ldots, y_{gn_g})$ and $\mu_i = \mu$ for $1 \le i \le g$ then we have $\mathbf{y} \sim N_n(\mu\,\mathbf{e}, \mathbf{W})$. Note that $\mathbf{y}'\mathbf{A}^*\mathbf{y} = SSR + SSE$. It follows from Remark 3.3.1 and Theorem 3.3.5 with $k = 2$ and $c_1 = c_2 = d > 0$ that (1), (2) and (3) hold if and only if $\mathbf{A}^*\mathbf{W}\mathbf{A}^* = d\,\mathbf{A}^*$ which is true if and only if $\mathbf{W} \in \mathcal{G}_{d,n}$ by Remark 3.2.3. This completes the proof of Theorem 3.3.6.

The next theorem concerns invariance property of the null distribution of the quadratic forms in one way ANOVA model where we assume that observations within each treatment are correlated but observations between different treatments are uncorrelated. Theorem 3.3.7 shows that the quadratic forms for testing the equality of g means are independent and have chi-square distributions if and only if all the observations are uncorrelated when g is greater than or equal to 3. The case $g = 2$ was already considered in Theorem 3.3.3.

Theorem 3.3.7 *Consider the one way ANOVA model as in Theorem* 3.3.6. *Let* $\varepsilon_i' = (\varepsilon_{i1}, \ldots, \varepsilon_{in_i})$ *and* $n = \sum_{i=1}^{g} n_i$. *Assume that* ε_i*'s are independent and* $\varepsilon_i \sim N_{n_i}(\mathbf{0}, \mathbf{W}_i)$ *where* $\mathbf{W}_i$*'s are n.n.d. matrices for* $1 \le i \le g$. *If* $g \ge 3$, *under the hypothesis* $\mu_i = \mu$

for $1 \le i \le g$, *we have* (1), (2) *and* (3) *of Theorem* 3.3.6 *hold if and only if* $\mathbf{W}_i = d\,\mathbf{I}_{n_i}$ *for* $1 \le i \le g$ *and for some* $d > 0$.

Proof: Let $\mathbf{y}' = (y_{11}, \ldots, y_{1n_1}, \ldots, y_{g1}, \ldots, y_{gn_g})$. Then, under the hypothesis $\mu_i = \mu$ for $1 \le i \le g$, we have $\mathbf{y} \sim N_n(\mu\,\mathbf{e}, \mathbf{W})$ where $\mathbf{W} = \bigoplus_{i=1}^{g} \mathbf{W}_i$ and $\oplus$ denoting the direct sum of $\mathbf{W}_i$'s. By Theorem 3.3.6, we have (1), (2) and (3) hold if and only if $\mathbf{W} \in \mathcal{G}_{d,n}$ for some $d > 0$, that is,

$$\mathbf{W} = d\left((\mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}') + \frac{1}{n}(\mathbf{e}\,\mathbf{a}' + \mathbf{a}\,\mathbf{e}') - \frac{\bar{a}}{n}\mathbf{e}\,\mathbf{e}'\right) \tag{3.3.13}$$

for some vector $\mathbf{a} \in \Re^n$ satisfying inequality (3.2.2). Let w_{ij} denotes the (i, j)th element of $\mathbf{W}$. Since ε_i's are uncorrelated we require

$$w_{ij} = d(a_i + a_j - (\bar{a} + 1)) = 0 \text{ for } 1 \le i \le n_1,\ n_1 + 1 \le j \le n \tag{3.3.14}$$

and

$$w_{ij} = d(a_i + a_j - (\bar{a} + 1)) = 0 \text{ for } \begin{array}{l} i = n_1 + 1, \ldots, n_1 + n_2, \\ j = 1, \ldots, n_1, n_1 + n_2 + 1, \ldots, n. \end{array} \tag{3.3.15}$$

Since $g \ge 3$, we have $n > n_1 + n_2$ and it is easy to check that (3.3.14) and (3.3.15) hold if and only if $\mathbf{a} = \mathbf{e}$. Therefore, from (3.3.13), we get $\mathbf{W} = d\mathbf{I}$ and hence $\mathbf{W}_i = d\,\mathbf{I}_{n_i}$ for $1 \le i \le g$. This completes the proof of the theorem.

Chapter 4

Wishartness and Independence of Quadratic Forms Under Special Covariance Structures: $\mathbf{W} \otimes \mathbf{V}$

4.1 Introduction

The purpose of this chapter is to study the multivariate generalizations of the results presented in Chapter 3. As preliminaries, we summarize comprehensively the properties of the Kronecker product and the *vec* operator in Section 4.2. In Sections 4.3 and 4.4, we introduce the definitions and prove some properties of the matrix normal distribution and the Wishart distribution. In Section 4.5, we study multivariate quadratic forms and state some known theorems relating to Wishartness and mutual independence of these quadratic forms.

The main results of this chapter are presented in Sections 4.6 and 4.7. Our results are very general and applicable to the singular and nonsingular Wishart distribution. In Section 4.6, we assume that the p-variate normal observations, $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are correlated, that is, $cov(\mathbf{x}_i, \mathbf{x}_j) = w_{ij}\mathbf{V}$ where $\mathbf{W} = (w_{ij})$ and $\mathbf{V}$ are n.n.d. matrices.

With this assumption, in Theorems 4.6.1 and 4.6.2, we give the necessary and sufficient conditions for Wishartness and mutual independence of multivariate quadratic forms. Also, in Theorem 4.6.3, we prove a version of the Cochran's theorem for multivariate observations with the above covariance structure. Finally, in Section 4.7, we obtain a characterization of the class of all covariance matrices $\mathbf{W}$ such that the distributions of test statistics occurring in MANOVA problems remain invariant.

4.2 The Kronecker Product and Vec Operator

In this section, we present some matrix theory results related to the Kronecker product and the *vec* operator. We state these results without giving any proofs.

Kronecker Product

Let $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$ be $m \times n$ and $p \times q$ matrices, respectively. Then the Kronecker product

$$\mathbf{A} \otimes \mathbf{B} = (a_{ij}\mathbf{B}) \tag{4.2.1}$$

is a $mp \times nq$ matrix expressible as a partitioned matrix with $a_{ij}\mathbf{B}$ as the (i,j)th partition for $i = 1, \ldots, m$ and $j = 1, \ldots, n$; that is,

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \cdots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix}. \tag{4.2.2}$$

The Kronecker product is also called the direct product. Some of the important and useful results pertaining to the Kronecker product are summarized below. These results are consequences of (4.2.1). For more discussion, we refer to Graybill (1969) and Henderson et al. (1983).

Result 4.2.1

(a) $\mathbf{0} \otimes \mathbf{A} = \mathbf{A} \otimes \mathbf{0} = \mathbf{0}$

(b) $(\mathbf{A}_1 + \mathbf{A}_2) \otimes \mathbf{B} = (\mathbf{A}_1 \otimes \mathbf{B}) + (\mathbf{A}_2 \otimes \mathbf{B})$

(c) $\mathbf{A} \otimes (\mathbf{B}_1 + \mathbf{B}_2) = (\mathbf{A} \otimes \mathbf{B}_1) + (\mathbf{A} \otimes \mathbf{B}_2)$

(d) $a\mathbf{A} \otimes b\mathbf{B} = ab\mathbf{A} \otimes \mathbf{B}$

(e) $(\mathbf{A}_1 \otimes \mathbf{B}_1)(\mathbf{A}_2 \otimes \mathbf{B}_2) = \mathbf{A}_1\mathbf{A}_2 \otimes \mathbf{B}_1\mathbf{B}_2$

(f) $(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C})$

(g) $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$

(h) $(\mathbf{A} \otimes \mathbf{B})' = \mathbf{A}' \otimes \mathbf{B}'$

(i) $[\mathbf{A}_1 \ \ \mathbf{A}_2] \otimes \mathbf{B} = [\mathbf{A}_1 \otimes \mathbf{B} \ \ \mathbf{A}_2 \otimes \mathbf{B}]$

(j) $r(\mathbf{A} \otimes \mathbf{B}) = r(\mathbf{A})\, r(\mathbf{B})$ and $tr(\mathbf{A} \otimes \mathbf{B}) = tr(\mathbf{A})\, tr(\mathbf{B})$

(k) Let the ith eigenvalue of $\mathbf{A}$ be λ_i with corresponding eigenvector $\mathbf{u}_i$ and jth eigenvalue of $\mathbf{B}$ be ν_j with corresponding eigenvector $\mathbf{v}_j$, then $\mathbf{A} \otimes \mathbf{B}$ has eigenvalues $\lambda_i \nu_j$ with corresponding eigenvectors $\mathbf{u}_i \otimes \mathbf{v}_j$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$.

(l) If $\mathbf{A}$ and $\mathbf{B}$ are n.n.d. matrices then $\mathbf{A} \otimes \mathbf{B}$ is also an n.n.d. matrix.

(m) $(\mathbf{A} \otimes \mathbf{B})^{-} = \mathbf{A}^{-} \otimes \mathbf{B}^{-}$

(n) $\mathbf{P}_{(\mathbf{A} \otimes \mathbf{B})} = \mathbf{P}_{\mathbf{A}} \otimes \mathbf{P}_{\mathbf{B}}$ where $\mathbf{P}_{\mathbf{X}}$ denotes orthogonal projection matrix onto $\mathcal{M}(\mathbf{X})$.

Vec Operator

A matrix operation dating back nearly a century is that of stacking the columns of a matrix one under the other to form a single column. Over the years it has had variety of names, the most recent being *vec*. Thus for a matrix

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$$

where $\mathbf{x}_i$ is $p \times 1$ vector for $i = 1, \ldots, n$; $vec(\mathbf{X})$ is a vector of order $np \times 1$ which is defined as

$$vec(\mathbf{X}) = \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{bmatrix}.$$

Henderson and Searle (1979) give history, properties and many applications of *vec* operator. In the following lemma, we have combined some of the results given in Henderson and Searle (1979) and Neudecker (1969). It is assumed that the sizes of the matrices are such that all the statements make sense.

Lemma 4.2.1

(a) $vec(\mathbf{A} + \mathbf{B}) = vec(\mathbf{A}) + vec(\mathbf{B})$

(b) $vec(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})\ vec(\mathbf{B})$

(c) $vec(\mathbf{PQ}) = (\mathbf{Q}' \otimes \mathbf{I})\ vec(\mathbf{P}) = (\mathbf{Q}' \otimes \mathbf{P})\ vec(\mathbf{I}) = (\mathbf{I} \otimes \mathbf{P})\ vec(\mathbf{Q})$

(d) $tr(\mathbf{ABC}) = vec(\mathbf{A}')'\,(\mathbf{I} \otimes \mathbf{B})\ vec(\mathbf{C})$

(e) $tr(\mathbf{PQ}) = vec(\mathbf{P}')'vec(\mathbf{Q})$

(f) $tr(\mathbf{AZ}'\mathbf{BZC}) = vec(\mathbf{Z})'(\mathbf{A}'\mathbf{C}' \otimes \mathbf{B})\ vec(\mathbf{Z}) = vec(\mathbf{Z})'(\mathbf{CA} \otimes \mathbf{B}')\ vec(\mathbf{Z})$

4.3 The Matrix Normal Distribution

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ be a sample of size n from a p-dimensional population. Define $p \times n$ matrix $\mathbf{X}$ such that

$$\mathbf{X} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \cdots & \vdots \\ x_{p1} & \cdots & x_{pn} \end{bmatrix} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n] = \begin{bmatrix} \mathbf{x}_1^{*\prime} \\ \vdots \\ \mathbf{x}_p^{*\prime} \end{bmatrix} \tag{4.3.1}$$

where $\mathbf{x}_j = [x_{1j}, \ldots, x_{pj}]'$ $(j = 1, \ldots, n)$ is the jth column of $\mathbf{X}$ and $\mathbf{x}_i^{*\prime} = [x_{i1}, \ldots, x_{in}]$ $(i = 1, \ldots, p)$ is the ith row of $\mathbf{X}$. Note that $\mathbf{x}_j$ represents p observations on the jth object or individual while $\mathbf{x}_i^*$ represents n observations on the ith variate. Geometrically, $\mathbf{x}_j$'s are n points in $\Re^p$ serving for an examination of the relationship among different objects. On the other hand, $\mathbf{x}_i^*$'s are p points in $\Re^n$ serving for an investigation of the relationship among different variates. We call $\mathbf{X}$, an observation matrix.

The sample mean vector and covariance matrix are denoted by a vector $\bar{\mathbf{x}}$ of order $p \times 1$ and a matrix $\mathbf{S}$ of order $p \times p$, respectively. The scatter of n points in $\Re^p$ provides information on their location and variability. If the points are regarded as solid spheres, the sample mean vector $\bar{\mathbf{x}}$ given by (4.3.2) is the center of balance. Variability occurs in more than one direction and it is quantified by the sample covariance matrix $\mathbf{S}$ given by (4.3.3). The sample mean vector and covariance matrix of $\mathbf{X}$ are defined below

$$\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_p \end{bmatrix} = \begin{bmatrix} \frac{1}{n}\sum_{j=1}^{n} x_{1j} \\ \vdots \\ \frac{1}{n}\sum_{j=1}^{n} x_{pj} \end{bmatrix} = \begin{bmatrix} \frac{1}{n}\mathbf{x}_1^{*\prime}\mathbf{e} \\ \vdots \\ \frac{1}{n}\mathbf{x}_p^{*\prime}\mathbf{e} \end{bmatrix} = \frac{1}{n}\mathbf{X}\mathbf{e} \tag{4.3.2}$$

and

$$\begin{aligned} \mathbf{S} &= (s_{ij}) = \left(\frac{1}{n-1}\sum_{k=1}^{n}(x_{ik}-\bar{x}_i)(x_{jk}-\bar{x}_j)\right) = \left(\frac{1}{n-1}\mathbf{x}_i^{*\prime}\,(\mathbf{I}-\frac{1}{n}\mathbf{e}\,\mathbf{e}')\,\mathbf{x}_j^{*}\right) \\ &= \frac{1}{n-1}\mathbf{X}\left(\mathbf{I}-\frac{1}{n}\mathbf{e}\,\mathbf{e}'\right)\mathbf{X}'. \end{aligned} \tag{4.3.3}$$

Now we are ready to define the matrix normal distribution.

Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$ be an observation matrix of order $p \times n$. Let $\mathbf{x}_j \sim N_p(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$ for $j = 1, \ldots, n$ and $\mathbf{M} = [\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_n]$ be a matrix of order $p \times n$. Also denote, $\tilde{\mathbf{x}} = vec(\mathbf{X})$, $\tilde{\mathbf{m}} = vec(\mathbf{M})$ and the covariance matrix of $\tilde{\mathbf{x}}$ by an n.n.d. matrix $\boldsymbol{\Sigma}$ of order $np \times np$.

Definition 4.3.1 *The $p \times n$ observation matrix $\mathbf{X}$ is said to have a matrix normal distribution and is denoted by $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \boldsymbol{\Sigma})$ if $\tilde{\mathbf{x}} \sim N_{pn}(\tilde{\mathbf{m}}, \boldsymbol{\Sigma})$.*

That is, the statements "$\mathbf{X} \sim N_{p,n}(\mathbf{M}, \boldsymbol{\Sigma})$" and "$\tilde{\mathbf{x}} \sim N_{pn}(\tilde{\mathbf{m}}, \boldsymbol{\Sigma})$" are equivalent and we say that random matrix $\mathbf{X}$ has a matrix normal distribution with matrix of means $\mathbf{M} = [\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_n]$ and that covariance matrix of $\tilde{\mathbf{x}}$ is $\boldsymbol{\Sigma}$. In this case, the observation matrix $\mathbf{X}$ is called the normal observation matrix. The matrix normal distribution is nonsingular(singular) as Σ is positive definite(positive semidefinite) matrix.

The singular matrix normal distribution is defined through the singular multivariate normal distribution. That is, if Σ is n.n.d. then by $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \Sigma)$ we mean that the probability distribution of $\tilde{\mathbf{x}}$ is same as the probability distribution of $\widetilde{\mathbf{m}}+\mathbf{L}\mathbf{z}$ where $\mathbf{L}$ is a matrix of full column rank and of order $np \times m$ such that $\Sigma = \mathbf{L}\mathbf{L}'$ and $\mathbf{z} \sim N_m(\mathbf{0}, \mathbf{I}_m)$. In all our discussion, in this chapter and in the next chapter Σ is assumed to be an n.n.d. matrix unless otherwise stated and $P(\mathbf{X} = \mathbf{O}) = 0$. In the following examples we consider two widely used covariance structures for Σ.

Example 4.3.1 Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$ be an observation matrix of order $p \times n$ such that $\mathbf{x}_j$'s are independent and identically distributed (i.i.d.) $N_p(\boldsymbol{\mu}, \mathbf{V})$ random vectors for $j = 1, \ldots, n$. Also, let $\mathbf{M} = [\boldsymbol{\mu}, \ldots, \boldsymbol{\mu}]$ be a matrix of order $p \times n$ and $\mathbf{V}$ be an n.n.d. matrix then $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \Sigma)$ where $\mathbf{M} = \boldsymbol{\mu}\,\mathbf{e}'$ and $\Sigma = \mathbf{I} \otimes \mathbf{V}$.

Example 4.3.2 Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$ be an observation matrix of order $p \times n$ such that $\mathbf{x}_j \sim N_p(\boldsymbol{\mu}_j, w_{jj}\mathbf{V})$ and $cov(\mathbf{x}_i, \mathbf{x}_j) = w_{ij}\,\mathbf{V}$ for $i, j = 1, \ldots, n$. Let $\mathbf{W} = (w_{ij})$ and $\mathbf{V}$ be n.n.d. matrices then $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \Sigma)$ where $\mathbf{M} = [\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_n]$ and $\Sigma = \mathbf{W} \otimes \mathbf{V}$.

In Theorem 4.3.1, we present a nice characterization of the singular matrix normal distribution given in Example 4.3.2. This characterization is used in Theorem 4.6.1

and 4.6.2 to obtain the necessary and sufficient conditions for the Wishartness and mutual independence of multivariate quadratic forms.

Let $\mathbf{W}$ and $\mathbf{V}$ be as defined in Example 4.3.2. By rank factorization of $\mathbf{W}$ and $\mathbf{V}$, we have

$$\mathbf{W} = \mathbf{T}\mathbf{T}' \text{ and } \mathbf{V} = \mathbf{U}\mathbf{U}' \tag{4.3.4}$$

where $\mathbf{T}$ and $\mathbf{U}$ are matrices of full column rank and of order $n \times r$ and $p \times s$, respectively. Using property (e) of Kronecker product, we can show that

$$\mathbf{W} \otimes \mathbf{V} = (\mathbf{T} \otimes \mathbf{U})\,(\mathbf{T}' \otimes \mathbf{U}') \tag{4.3.5}$$

$$= \mathbf{L}\mathbf{L}' \tag{4.3.6}$$

where $\mathbf{L} = \mathbf{T} \otimes \mathbf{U}$.

Theorem 4.3.1 *Let* $\mathbf{X}$, $\mathbf{M}$, $\mathbf{W}$ *and* $\mathbf{V}$ *be as given in Example* 4.3.2. *Let* $\mathbf{W}$ *be of rank* r *and* $\mathbf{V}$ *be of rank* s *then* $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ *if and only if* $\mathbf{X} = \mathbf{M} + \mathbf{U}\,\mathbf{Z}\,\mathbf{T}'$*; that is,* $\mathbf{X}$ *has the same probability distribution as that of* $\mathbf{M} + \mathbf{U}\,\mathbf{Z}\,\mathbf{T}'$ *where* $\mathbf{T}$ *and* $\mathbf{U}$ *are defined in* (4.3.4) *and* $\mathbf{Z} \sim N_{s,r}(\mathbf{O}, \mathbf{I}_r \otimes \mathbf{I}_s)$.

Proof: If $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ then from Definition 4.3.1, we have $\tilde{\mathbf{x}} \sim N_{pn}(\widetilde{\mathbf{m}}, \mathbf{W} \otimes \mathbf{V})$. It follows from a result given in Anderson (1984), page 32-33, that with probability one

$$\tilde{\mathbf{x}} = \widetilde{\mathbf{m}} + (\mathbf{T} \otimes \mathbf{U})\,\tilde{\mathbf{z}} \tag{4.3.7}$$

where $\tilde{z} \sim N_{rs}(\mathbf{0}, \mathbf{I}_r \otimes \mathbf{I}_s)$. Rewriting (4.3.7) in matrix form, we have with probability one $\mathbf{X} = \mathbf{M} + \mathbf{U}\,\mathbf{Z}\,\mathbf{T}'$ where $\mathbf{Z} \sim N_{s,r}(\mathbf{O}, \mathbf{I}_r \otimes \mathbf{I}_s)$.

To prove the converse, note that

$$vec(\mathbf{X}) = vec(\mathbf{M} + \mathbf{U}\,\mathbf{Z}\,\mathbf{T}') = \widetilde{\mathbf{m}} + (\mathbf{T} \otimes \mathbf{U})\,\tilde{\mathbf{z}} \tag{4.3.8}$$

where $\tilde{\mathbf{z}} = vec(\mathbf{Z})$. Hence $\tilde{\mathbf{x}} = \widetilde{\mathbf{m}} + (\mathbf{T} \otimes \mathbf{U})\,\tilde{\mathbf{z}} \sim N_{np}(\widetilde{\mathbf{m}}, \mathbf{W} \otimes \mathbf{V})$, since $\tilde{\mathbf{z}} \sim N_{rs}(\mathbf{0}, \mathbf{I}_r \otimes \mathbf{I}_s)$. It follows from Definition 4.3.1 that $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$. This completes the proof of the theorem.

Properties of Matrix Normal Distribution

There are several interesting properties of matrix normal distribution. We state two important properties which we use later in this chapter. We begin with the definition of the commutation matrix, $\mathbf{K}_{pn}$. The matrix $\mathbf{K}_{pn}$ is called "commutation matrix" because of its role in reversing ("commuting") the order of Kronecker products. Henderson and Searle (1981) call it a permutation matrix. We refer to Magnus and Neudecker (1979) and Henderson and Searle (1981) for various properties of $\mathbf{K}_{pn}$. The commutation matrix $\mathbf{K}_{pn}$ is defined as

$$\mathbf{K}_{pn} = \sum_{i=1}^{p} \sum_{j=1}^{n} \left(\mathbf{H}_{ij} \otimes \mathbf{H}'_{ij}\right) \tag{4.3.9}$$

where $\mathbf{H}_{ij}$ is a $p \times n$ matrix with a one at the (i, j)th position and zeros elsewhere. It follows from Theorem 3.1 of Magnus and Neudecker (1979) that

$$\mathbf{K}_{pn}\, vec(\mathbf{X}) \;=\; vec(\mathbf{X}') \tag{4.3.10}$$

and

$$\mathbf{K}_{pn}\,(\mathbf{W}\otimes\mathbf{V})\,\mathbf{K}'_{pn} \;=\; \mathbf{V}\otimes\mathbf{W} \tag{4.3.11}$$

where $\mathbf{X}$, $\mathbf{W}$ and $\mathbf{V}$ are matrices of order $p\times n$, $n\times n$ and $p\times p$, respectively.

We state some useful properties of matrix normal distribution in the following two theorems.

Theorem 4.3.2 *Let* $\mathbf{X}$*,* $\mathbf{M}$*,* $\mathbf{W}$ *and* $\mathbf{V}$ *be as given in Example* 4.3.2. *If* $\mathbf{X}\sim N_{p,n}(\mathbf{M},\mathbf{W}\otimes\mathbf{V})$ *then* $\mathbf{X}'\sim N_{n,p}(\mathbf{M}',\mathbf{V}\otimes\mathbf{W})$.

Proof: If $\mathbf{X}\sim N_{p,n}(\mathbf{M},\mathbf{W}\otimes\mathbf{V})$ then it follows from Definition 4.3.1 that $vec(\mathbf{X})\sim N_{pn}(vec(\mathbf{M}),\mathbf{W}\otimes\mathbf{V})$. From (4.3.10), we have

$$vec(\mathbf{X}')\sim N_{np}\left(\mathbf{K}_{pn}vec(\mathbf{M}),\, \mathbf{K}_{pn}\,(\mathbf{W}\otimes\mathbf{V})\,\mathbf{K}'_{pn}\right). \tag{4.3.12}$$

It follows from (4.3.10) and (4.3.11) that $vec(\mathbf{X}')\sim N_{np}(vec(\mathbf{M}'),\,\mathbf{V}\otimes\mathbf{W})$ which implies $\mathbf{X}'\sim N_{n,p}(\mathbf{M}',\mathbf{V}\otimes\mathbf{W})$.

In the following theorem, we show that if $\mathbf{X}$ has a matrix normal distribution then the matrix of linear combinations of columns and rows of $\mathbf{X}$ also has a matrix normal distribution.

Theorem 4.3.3 *Let* $\mathbf{X}$*,* $\mathbf{M}$*,* $\mathbf{W}$ *and* $\mathbf{V}$ *be as given in Example* 4.3.2. *If* $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ *then* $\mathbf{A}\mathbf{X}\mathbf{B}' \sim N_{m,s}(\mathbf{A}\mathbf{M}\mathbf{B}', \mathbf{B}\mathbf{W}\mathbf{B}' \otimes \mathbf{A}\mathbf{V}\mathbf{A}')$ *where* $\mathbf{A}$ *and* $\mathbf{B}$ *are matrices of known constants and of order* $m \times p$ *and* $s \times n$*, respectively.*

Proof: From Lemma 4.2.1(b) it follows that

$$\begin{aligned} vec(\mathbf{A}\mathbf{X}\mathbf{B}') &= (\mathbf{B} \otimes \mathbf{A})\, vec(\mathbf{X}) && (4.3.13) \\ &= \mathbf{T}\, vec(\mathbf{X}) && (4.3.14) \end{aligned}$$

where $\mathbf{T} = \mathbf{B} \otimes \mathbf{A}$.

If $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ then $\mathbf{T}\, vec(\mathbf{X}) \sim N_{ms}(\mathbf{T}\, vec(\mathbf{M}), \mathbf{T}(\mathbf{W} \otimes \mathbf{V})\mathbf{T}')$. It is easy to verify that

$$\begin{aligned} \mathbf{T}\, vec(\mathbf{M}) &= vec(\mathbf{A}\mathbf{M}\mathbf{B}') \text{ and} && (4.3.15) \\ \mathbf{T}(\mathbf{W} \otimes \mathbf{V})\mathbf{T}' &= \mathbf{B}\mathbf{W}\mathbf{B}' \otimes \mathbf{A}\mathbf{V}\mathbf{A}'. && (4.3.16) \end{aligned}$$

Now, from Definition 4.3.1 we have the required result.

4.4 The Wishart Distribution

In this section, we present the definition of Wishart distribution along with some of its interesting properties. The definition given here is very general in the sense that the Wishart distribution is defined through matrix normal distribution without assuming the existence of the density function. Also, the following definition allows for Wishart distribution that are singular. An additional advantage of using this definition is that we can derive most of the properties of Wishart distribution by exploiting its representation in the form of normal random vectors.

Definition 4.4.1 *Let* $\mathbf{x}_1, \ldots, \mathbf{x}_n$ *be i.i.d.* $N_p(\mathbf{0}, \mathbf{V})$ *random vectors, that is,* $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_n] \sim N_{p,n}(\mathbf{O}, \mathbf{I} \otimes \mathbf{V})$ *then* $\mathcal{W} = \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i' = \mathbf{X}\mathbf{X}'$ *is said to have the Wishart distribution with* n *degrees of freedom and covariance matrix* $\mathbf{V}$*. We will write that* $\mathcal{W} \sim W_p(n, \mathbf{V})$*, the subscript on* W *denoting the size of the matrix* $\mathcal{W}$*.*

Remark 4.4.1 In the above definition, p and n are positive integers and $\mathbf{V}$ is an n.n.d. matrix. The random matrix $\mathcal{W}$ has a nonsingular Wishart distribution if and only if $n \geq p$ and $\mathbf{V}$ is a positive definite matrix, see Eaton 1983, page 304. From now onwards, in all our discussion, the covariance matrix $\mathbf{V}$ is assumed to be n.n.d. unless otherwise stated. Also, since the definition of Wishart distribution does not require $n \geq p$, we will not assume $n \geq p$ unless we need this restriction.

Remark 4.4.2 Since $\mathbf{X}\mathbf{X}'$ is n.n.d., the Wishart distribution has all of its mass on the set of n.n.d. matrices. If $p = 1$ then $\mathcal{W} = x_1^2 + \cdots + x_n^2$ where x_i's are i.i.d. $N(0, \sigma^2)$, in this case it is clear that $W_1(n, \sigma^2)$ is same as the $\sigma^2 \chi^2(n)$ distribution hence the Wishart distribution is a matrix generalization of the chi-square distribution. In fact, it plays the same role as that of a χ^2-distribution in multivariate regression analysis, MANOVA and more generally, in multivariate linear models.

Remark 4.4.3 If $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{I} \otimes \mathbf{V})$ then $\mathbf{X}\mathbf{X}'$ is said to have a noncentral Wishart distribution, denoted by $W_p(n, \mathbf{V}; \Omega)$ where $\Omega = \mathbf{M}\mathbf{M}'$ is known as the noncentrality matrix. Note that the distribution of $\mathbf{XX}'$ depends on $\mathbf{M}$ only through Ω, see Eaton 1983, page 316.

Remark 4.4.4 If $\mathcal{W} \sim W_p(n, \mathbf{V})$ then we often write $\mathcal{W} = \mathbf{XX}'$ where $\mathbf{X} \sim N_{p,n}(\mathbf{O}, \mathbf{I} \otimes \mathbf{V})$, that is, the distribution of $\mathcal{W}$ is same as the distribution of $\mathbf{XX}'$.

Remark 4.4.5 The density function of a nonsingular Wishart distribution was first obtained by Fisher (1915) when $p = 2$ and for a general p by Wishart (1928) using a geometrical argument. It seems that the multivariate analysis has begun its successive and rapid progress with this discovery. There are various methods available for the derivation of the Wishart distribution, see for example, Wishart and Bartlett (1933), Madow (1938), Hsu (1939a), Olkin and Roy (1954) and James (1954).

Properties of Wishart Distribution

The Wishart distribution has several interesting properties. In this section, we present some of its properties that are useful for us in deriving the results given here and in the next chapter. The first property is important because it gives us a method to generate the family of singular Wishart distributions.

Theorem 4.4.1 *If $\mathcal{W} \sim W_p(n, \mathbf{V})$ and $\mathbf{C}$ is a $r \times p$ matrix of constants then $\mathbf{C}\,\mathcal{W}\,\mathbf{C}' \sim W_r(n, \mathbf{C}\,\mathbf{V}\,\mathbf{C}')$.*

Proof: Since $\mathcal{W} \sim W_p(n, \mathbf{V})$, we have $\mathcal{W} = \mathbf{XX}'$ where $\mathbf{X} \sim N_{p,n}(\mathbf{O}, \mathbf{I} \otimes \mathbf{V})$. Thus $\mathbf{C}\,\mathcal{W}\,\mathbf{C}' = \mathbf{C}\mathbf{X}\mathbf{X}'\mathbf{C}' = \mathbf{Y}\mathbf{Y}'$ where $\mathbf{Y} = \mathbf{C}\mathbf{X}$. From Theorem 4.3.3, we have

$\mathbf{Y} \sim N_{r,n}(\mathbf{O}, \mathbf{I} \otimes \mathbf{C}\,\mathbf{V}\,\mathbf{C}')$. Hence by Definition 4.4.1 of Wishart distribution, we have the result.

Remark 4.4.6 We can easily show that if $\mathcal{W} \sim W_p(n, \mathbf{V}; \Omega)$ then $\mathbf{C}\,\mathcal{W}\,\mathbf{C}' \sim W_r(n, \mathbf{C}\,\mathbf{V}\,\mathbf{C}'; \mathbf{C}\,\Omega\,\mathbf{C}')$.

The following corollary is an easy consequence of Theorem 4.4.1.

Corollary 4.4.1 *Let* $\mathbf{U}$ *be a matrix of full column rank and of order* $p \times s$ *such that* $\mathbf{V} = \mathbf{U}\,\mathbf{U}'$ *then* $\mathcal{W} \sim W_s(n, \mathbf{I}_s)$ *if and only if* $\mathbf{U}\,\mathcal{W}\,\mathbf{U}' \sim W_p(n, \mathbf{V})$.

It is clear from the above corollary that the family of singular Wishart distributions, $W_p(n, \mathbf{V})$, can be generated from the $W_s(n, \mathbf{I}_s)$ distribution where s is the rank of $\mathbf{V}$. Now, we state a straightforward corollary of Theorem 4.4.1.

Corollary 4.4.2 *If* $\mathcal{W} \sim W_p(n, \mathbf{V})$ *then* $\mathbf{c}'\,\mathcal{W}\,\mathbf{c} \sim (\mathbf{c}'\,\mathbf{V}\,\mathbf{c})\,\chi^2(n)$ *where* $\mathbf{c}$ *is a vector of constants of order* $p \times 1$.

The following theorem is stated without a proof and is useful to prove some of the results related to the Wishartness of multivariate quadratic forms. For a proof, see Siotani et.al. (1985), page 66.

Theorem 4.4.2 *Let* $\mathcal{W} \sim W_p(n, \mathbf{V}; \Omega)$ *where* $\mathbf{V}$ *is a diagonal matrix,* $\mathbf{V} = \mathrm{diag}(\sigma_{11}, \ldots, \sigma_{pp})$, *then the diagonal elements* $w_{11}, \ldots, w_{pp}$ *of* $\mathcal{W}$ *are all independent and* $w_{ii} \sim \sigma_{ii}\chi^2(n; \delta_{ii})$ *for* $i = 1, \ldots, p$ *where* δ_{ii} *is the* i*th diagonal element of* Ω.

The following corollary is an easy consequence of the above theorem.

Corollary 4.4.3 *Let* $\mathcal{W} \sim W_p(n, \mathbf{I}; \Omega)$ *then* $tr(\mathcal{W}) \sim \chi^2(np; tr(\Omega))$.

We refer to Rao (1973), page 537-538 for a proof of the following theorem.

Theorem 4.4.3 *Let* $\mathcal{W}_1 \sim W_p(n_1, \mathbf{V}; \Omega_1)$ *and* $\mathcal{W}_2 \sim W_p(n_2, \mathbf{V}; \Omega_2)$ *be independently distributed then* $\mathcal{W} = \mathcal{W}_1 + \mathcal{W}_2 \sim W_p(n_1 + n_2, \mathbf{V}; \Omega_1 + \Omega_2)$.

We now proceed to define the Hotelling's generalized T^2 statistic. Let $\mathcal{W} \sim W_p(m, \mathbf{V})$ where $\mathbf{V}$ is a positive definite matrix and $m \geq p$. Suppose, $\mathbf{y} \sim N_p(\boldsymbol{\mu}, \frac{1}{c}\mathbf{V})$ where $c > 0$. Let $\mathbf{y}$ and $\mathcal{W}$ be independently distributed then Hotelling's generalized T^2 statistic is defined by

$$T^2 = cm\mathbf{y}'\mathcal{W}^{-1}\mathbf{y}. \tag{4.4.1}$$

Theorem 4.4.4 *Let* T^2 *be as defined in* (4.4.1) *then*

$$\frac{T^2}{m}\frac{m-p+1}{p} \sim F(p, m-p+1; \delta) \tag{4.4.2}$$

where $\delta = c\boldsymbol{\mu}'\mathbf{V}^{-1}\boldsymbol{\mu}$ *and* $F(m_1, m_2; \omega)$ *denotes a noncentral* F*- distribution with* (m_1, m_2) *degrees of freedom and noncentrality parameter* ω.

Proof: See, Rao (1973), page 541-542 for a proof of the theorem. It is easy to see that the distribution is central F if and only if $\boldsymbol{\mu} = \mathbf{0}$.

4.5 Multivariate Quadratic Forms

Multivariate quadratic forms are used in multivariate linear models in the construction of various test statistics such as Wilk's lambda and the likelihood ratio tests. In this section, we first define a multivariate quadratic form and then present the necessary and sufficient conditions for certain quadratic forms to be mutually independent and to follow a Wishart distribution.

Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$ be an observation matrix of order $p \times n$. Let $\mathbf{A}$ be a symmetric matrix of order $n \times n$ then the multivariate quadratic form, $\mathbf{Q}(\mathbf{X})$, is defined as

$$\mathbf{Q}(\mathbf{X}) = (q_{ij}) = (\mathbf{x}_i^{*\prime}\,\mathbf{A}\,\mathbf{x}_j^{*}) = \mathbf{X}\,\mathbf{A}\,\mathbf{X}' \tag{4.5.1}$$

where $\mathbf{x}_i^{*\prime}$ represents the ith row of $\mathbf{X}$. Note that $\mathbf{Q}(\mathbf{X})$ is a symmetric matrix of order $p \times p$. In the rest of the section, we present some basic results due to Khatri (1962).

Theorem 4.5.1 *Let* $\mathbf{X}$ *be an observation matrix of order* $p \times n$ *such that* $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{I} \otimes \mathbf{V})$ *where* $\mathbf{V}$ *is a positive definite matrix. Let* $\mathbf{P}(\mathbf{X}) = \mathbf{X}\,\mathbf{A}\,\mathbf{X}' + \frac{1}{2}(\mathbf{L}\,\mathbf{X}' + \mathbf{X}\,\mathbf{L}') + \mathbf{C}$ *where* $\mathbf{A}$ *and* $\mathbf{C}$ *are symmetric matrices of order* $n \times n$ *and* $p \times p$*, respectively and* $\mathbf{L}$ *is of order* $p \times n$*. Then* $\mathbf{P}(\mathbf{X}) \sim W_p(r(\mathbf{A}), \mathbf{V}; \Omega)$ *if and only if*

$$\text{(i) } \mathbf{A}^2 = \mathbf{A}, \quad \text{(ii) } \mathbf{L} = \mathbf{L}\,\mathbf{A} \text{ and } \text{(iii) } \mathbf{C} = \frac{1}{4}\mathbf{L}\,\mathbf{A}\,\mathbf{L}.' \tag{4.5.2}$$

Also, if conditions (i), (ii) *and* (iii) *are satisfied then* $\Omega = \mathbf{P}(\mathbf{M})$.

A proof of the corollary given below can be found in Siotani et.al. (1985), page 94.

Corollary 4.5.1 *The distribution of* $\mathbf{P}(\mathbf{X})$ *in Theorem* 4.5.1 *is a central Wishart distribution,* $W_p(r(\mathbf{A}), \mathbf{V})$*, if and only if*

$$\text{(i) } \mathbf{A}^2 = \mathbf{A}, \quad \text{(ii) } \mathbf{L} = -2\mathbf{M}\,\mathbf{A} \text{ and (iii) } \mathbf{C} = \mathbf{M}\,\mathbf{A}\,\mathbf{M}.' \qquad (4.5.3)$$

The following corollary is of special interest to us because in many practical situations (such as MANOVA problems) we work with multivariate quadratic forms of the type $\mathbf{Q}(\mathbf{X})$ defined in (4.5.1). We also know, from (4.3.3) that the sample covariance matrix $\mathbf{S}$ is in the form of $\mathbf{Q}(\mathbf{X})$. Hence it is important to study the distributional properties of $\mathbf{Q}(\mathbf{X})$. Moreover, in the next section we generalize the following corollary.

Corollary 4.5.2 *Let the distribution of* $\mathbf{X}$ *be as given in Theorem* 4.5.1 *and* $\mathbf{Q}(\mathbf{X})$ *be as defined in* (4.5.1)*, then* $\mathbf{Q}(\mathbf{X}) \sim W_p(r(\mathbf{A}), \mathbf{V}; \Omega)$ *if and only if* $\mathbf{A}^2 = \mathbf{A}$*. Further, if* $\mathbf{A}^2 = \mathbf{A}$ *then* $\Omega = \mathbf{Q}(\mathbf{M})$.

Remark 4.5.1 The distribution of $\mathbf{Q}(\mathbf{X})$ in the above corollary is a central Wishart distribution, $W_p(r(\mathbf{A}), \mathbf{V})$, if and only if $\mathbf{A}\,\mathbf{M}' = \mathbf{O}$.

We now present some results due to Khatri (1962) pertaining to the independence of two multivariate quadratic forms. We also give the necessary and sufficient conditions for the independence of a multivariate quadratic form and a linear function of the observation matrix $\mathbf{X}$ such as $\mathbf{X}\,\mathbf{K}'$. These results are multivariate generalizations of similar results obtained by many authors, for example, see Shanbhag (1966), Craig (1943) and Lancaster (1954) for the case when $p = 1$.

Theorem 4.5.2 *Let* $\mathbf{X}$ *be an observation matrix of order* $p \times n$ *such that* $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{I} \otimes \mathbf{V})$ *where* $\mathbf{V}$ *is a positive definite matrix. Let* $\mathbf{P}_1(\mathbf{X}) = \mathbf{X}\,\mathbf{A}\,\mathbf{X}' + \frac{1}{2}(\mathbf{L}\,\mathbf{X}' + \mathbf{X}\,\mathbf{L}') + \mathbf{C}$ *and* $\mathbf{P}_2(\mathbf{X}) = \mathbf{X}\,\mathbf{B}\,\mathbf{X}' + \frac{1}{2}(\mathbf{K}\,\mathbf{X}' + \mathbf{X}\,\mathbf{K}') + \mathbf{D}$ *where* $\mathbf{A}$ *and* $\mathbf{B}$ *are symmetric matrices of order* $n \times n$*,* $\mathbf{C}$ *and* $\mathbf{D}$ *are symmetric matrices of order* $p \times p$*,* $\mathbf{L}$ *and* $\mathbf{K}$ *are matrices of order* $p \times n$*. Then* $\mathbf{P}_1(\mathbf{X})$ *and* $\mathbf{P}_2(\mathbf{X})$ *are independently distributed if and only if*

$$\text{(i) } \mathbf{A}\,\mathbf{B} = \mathbf{O}, \quad \text{(ii) } \mathbf{L}\,\mathbf{B} = \mathbf{O} = \mathbf{K}\,\mathbf{A} \ \text{ and } \ \text{(iii) } \mathbf{K}\,\mathbf{L}' = \mathbf{O}. \qquad (4.5.4)$$

Corollary 4.5.3 *Let* $\mathbf{X}$ *be an observation matrix of order* $p \times n$ *such that* $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{I} \otimes \mathbf{V})$ *where* $\mathbf{V}$ *is a positive definite matrix. Let* $\mathbf{P}(\mathbf{X}) = \mathbf{X}\,\mathbf{A}\,\mathbf{X}' + \frac{1}{2}(\mathbf{L}\,\mathbf{X}' + \mathbf{X}\,\mathbf{L}') + \mathbf{C}$*. Then* $\mathbf{P}(\mathbf{X})$ *and* $\mathbf{X}\,\mathbf{K}'$ *are independently distributed if and only if*

$$\text{(i) } \mathbf{K}\,\mathbf{A} = \mathbf{O} \quad \text{and} \quad \text{(ii) } \mathbf{K}\,\mathbf{L}' = \mathbf{O} \qquad (4.5.5)$$

where $\mathbf{A}$*,* $\mathbf{L}$*,* $\mathbf{C}$ *and* $\mathbf{K}$ *are as defined in Theorem* 4.5.2.

4.6 The Covariance Structure $\mathbf{W} \otimes \mathbf{V}$

Let $\mathbf{X}$ be an observation matrix of order $p \times n$ such that $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ where $\mathbf{M}$, $\mathbf{W}$ and $\mathbf{V}$ be as given in Example 4.3.2. In this section, we first obtain the distribution of $\mathbf{Q}(\mathbf{X}) = \mathbf{X}\,\mathbf{A}\,\mathbf{X}'$ and later give the necessary and sufficient conditions for the independence of two multivariate quadratic forms. We then present a multivariate version of the Cochran's theorem; and some statistical applications related to the invariance properties of the Hotelling's T^2 statistic and the MANOVA problems. For earlier works, we refer to Khatri (1962, 1959), Siotani et.al. (1985) and Roy and Gnanadesikan (1959).

Khatri (1962) has given the necessary and sufficient conditions for the Wishartness and independence of multivariate quadratic forms for the case when $\mathbf{W}$ and $\mathbf{V}$ are positive definite matrices whereas Siotani et.al. (1985) discussed the case when $\mathbf{W}$ is n.n.d. and $\mathbf{V}$ is a positive definite matrix. We first present two important lemmas. The following lemma is used in the proofs of some theorems related to Wishartness of multivariate quadratic forms in singular normal observation matrix.

Lemma 4.6.1 *Let* $\mathbf{A}$ *and* $\mathbf{W}$ *be symmetric matrices of order* $n \times n$*. Consider the following two conditions:*

(a) $\mathbf{W}$ *is an n.n.d. matrix such that* $tr(\mathbf{A}\,\mathbf{W}) = r(\mathbf{A})$

(b) $r(\mathbf{A}\,\mathbf{W}) = r(\mathbf{A})$.

If the condition (a) *or* (b) *holds then*

$$\text{(i) } \mathbf{W}\,\mathbf{A}\,\mathbf{W}\,\mathbf{A}\,\mathbf{W} = \mathbf{W}\,\mathbf{A}\,\mathbf{W}, \quad \text{(ii) } \mathbf{M}\,\mathbf{A}\,\mathbf{W} = \mathbf{M}\,\mathbf{A}\,\mathbf{W}\,\mathbf{A}\,\mathbf{W} \quad \text{and}$$

$$\text{(iii) } \mathbf{M}\,\mathbf{A}\,\mathbf{M}' = \mathbf{M}\,\mathbf{A}\,\mathbf{W}\,\mathbf{A}\,\mathbf{M}' \tag{4.6.6}$$

if and only if

$$\mathbf{A}\,\mathbf{W}\,\mathbf{A} = \mathbf{A} \tag{4.6.7}$$

where $\mathbf{M}$ *is a matrix of order* $p \times n$.

Proof: It is easy to see that (4.6.7) implies (i), (ii) and (iii) of (4.6.6). To prove the converse, let (a) be given then from (4.6.6)(i), we get

$$\mathbf{T}'\,\mathbf{A}\,\mathbf{T}\,\mathbf{T}'\,\mathbf{A}\,\mathbf{T} = \mathbf{T}'\,\mathbf{A}\,\mathbf{T} \tag{4.6.8}$$

where $\mathbf{T}$ is defined in (4.3.4). Hence, we have $r(\mathbf{T}'\,\mathbf{A}\,\mathbf{T}) = r(\mathbf{A})$ which implies

$$r(\mathbf{A}) = r(\mathbf{T}'\,\mathbf{A}\,\mathbf{T}) \leq r(\mathbf{A}\,\mathbf{T}) \leq r(\mathbf{A}). \tag{4.6.9}$$

From (4.6.9), we have $\mathcal{M}(\mathbf{A}\,\mathbf{T}) = \mathcal{M}(\mathbf{A})$. Hence $\mathbf{A} = \mathbf{A}\,\mathbf{T}\,\mathbf{C} = \mathbf{C}'\,\mathbf{T}'\,\mathbf{A}$ for some matrix $\mathbf{C}$. We get (4.6.7), if we pre- and postmultiply (4.6.8) by $\mathbf{C}'$ and $\mathbf{C}$, respectively.

Suppose (b) is given, then $\mathcal{M}(\mathbf{A}\,\mathbf{W}) = \mathcal{M}(\mathbf{A})$. Hence, we have $\mathbf{A}\,\mathbf{W}\,\mathbf{D} = \mathbf{A}$ for some matrix $\mathbf{D}$ such that $\mathbf{A}\,\mathbf{W}\,\mathbf{D} = \mathbf{D}'\,\mathbf{W}\,\mathbf{A}$. Pre- and postmultiplying (4.6.6)(i) by $\mathbf{D}'$ and $\mathbf{D}$, respectively, we get (4.6.7). This completes the proof of the lemma.

Lemma 4.6.2 is used to obtain the necessary and sufficient conditions for two multivariate quadratic forms to be independently distributed.

Lemma 4.6.2 *Let* $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{W}$ *be symmetric matrices of order* $n \times n$. *Consider the following two conditions:*

(a) $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{W}$ *are n.n.d. matrices*

(b) $r(\mathbf{A}\,\mathbf{W}) = r(\mathbf{A})$ *and* $r(\mathbf{B}\,\mathbf{W}) = r(\mathbf{B})$.

If the condition (a) *or* (b) *holds then*

$$\text{(i) } \mathbf{W}\,\mathbf{A}\,\mathbf{W}\,\mathbf{B}\,\mathbf{W} = \mathbf{O}, \ \text{(ii) } \mathbf{W}\,\mathbf{A}\,\mathbf{W}\,\mathbf{B}\,\mathbf{M}' = \mathbf{O} = \mathbf{W}\,\mathbf{B}\,\mathbf{W}\,\mathbf{A}\,\mathbf{M}' \ \text{ and}$$

$$\text{(iii) } \mathbf{M}\,\mathbf{A}\,\mathbf{W}\,\mathbf{B}\,\mathbf{M}' = \mathbf{O} \tag{4.6.10}$$

if and only if

$$\mathbf{A}\,\mathbf{W}\,\mathbf{B} = \mathbf{O} \tag{4.6.11}$$

where $\mathbf{M}$ *is a matrix of order* $p \times n$.

Proof: It is easy to see that (4.6.11) implies (i), (ii) and (iii) of (4.6.10). Let (a) be given, then it follows from Shanbhag (1966) that (4.6.10)(i) and (4.6.11) are equivalent conditions.

Suppose now (b) is given then as shown in Lemma 4.6.1, we have $\mathbf{A}\,\mathbf{W}\,\mathbf{C} = \mathbf{A}$ and $\mathbf{B}\,\mathbf{W}\,\mathbf{D} = \mathbf{B}$ for some matrices $\mathbf{C}$ and $\mathbf{D}$. Pre- and postmultiplying (4.6.10)(i) by $\mathbf{C}'$ and $\mathbf{D}$, respectively, we get (4.6.11).

We now present one of our main theorems. This theorem generalizes Theorem 2.8.5 of Siotani et.al. (1985) for the case of singular Wishart distribution and it is also a multivariate generalization of Theorem 2s of Searle (1971). For related works, we refer to Khatri (1963), Rayner and Livingstone (1965), Khatri (1968) and Shanbhag (1968).

Theorem 4.6.1 *Let* $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$ *be an observation matrix of order* $p \times n$ *such that* $\mathbf{x}_j \sim N_p(\boldsymbol{\mu}_j, w_{jj}\mathbf{V})$ *and* $cov(\mathbf{x}_i, \mathbf{x}_j) = w_{ij}\,\mathbf{V}$ *for* $i, j = 1, \ldots, n$*; that is,* $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ *where* $\mathbf{W} = (w_{ij})$ *and* $\mathbf{M} = [\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_n]$*. Let* $\mathbf{W}$*,* $\mathbf{V}$ *be n.n.d. matrices and* $\mathbf{A}$ *be a symmetric matrix of order* $n \times n$*. Then* $\mathbf{Q}(\mathbf{X}) = \mathbf{X}\mathbf{A}\mathbf{X}' \sim W_p(r(\mathbf{A}), \mathbf{V}; \mathbf{Q}(\mathbf{M}))$ *if and only if* $\mathbf{A}\mathbf{W}\mathbf{A} = \mathbf{A}$*.*

Proof: Let $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ and $\mathbf{A}\mathbf{W}\mathbf{A} = \mathbf{A}$ then from Theorem 4.3.1 and 4.3.3, we have $\mathbf{X} = \mathbf{M} + \mathbf{Y}\mathbf{T}'$ where $\mathbf{Y} \sim N_{p,r}(\mathbf{O}, \mathbf{I}_r \otimes \mathbf{V})$, $\mathbf{O}$ is a matrix zeros of order $p \times r$ and $\mathbf{T}$ is as given in (4.3.4). Rewriting $\mathbf{X}\mathbf{A}\mathbf{X}'$ in terms of $\mathbf{Y}$, we get

$$\begin{aligned}\mathbf{X}\mathbf{A}\mathbf{X}' &= (\mathbf{M} + \mathbf{Y}\mathbf{T}')\mathbf{A}(\mathbf{M} + \mathbf{Y}\mathbf{T}')' \\ &= (\mathbf{M} + \mathbf{Y}\mathbf{T}')\mathbf{A}\mathbf{T}\mathbf{T}'\mathbf{A}(\mathbf{M} + \mathbf{Y}\mathbf{T}')' \\ &= (\mathbf{Y}\mathbf{T}'\mathbf{A}\mathbf{T} + \mathbf{M}\mathbf{A}\mathbf{T})(\mathbf{Y}\mathbf{T}'\mathbf{A}\mathbf{T} + \mathbf{M}\mathbf{A}\mathbf{T})'. \qquad (4.6.12)\end{aligned}$$

Let $r(\mathbf{T}'\mathbf{A}\mathbf{T}) = t$, since the matrix $\mathbf{T}'\mathbf{A}\mathbf{T}$ is idempotent and symmetric, we have $\mathbf{T}'\mathbf{A}\mathbf{T} = \mathbf{H}\mathbf{H}'$ where $\mathbf{H}$ is of order $r \times t$ such that $\mathbf{H}'\mathbf{H} = \mathbf{I}_t$. Hence from (4.6.12), we have

$$\mathbf{X}\mathbf{A}\mathbf{X}' = (\mathbf{Y}\mathbf{H} + \mathbf{M}\mathbf{A}\mathbf{T}\mathbf{H})(\mathbf{Y}\mathbf{H} + \mathbf{M}\mathbf{A}\mathbf{T}\mathbf{H})' \qquad (4.6.13)$$

since $\mathbf{M}\mathbf{A}\mathbf{T} = \mathbf{M}\mathbf{A}\mathbf{T}\mathbf{T}'\mathbf{A}\mathbf{T}$. From Theorem 4.3.3, we have $\mathbf{Y}\mathbf{H} + \mathbf{M}\mathbf{A}\mathbf{T}\mathbf{H} \sim N_{p,t}(\mathbf{M}\mathbf{A}\mathbf{T}\mathbf{H}, \mathbf{I}_t \otimes \mathbf{V})$. Now from Remark 4.4.3, we get $\mathbf{X}\mathbf{A}\mathbf{X}' \sim W_p(r(\mathbf{A}), \mathbf{V}; \Omega)$ where $\Omega = (\mathbf{M}\mathbf{A}\mathbf{T}\mathbf{H})(\mathbf{M}\mathbf{A}\mathbf{T}\mathbf{H})' = \mathbf{M}\mathbf{A}\mathbf{M}' = \mathbf{Q}(\mathbf{M})$. Also note that $r(\mathbf{T}'\mathbf{A}\mathbf{T}) = tr(\mathbf{T}'\mathbf{A}\mathbf{T}) = tr(\mathbf{A}\mathbf{W}) = r(\mathbf{A})$ since we have assumed $\mathbf{A}\mathbf{W}\mathbf{A} = \mathbf{A}$.

To prove the converse, let $\mathbf{V} = \mathbf{U}\,\mathbf{U}'$ and $\mathbf{C} = (\mathbf{U}'\,\mathbf{U})^{-1}\,\mathbf{U}'$ where $\mathbf{U}$ is as given in (4.3.4). Let $\mathbf{Q}(\mathbf{X}) \sim W_p(r(\mathbf{A}), \mathbf{V}; \mathbf{Q}(\mathbf{M}))$ then from Remark 4.4.6, we have

$$\mathbf{C}\,\mathbf{Q}(\mathbf{X})\,\mathbf{C}' \sim W_s(r(\mathbf{A}), \mathbf{I}_s; \mathbf{C}\,\mathbf{Q}(\mathbf{M})\mathbf{C}'). \tag{4.6.14}$$

It follows from Corollary 4.4.3 that

$$tr(\mathbf{C}\,\mathbf{Q}(\mathbf{X})\mathbf{C}') \sim \chi^2(s\; r(\mathbf{A}); tr(\mathbf{C}\,\mathbf{Q}(\mathbf{M})\mathbf{C}')). \tag{4.6.15}$$

From Lemma 4.2.1(f), we obtain

$$\begin{aligned} tr(\mathbf{C}\,\mathbf{Q}(\mathbf{X})\mathbf{C}') &= tr(\mathbf{C}\,\mathbf{X}\,\mathbf{A}\,\mathbf{X}'\mathbf{C}') = tr(\mathbf{A}\,\mathbf{X}'\mathbf{C}'\,\mathbf{C}\,\mathbf{X}) \\ &= vec(\mathbf{X})'\,(\mathbf{A} \otimes \mathbf{C}'\,\mathbf{C})\,vec(\mathbf{X}). \end{aligned} \tag{4.6.16}$$

Let $\boldsymbol{\Sigma} = \mathbf{W} \otimes \mathbf{V}$ then from the definition of matrix normal distribution, we have $vec(\mathbf{X}) \sim N_{pn}\,(vec(\mathbf{M}), \boldsymbol{\Sigma})$. From Corollary 2s.1 of Searle (1971) and Lemma 4.6.1 for $p = 1$, we get from (4.6.15) and (4.6.16) that

$$vec(\mathbf{X})'\,(\mathbf{A} \otimes \mathbf{C}'\,\mathbf{C})\,vec(\mathbf{X}) \sim \chi^2(s\; r(\mathbf{A}); \delta)$$

which is equivalent to

$$(\mathbf{A} \otimes \mathbf{C}'\,\mathbf{C})\,\boldsymbol{\Sigma}\,(\mathbf{A} \otimes \mathbf{C}'\,\mathbf{C}) = \mathbf{A} \otimes \mathbf{C}'\,\mathbf{C} \tag{4.6.17}$$

where $\delta = vec(\mathbf{M})'\,(\mathbf{A} \otimes \mathbf{C}'\,\mathbf{C})\,vec(\mathbf{M})$. It is clear from (4.6.17) that $\mathbf{A}\,\mathbf{W}\,\mathbf{A} = \mathbf{A}$. This completes the proof of the theorem.

Remark 4.6.1 We also have a direct proof of the converse given above. This proof is more informative and is straightforward too. The proof is based on Theorem 4.4.2. We have to show that if $\mathbf{Q}(\mathbf{X}) \sim W_p(r(\mathbf{A}), \mathbf{V}; \mathbf{Q}(\mathbf{M}))$ then $\mathbf{A}\mathbf{W}\mathbf{A} = \mathbf{A}$. Since $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$, from Theorem 4.3.2, we have $\mathbf{X}' \sim N_{n,p}(\mathbf{M}', \mathbf{V} \otimes \mathbf{W})$. Hence $\mathbf{x}_i^* \sim N_n(\mathbf{m}_i^*, v_{ii}\mathbf{W})$ where $\mathbf{x}_i^*$ and $\mathbf{m}_i^*$ are the ith columns of $\mathbf{X}'$ and $\mathbf{M}'$, respectively and v_{ii} is the ith diagonal element of $\mathbf{V}$. It follows from Theorem 4.4.1 that $\mathbf{x}_i^{*\prime}\mathbf{A}\mathbf{x}_i^* \sim v_{ii}\chi^2(r(\mathbf{A}); \delta_{ii})$ where $\delta_{ii} = \mathbf{m}_i^{*\prime}\mathbf{A}\mathbf{m}_i^*$ and v_{ii} is a nonzero diagonal element of $\mathbf{V}$. Note that, we can always find a nonzero v_{ii} since $P(\mathbf{X} = \mathbf{O}) = 0$. It follows from Corollary 2s.1 of Searle (1971) and Lemma 4.6.1 for $p = 1$ that $\mathbf{A}\mathbf{W}\mathbf{A} = \mathbf{A}$.

Remark 4.6.2 If $\mathbf{V}$ is a positive definite matrix or $\mathbf{A}\mathbf{M}' = \mathbf{O}$ then the proof of the above theorem follows from Theorem 4.5.1 and Corollary 4.4.1.

Corollary 4.6.1 *The distribution of* $\mathbf{Q}(\mathbf{X})$ *in Theorem* 4.6.1 *is a central Wishart distribution,* $W_p(r(\mathbf{A}), \mathbf{V})$*, if and only if* $\mathbf{A}\mathbf{M}' = \mathbf{O}$.

Proof: It is easy to see that if $\mathbf{A}\mathbf{M}' = \mathbf{O}$ then $\mathbf{M}\mathbf{A}\mathbf{M}' = \mathbf{O}$. Hence $\mathbf{Q}(\mathbf{X})$ has a central Wishart distribution. To show the converse, let $\mathbf{W} = \mathbf{T}\mathbf{T}'$ where $\mathbf{T}$ is given in (4.3.4) then $\mathbf{M}\mathbf{A}\mathbf{M}' = \mathbf{O}$ implies $\mathbf{M}\mathbf{A}\mathbf{T} = \mathbf{O}$. Which in turn implies that $\mathbf{M}\mathbf{A}\mathbf{T}\mathbf{T}'\mathbf{A} = \mathbf{O}$. Hence $\mathbf{A}\mathbf{M}' = \mathbf{O}$ since $\mathbf{A}\mathbf{W}\mathbf{A} = \mathbf{A}$.

Corollary 4.6.2 *Under the assumptions of Theorem* 4.6.1*, we have* $\mathbf{Q}(\mathbf{X}) \sim d\,W_p(r(\mathbf{A}), \mathbf{V}; \Omega)$ *if and only if* $\mathbf{A}\mathbf{W}\mathbf{A} = d\,\mathbf{A}$ *where* $d > 0$ *and* $\Omega = \frac{1}{d}\mathbf{Q}(\mathbf{M})$.

In our next theorem, we present a multivariate generalization of Theorem 4s of Searle (1971). For earlier works, we refer to Khatri (1962, 1963), Shanbhag (1966), Good (1966) and Styan (1969).

Theorem 4.6.2 *Let* $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ *where* $\mathbf{X}$, $\mathbf{M}$, $\mathbf{W}$ *and* $\mathbf{V}$ *are as defined in Theorem* 4.6.1. *Let* $\mathbf{Q}_1(\mathbf{X}) = \mathbf{X}\,\mathbf{A}\,\mathbf{X}'$ *and* $\mathbf{Q}_2(\mathbf{X}) = \mathbf{X}\,\mathbf{B}\,\mathbf{X}'$ *where* $\mathbf{A}$ *and* $\mathbf{B}$ *are symmetric matrices of order* $n \times n$. *Consider the conditions:* (a) $\mathbf{A}$ *and* $\mathbf{B}$ *are n.n.d. matrices and* (b) $r(\mathbf{A}\,\mathbf{W}) = r(\mathbf{A})$ *and* $r(\mathbf{B}\,\mathbf{W}) = r(\mathbf{B})$. *If the condition* (a) *or* (b) *holds then* $\mathbf{Q}_1(\mathbf{X})$ *and* $\mathbf{Q}_2(\mathbf{X})$ *are independently distributed if and only if* $\mathbf{A}\,\mathbf{W}\,\mathbf{B} = \mathbf{O}$.

Proof: Let $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ then from Theorem 4.3.2, we have $\mathbf{X}' \sim N_{n,p}(\mathbf{M}', \mathbf{V} \otimes \mathbf{W})$. Hence $\mathbf{x}_i^* \sim N_n(\mathbf{m}_i^*, v_{ii}\mathbf{W})$ where $\mathbf{x}_i^*$ and $\mathbf{m}_i^*$ are the ith columns of $\mathbf{X}'$ and $\mathbf{M}'$, respectively and v_{ii} is the ith nonzero diagonal element of $\mathbf{V}$. If $\mathbf{Q}_1(\mathbf{X})$ and $\mathbf{Q}_2(\mathbf{X})$ are independently distributed then their ith diagonal elements, $\mathbf{x}_i^{*\prime}\,\mathbf{A}\,\mathbf{x}_i^*$ and $\mathbf{x}_i^{*\prime}\,\mathbf{B}\,\mathbf{x}_i^*$, are also independently distributed. Let (a) or (b) be given, then it follows from Theorem 4s of Searle (1971) and Lemma 4.6.2 with $p = 1$ that $\mathbf{A}\,\mathbf{W}\,\mathbf{B} = \mathbf{O}$.

To prove the converse, let $\mathbf{A}\,\mathbf{W}\,\mathbf{B} = \mathbf{O}$ then condition (i), (ii) and (iii) of Lemma 4.6.2 are trivially satisfied. Since $\mathbf{W} = \mathbf{T}\,\mathbf{T}'$, it is easy to see that conditions (i), (ii) and (iii) are equivalent to

$$\text{(ia) } \mathbf{A}_1\,\mathbf{B}_1 = \mathbf{O}, \quad \text{(iia) } \mathbf{L}_1\,\mathbf{B}_1 = \mathbf{O} = \mathbf{K}_1\,\mathbf{A}_1 \text{ and (iiia) } \mathbf{K}_1\,\mathbf{L}_1' = \mathbf{O} \tag{4.6.18}$$

where $\mathbf{A}_1 = \mathbf{T}'\mathbf{A}\,\mathbf{T}$, $\mathbf{B}_1 = \mathbf{T}'\mathbf{B}\,\mathbf{T}$, $\mathbf{L}_1 = \mathbf{M}\mathbf{A}\mathbf{T}$ and $\mathbf{K}_1 = \mathbf{M}\mathbf{B}\mathbf{T}$. Since $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W}\otimes\mathbf{V})$, from Theorem 4.3.1 and 4.3.3, we have $\mathbf{X} = \mathbf{M} + \mathbf{Y}\,\mathbf{T}'$ where $\mathbf{Y} \sim N_{p,r}(\mathbf{O}, \mathbf{I}_r \otimes \mathbf{V})$, $\mathbf{O}$ is a matrix zeros of order $p \times r$ and $\mathbf{T}$ is as given in (4.3.4). If we represent $\mathbf{X}\,\mathbf{A}\,\mathbf{X}'$ and $\mathbf{X}\,\mathbf{B}\,\mathbf{X}'$ in terms of $\mathbf{Y}$, we get

$$\mathbf{Q}_1(\mathbf{X}) = \mathbf{Y}\,\mathbf{A}_1\,\mathbf{Y}' + \mathbf{Y}\,\mathbf{L}_1' + \mathbf{L}_1\,\mathbf{Y}' + \mathbf{C}_1 \tag{4.6.19}$$

and

$$\mathbf{Q}_2(\mathbf{X}) = \mathbf{Y}\,\mathbf{B}_1\,\mathbf{Y}' + \mathbf{Y}\,\mathbf{K}_1' + \mathbf{K}_1\,\mathbf{Y}' + \mathbf{D}_1 \tag{4.6.20}$$

where $\mathbf{C}_1 = \mathbf{M}\,\mathbf{A}\,\mathbf{M}'$ and $\mathbf{D}_1 = \mathbf{M}\,\mathbf{B}\,\mathbf{M}'$. From Lemma 1 of Khatri (1962), we get

$$\mathbf{L}_1 = [\mathbf{L}_1^* \;\; \mathbf{O}]\,\mathbf{H}' \quad \text{and} \quad \mathbf{K}_1 = [\mathbf{O} \;\; \mathbf{K}_1^*]\,\mathbf{H}'; \tag{4.6.21}$$

$$\mathbf{A}_1 = \mathbf{H}\begin{bmatrix} \mathbf{A}_1^* & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}\mathbf{H}' \quad \text{and} \quad \mathbf{B}_1 = \mathbf{H}\begin{bmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{B}_1^* \end{bmatrix}\mathbf{H}' \tag{4.6.22}$$

where $\mathbf{H}$ is of order $r \times (t+u)$ such that $\mathbf{H}'\mathbf{H} = \mathbf{I}_t$, $t = r([\mathbf{A}_1 \;\; \mathbf{L}_1'])$, $u = r([\mathbf{B}_1 \;\; \mathbf{K}_1'])$, $\mathbf{A}_1^*$ and $\mathbf{B}_1^*$ are symmetric matrices of order $t \times t$ and $u \times u$, respectively ; $\mathbf{L}_1^*$ is of order $p \times t$, $\mathbf{K}_1^*$ of order $p \times u$ and $\mathbf{O}$ is matrix of zeros of an appropriate order. Let $\mathbf{Y}^* = \mathbf{Y}\,\mathbf{H}$ then $\mathbf{Y}^* \sim N_{p,t+u}(\mathbf{O}, \mathbf{I}_{t+u} \otimes \mathbf{V})$. Let $\mathbf{Y}^* = [\mathbf{Y}_1^* \;\; \mathbf{Y}_2^*]$ where $\mathbf{Y}_1^*$ is of order $p \times t$ and $\mathbf{Y}_2^*$ is of order $p \times u$. It is easy to see that $\mathbf{Y}_1^*$ and $\mathbf{Y}_2^*$ are independently distributed. Now it follows from (4.6.19), (4.6.20), (4.6.21) and (4.6.22) that

$$\mathbf{Q}_1(\mathbf{X}) = \mathbf{Y}_1^* \mathbf{A}_1^* \mathbf{Y}_1^{*'} + \mathbf{Y}_1^* \mathbf{L}_1^{*'} + \mathbf{L}_1^* \mathbf{Y}_1^{*'} + \mathbf{C}_1 \tag{4.6.23}$$

and

$$\mathbf{Q}_2(\mathbf{X}) = \mathbf{Y}_2^* \mathbf{B}_1^* \mathbf{Y}_2^{*'} + \mathbf{Y}_2^* \mathbf{K}_1^{*'} + \mathbf{K}_1^* \mathbf{Y}_2^{*'} + \mathbf{D}_1. \tag{4.6.24}$$

Since $\mathbf{Q}_1(\mathbf{X})$ depends only on $\mathbf{Y}_1^*$ and $\mathbf{Q}_2(\mathbf{X})$ on $\mathbf{Y}_2^*$, we have $\mathbf{Q}_1(\mathbf{X})$ and $\mathbf{Q}_2(\mathbf{X})$ independently distributed. This completes the proof of the theorem.

Remark 4.6.3 If $\mathbf{V}$ is a positive definite matrix or $\mathbf{A}\,\mathbf{M}' = \mathbf{O}$ then the proof of the above theorem follows from Theorem 4.5.2.

The proof of the following corollary follows from the above theorem and Corollary 4.5.3.

Corollary 4.6.3 *Let* $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ *where* $\mathbf{X}$, $\mathbf{M}$, $\mathbf{W}$ *and* $\mathbf{V}$ *are as defined in Theorem* 4.6.1. *Then* $\mathbf{X}\,\mathbf{A}\,\mathbf{X}'$ *and* $\mathbf{X}\,\mathbf{L}'$ *are independently distributed if and only if* $\mathbf{A}\,\mathbf{W}\,\mathbf{L}' = \mathbf{O}$ *where* $\mathbf{A}$ *is an n.n.d. matrix of order* $n \times n$ *and* $\mathbf{L}$ *is of order* $p \times n$.

Remark 4.6.4 Let $\mathbf{Q}_1(\mathbf{X})$ and $\mathbf{Q}_2(\mathbf{X})$ be as defined in Theorem 4.6.2. Let $\mathbf{Q}_1(\mathbf{X}) \sim W_p(r(\mathbf{A}), \mathbf{V}; \mathbf{Q}_1(\mathbf{M}))$ and $\mathbf{Q}_2(\mathbf{X}) \sim W_p(r(\mathbf{B}), \mathbf{V}; \mathbf{Q}_2(\mathbf{M}))$ then $\mathbf{Q}_1(\mathbf{X})$ and $\mathbf{Q}_2(\mathbf{X})$ are independently distributed if and only if $\mathbf{A}\,\mathbf{W}\,\mathbf{B} = \mathbf{O}$.

Remark 4.6.5 It is interesting to note that the necessary and sufficient conditions in Theorem 4.6.1 and 4.6.2 do not depend on $\mathbf{M}$. Hence without loss of generality we can assume $\mathbf{M} = \mathbf{O}$.

We now present a multivariate analogue of the Cochran's theorem, see Cochran (1934). The case $p = 1$ for singular normal random vector is discussed by Styan (1969), Rao and Mitra (1971, section 9.3) and for nonsingular normal random vector by several authors including Graybill and Marsaglia (1957), Banerjee (1964) and Loynes (1966). Khatri (1962) has given a multivariate generalization of the Cochran's theorem for nonsingular matrix normal distribution.

Theorem 4.6.3 *Let* $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ *where* $\mathbf{X}$, $\mathbf{M}$, $\mathbf{W}$ *and* $\mathbf{V}$ *are as defined in Theorem* 4.6.1. *Let* $\mathbf{A}_i$ $(i = 1, \ldots, k)$ *and* $\mathbf{A}$ *be symmetric matrices of order* $n \times n$ *such that* $\mathbf{A} = \sum_{i=1}^{k} \mathbf{A}_i$. *Consider the following conditions:*

(a_1) $\mathbf{X}\,\mathbf{A}_i\,\mathbf{X}' \sim W_p(r(\mathbf{A}_i), \mathbf{V}; \Omega_i)$ *where* $\Omega_i = \mathbf{M}\,\mathbf{A}_i\,\mathbf{M}'$ *for* $i = 1, \ldots, k$

(a_2) $\mathbf{X}\,\mathbf{A}_i\,\mathbf{X}'$ *and* $\mathbf{X}\,\mathbf{A}_j\,\mathbf{X}'$ *are mutually independent for* $i \neq j = 1, \ldots, k$

(a_3) $\mathbf{X}\,\mathbf{A}\,\mathbf{X}' \sim W_p(r(\mathbf{A}), \mathbf{V}; \Omega)$ *where* $\Omega = \mathbf{M}\,\mathbf{A}\,\mathbf{M}'$

(b_1) $\mathbf{A}_i\,\mathbf{W}\,\mathbf{A}_i = \mathbf{A}_i$ *for* $i = 1, \ldots, k$

(b_2) $\mathbf{A}_i\,\mathbf{W}\,\mathbf{A}_j = \mathbf{O}$ *for* $i \neq j = 1, \ldots, k$

(b_3) $\mathbf{A}\,\mathbf{W}\,\mathbf{A} = \mathbf{A}$

(b_4) $\sum_{i=1}^{k} r(\mathbf{A}_i) = r(\mathbf{A})$.

Then

(1) *any two of the three conditions* (a_1), (a_2), (a_3) *or*

(2) *any two of the three conditions* (b_1), (b_2), (b_3) *or*

(3) *any two conditions* (a_i) *and* (b_j) *for* $i \neq j = 1, 2, 3$ *or*

(4) (b_3) *and* (b_4) *or*

(5) (a_3) *and* (b_4)

are necessary and sufficient for all the remaining conditions: (a_1) - (b_4).

Proof: The proof is based on Theorems 4.6.1, 4.6.2 and Theorem 1 of Graybill and Marsaglia (1957). We only prove (5), that is, (a_3) and (b_4) are necessary and sufficient for all the remaining conditions. We show that (a_3) and (b_4) imply all the remaining conditions since converse is trivially true.

Let (a_3) and (b_4) be given, then from Theorem 4.6.1, we get (b_3). Let $\mathbf{B} = \mathbf{T}'\mathbf{A}\,\mathbf{T}$ and $\mathbf{B}_i = \mathbf{T}'\mathbf{A}_i\mathbf{T}$ for $i = 1, \ldots, k$ where $\mathbf{W} = \mathbf{T}\mathbf{T}'$ and $\mathbf{T}$ is as defined in (4.3.4) then $\mathbf{B} = \sum_{i=1}^{k} \mathbf{B}_i$. Also from (b_3), we get $\mathbf{B}^2 = \mathbf{B}$. On using condition (b_4), we have

$$r(\mathbf{A}) = tr(\mathbf{A}\,\mathbf{W}) = tr(\mathbf{B}) = r(\mathbf{B}) \leq \sum_{i=1}^{k} r(\mathbf{B}_i) \leq \sum_{i=1}^{k} r(\mathbf{A}_i) = r(\mathbf{A}) \tag{4.6.25}$$

since $tr(\mathbf{A}\,\mathbf{W}) = r(\mathbf{A}\,\mathbf{W}) = r(\mathbf{A})$ from (b_3). Hence $r(\mathbf{B}) = \sum_{i=1}^{k} r(\mathbf{B}_i)$. From Theorem 1 of Graybill and Marsaglia (1957), we get $\mathbf{B}_i^2 = \mathbf{B}_i$ and $\mathbf{B}_i\mathbf{B}_j = \mathbf{O}$ for $i \neq j = 1, \ldots, k$. It follows from the following lemma that $r(\mathbf{B}_i) = r(\mathbf{A}_i)$ for $i = 1, \ldots, k$. Hence from Lemmas 4.6.1 and 4.6.2, we get (b_1) and (b_2) and (a_1) and (a_2) follow from Theorems 4.6.1 and 4.6.2.

Lemma 4.6.3 *Let* $\mathbf{A}_i$*,* $\mathbf{B}_i$*,* $\mathbf{A}$ *and* $\mathbf{B}$ *be as defined in Theorem* 4.6.3. *Let* $r(\mathbf{A}) = \sum_{i=1}^{k} r(\mathbf{A}_i) = \sum_{i=1}^{k} r(\mathbf{B}_i) = r(\mathbf{B})$ *then* $r(\mathbf{B}_i) = r(\mathbf{A}_i)$ *for* $i = 1, \ldots, k$.

Proof: It is obvious from the definition of $\mathbf{B}_i$'s that $r(\mathbf{B}_i) \leq r(\mathbf{A}_i)$ for $i = 1, \ldots, k$. Also, $r(\mathbf{B}_i) = r(\mathbf{B}) - \sum_{j\neq i=1}^{k} r(\mathbf{B}_j) = r(\mathbf{A}) - \sum_{j\neq i=1}^{k} r(\mathbf{B}_j) \geq r(\mathbf{A}) - \sum_{j\neq i=1}^{k} r(\mathbf{A}_j) = r(\mathbf{A}_i)$. Hence, repeating the same argument we can show that $r(\mathbf{B}_i) \geq r(\mathbf{A}_i)$ for $i = 1, \ldots, k$.

4.7 Statistical Applications

In this section, we present multivariate generalizations of the applications presented in Chapter 3. Let $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ where $\mathbf{X}$, $\mathbf{M}$, $\mathbf{W}$ and $\mathbf{V}$ are as defined in Theorem 4.6.1. Our goal in this section is to characterize the class of all n.n.d. $\mathbf{W}$'s such that the distribution of matrices corresponding to various sum of squares and cross products in MANOVA problems remain invariant, that is, the distributions are preserved except for a scale factor.

Basu et al. (1974) defined a covariance structure called simply equicorrelated by taking $\mathbf{W} = (1-\rho)\,\mathbf{I} + \rho\,\mathbf{e}\,\mathbf{e}'$ and derived the distributions of the sample mean vector

$\bar{\mathbf{x}}$, sample covariance matrix $\mathbf{S}$ and the Hotelling's T^2 statistic. They showed that for this choice of $\mathbf{W}$ the sample mean vector and covariance matrix are independently distributed and their distributions are also preserved except for a constant factor. The converse of this result is given in Theorem 4.7.1. In Theorem 4.7.1, we show that $\bar{\mathbf{x}}$ and $\mathbf{S}$ are independently distributed and their distributions remain invariant if and only if the observations are equicorrelated (or simply equicorrelated as mentioned above). As a simple consequence of Theorem 4.7.1, we show in Theorem 4.7.2 that the distribution of commonly used one sample Hotelling's T^2 statistic remains invariant except for a constant factor when the observations in the sample are equicorrelated. We also show that the distribution of usual two sample Hotelling's T^2 statistic remains invariant if the observations in one of the samples are positively equicorrelated and those in the other sample are negatively equicorrelated with the same correlation in absolute value. This result is contained in Theorem 4.7.3.

The next two examples give a characterization of the class of all n.n.d. $\mathbf{W}$'s such that the distribution of a given multivariate quadratic form remains invariant.

Example 4.7.1 Let $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ where $\mathbf{X}$, $\mathbf{M}$, $\mathbf{W}$ and $\mathbf{V}$ are as defined in Theorem 4.6.1. Let $\mathbf{A}$ be an n.n.d. matrix of order $n \times n$. Then it follows from Theorem 4.6.1 that $\mathbf{X}\,\mathbf{A}\,\mathbf{X}' \sim W_p(r(\mathbf{A}), \mathbf{V}; \mathbf{Q}(\mathbf{M}))$ if and only if $\mathbf{A}\,\mathbf{W}\,\mathbf{A} = \mathbf{A}$ which is equivalent to characterizing the class of all n.n.d. g-inverses of the matrix $\mathbf{A}$ and such a class is given in Example 3.2.1.

Example 4.7.2 In Example 4.7.1, $\mathbf{Q}(\mathbf{X}) \sim d\,W_p(r(\mathbf{A}), \mathbf{V}; \Omega)$ if and only if $\mathbf{A}\,\mathbf{W}\,\mathbf{A} = d\,\mathbf{A}$ where $d > 0$ and $\Omega = \frac{1}{d}\mathbf{Q}(\mathbf{M})$. Therefore, from Theorem 3.2.2 we can obtain the class of all n.n.d. $\mathbf{W}$'s such that $\mathbf{A}\,\mathbf{W}\,\mathbf{A} = d\,\mathbf{A}$.

In the next example we examine an invariance property of the distribution of sample covariance matrix $\mathbf{S}$.

Example 4.7.3 Let $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ where $\mathbf{X}$, $\mathbf{M}$, $\mathbf{W}$ and $\mathbf{V}$ are as defined in Theorem 4.6.1. Then for any $d > 0$, we have $(n-1)\,\mathbf{S} \sim d\,W_p(n-1, \mathbf{V}; \Omega)$ if and only if $\mathbf{W} \in \mathcal{G}_{d,n}$ where class $\mathcal{G}_{d,n}$ is as defined in Remark 3.2.3, $\mathbf{S}$ is defined in (4.3.3) and $\Omega = \frac{1}{d}\mathbf{M}\,(\mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}')\,\mathbf{M}'$. From Corollary 4.6.2, we get $(n-1)\,\mathbf{S} \sim d\,W_p(n-1, \mathbf{V}; \Omega)$ if and only if $(\mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}')\,\mathbf{W}\,(\mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}') = d\,(\mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}')$ which is true if and only if $\mathbf{W} \in \mathcal{G}_{d,n}$.

We now characterize the class of all covariance matrices $\mathbf{W}$ such that the sample mean vector and covariance matrix are independently distributed and their distributions are also preserved except for a constant factor.

Theorem 4.7.1 *Let $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ where $\mathbf{X}$, $\mathbf{M}$, $\mathbf{W}$ and $\mathbf{V}$ are as defined in Theorem 4.6.1. Then $(n-1)\,\mathbf{S} \sim d\,W_p(n-1, \mathbf{V}; \Omega)$ and $\bar{\mathbf{x}}$ is independent of $\mathbf{S}$ if and only if $\mathbf{W} = d\,(\mathbf{I} - \frac{(1-c)}{n}\mathbf{e}\,\mathbf{e}')$ for some $c \geq 0$, $d > 0$ and Ω is as given in Example 4.7.3.*

Proof: From Example 4.7.3 for any $d > 0$, we have $(n-1)\,\mathbf{S} \sim d\,W_p(n-1, \mathbf{V}; \Omega)$ if and only if $\mathbf{W} \in \mathcal{G}_{d,n}$. From Corollary 4.6.3, $\bar{\mathbf{x}}$ is independent of $\mathbf{S}$ if and only if

$$(\mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}')\,\mathbf{W}\,\mathbf{e} = \mathbf{0}. \tag{4.7.26}$$

It is easy to check that $\mathbf{W} \in \mathcal{G}_{d,n}$ and satisfies (4.7.26) if and only if

$$(\mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}')\,\mathbf{a} = \mathbf{0} \tag{4.7.27}$$

which in turn holds if and only if $\mathbf{a} = c\,\mathbf{e}$ where $c = \bar{a} \geq 0$. Thus $(n-1)\,\mathbf{S} \sim d\,W_p(n-1, \mathbf{V};\, \Omega)$ and $\bar{\mathbf{x}}$ is independent of $\mathbf{S}$ if and only if

$$\begin{aligned}
\mathbf{W} &= d\left((\mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}') + \frac{1}{n}(\mathbf{e}\,\mathbf{a}' + \mathbf{a}\,\mathbf{e}') - \frac{\bar{a}}{n}\mathbf{e}\,\mathbf{e}'\right) \\
&= d\left((\mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}') + \frac{2c}{n}\mathbf{e}\,\mathbf{e}' - \frac{c}{n}\mathbf{e}\,\mathbf{e}'\right) \\
&= d\left(\mathbf{I} - \frac{(1-c)}{n}\mathbf{e}\,\mathbf{e}'\right)
\end{aligned} \tag{4.7.28}$$

where $c \geq 0$. This completes the proof of the theorem.

In the following theorem, we show that the distribution of commonly used one sample Hotelling's T^2 statistic remains invariant except for a constant factor when the observations in the sample are equicorrelated.

Theorem 4.7.2 *Let* $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ *where* $\mathbf{X}$, $\mathbf{M}$, $\mathbf{W}$ *and* $\mathbf{V}$ *are as defined in Theorem* 4.6.1. *Let* $\mathbf{M} = \boldsymbol{\mu}\,\mathbf{e}'$ *where* $\boldsymbol{\mu}$ *is a vector of order* $p \times 1$, $\mathbf{V}$ *be a positive definite matrix and* $n - 1 \geq p$. *Then*

$$\frac{n\,(\bar{\mathbf{x}} - \boldsymbol{\mu})'\,\mathbf{S}^{-1}\,(\bar{\mathbf{x}} - \boldsymbol{\mu})}{n-1}\,\frac{n-p}{p} \sim c\,F(p, n-p) \tag{4.7.29}$$

if $\mathbf{W} = d(\mathbf{I} - \frac{(1-c)}{n}\mathbf{e}\,\mathbf{e}')$ *for some* $c \geq 0$ *and* $d > 0$.

Proof: It follows from Theorem 4.7.1 that for any $d > 0$, $(n-1)\,\mathbf{S} \sim d\,W_p(n-1, \mathbf{V})$ and $\bar{\mathbf{x}}$ is independent of $\mathbf{S}$ if and only if

$$\mathbf{W} = d(\mathbf{I} - \frac{(1-c)}{n}\mathbf{e}\,\mathbf{e}') \tag{4.7.30}$$

for some $c \geq 0$. From Theorem 4.3.3, we have $\bar{\mathbf{x}} \sim N_p(\boldsymbol{\mu}, \frac{d\,c}{n}\mathbf{V})$. Therefore, if $c = 0$ then $\bar{\mathbf{x}}$ is degenerate at $\boldsymbol{\mu}$ and (4.7.29) trivially holds. If $c > 0$ then the result follows from Theorem 4.4.4.

Remark 4.7.1 It is easy to see that if $c \neq 1$ then all the observations are either positively equicorrelated or negatively equicorrelated according as $c > 1$ or $0 \leq c < 1$. Hence from the above theorem we can see that if in a sample all the observations are either positively equicorrelated or negatively equicorrelated then the distribution of usual one sample Hotelling's T^2 statistic is changed only by a constant factor. If $c = 1$ then observations are independent and in that case we have an exact distribution.

We now examine the invariance property of two sample Hotelling's T^2 statistic. In the following theorem we show that the distribution of usual two sample Hotelling's T^2 statistic remains invariant if the observations in one of the samples are positively equicorrelated and those in the other sample are negatively equicorrelated with the same correlation in absolute value.

Theorem 4.7.3 *Let* $\mathbf{X}_k = \left[\mathbf{x}_{k_1}, \mathbf{x}_{k_2}, \ldots, \mathbf{x}_{k_{n_k}}\right]$ *be an observation matrix of order* $p \times n_k$ *such that* $\mathbf{x}_{k_j} \sim N_p(\boldsymbol{\mu}_k, w_{k_{jj}} \mathbf{V})$ *and* $cov(\mathbf{x}_{k_i} \mathbf{x}_{k_j}) = w_{k_{ij}} \mathbf{V}$ *for* $i, j = 1, \ldots, n_k$*; that is,* $\mathbf{X}_k \sim N_{p,n_k}\left(\boldsymbol{\mu}_k \mathbf{e}'_{n_k}, \mathbf{W}_k \otimes \mathbf{V}\right)$ *where* $\mathbf{W}_k = (w_{k_{ij}})$ *for* $k = 1, 2$*. Let* $\mathbf{W}_1$ *and* $\mathbf{W}_2$ *be n.n.d. matrices,* $\mathbf{V}$ *be a positive definite matrix,* $\mathbf{X}_1$ *and* $\mathbf{X}_2$ *be mutually independent and* $n_1 + n_2 - 2 \geq p$*. Let* $\bar{\mathbf{x}}_1$*,* $\bar{\mathbf{x}}_2$ *be the sample mean vectors and* $\mathbf{S}_1$*,* $\mathbf{S}_2$ *be the sample covariance matrices of* $\mathbf{X}_1$ *and* $\mathbf{X}_2$*, respectively. Also, let* $\mathbf{S}_p = [(n_1 - 1)\mathbf{S}_1 + (n_2 - 1)\mathbf{S}_2]/(n_1 + n_2 - 2)$ *and* $\boldsymbol{\theta} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$*. Then*

$$\frac{n_1 n_2}{n_1 + n_2} \frac{((\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) - \boldsymbol{\theta})' \, \mathbf{S}_p^{-1} ((\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) - \boldsymbol{\theta})}{n_1 + n_2 - 2} \frac{(n_1 + n_2 - p - 1)}{p} \sim F(p, n_1 + n_2 - p - 1) \qquad (4.7.31)$$

if

$$\mathbf{W}_1 = d\,(\mathbf{I}_{n_1} + \beta\, \mathbf{e}_{n_1} \mathbf{e}'_{n_1}) \ \ \text{and} \ \ \mathbf{W}_2 = d\,(\mathbf{I}_{n_2} - \beta\, \mathbf{e}_{n_2} \mathbf{e}'_{n_2}) \qquad (4.7.32)$$

for some constants β *and* d *such that* $d > 0$ *and* $-1/n_1 < \beta < 1/n_2$*.*

Proof: If $d > 0$ then it follows from Theorem 4.7.1 that $(n_k - 1)\,\mathbf{S}_k \sim d\,W_p(n_k - 1, \mathbf{V})$ and $\bar{\mathbf{x}}_k$ is independent of $\mathbf{S}_k$ if and only if

$$\mathbf{W}_k = d\,(\mathbf{I}_{n_k} - \frac{(1 - c_k)}{n_k} \mathbf{e}_{n_k} \mathbf{e}'_{n_k}) \qquad (4.7.33)$$

where $c_k \geq 0$ for $k = 1, 2$. From Theorems 4.4.3 and 4.3.3, we get

$$\frac{n_1 + n_2 - 2}{d} \,\mathbf{S}_p \sim W_p(n_1 + n_2 - 2, \mathbf{V}) \qquad (4.7.34)$$

and

$$\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2 - \boldsymbol{\theta} \sim N_p\left(\mathbf{0}, d\,(\frac{c_1}{n_1} + \frac{c_2}{n_2})\,\mathbf{V}\right) \tag{4.7.35}$$

where $\boldsymbol{\theta} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$. Thus for $c_1 > 0$ and $c_2 > 0$, from Theorem 4.4.4, we get

$$\frac{n_1\, n_2}{n_1 + n_2}\, \frac{((\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) - \boldsymbol{\theta})'\, \mathbf{S}_p^{-1}\, ((\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) - \boldsymbol{\theta})}{n_1 + n_2 - 2}\, \frac{(n_1 + n_2 - p - 1)}{p} \sim \frac{c_1\, n_2 + c_2\, n_1}{n_1 + n_2}\, F(p, n_1 + n_2 - p - 1) \tag{4.7.36}$$

if $\mathbf{W}_k$'s are given by (4.7.33). Note that (4.7.36) is also true for $c_1 = 0$ or $c_2 = 0$. If we choose $c_1 = n_1\,\beta + 1$ and $c_2 = 1 - n_2\,\beta$ where $-1/n_1 < \beta < 1/n_2$, we get (4.7.32) and in that case (4.7.31) also holds.

Remark 4.7.2 If $c_1 = c_2 = c$ where $c \geq 0$ then from (4.7.33), we get

$$\mathbf{W}_k = d\,(\mathbf{I}_{n_k} - \frac{(1-c)}{n_k}\, \mathbf{e}_{n_k}\, \mathbf{e}_{n_k}') \tag{4.7.37}$$

for $k = 1, 2$ and in that case from (4.7.36) we have

$$\frac{n_1\, n_2}{n_1 + n_2}\, \frac{((\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) - \boldsymbol{\theta})'\, \mathbf{S}_p^{-1}\, ((\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) - \boldsymbol{\theta})}{n_1 + n_2 - 2}\, \frac{(n_1 + n_2 - p - 1)}{p} \sim c\, F(p, n_1 + n_2 - p - 1). \tag{4.7.38}$$

It is now clear from Remark 4.7.2 that if the observations in both the samples are positively equicorrelated ($c > 1$) or negatively equicorrelated ($0 \leq c < 1$) then the distribution of usual two sample Hotelling's T^2 statistic is changed only by a constant factor. On the other hand, if the observations in one of the samples are positively equicorrelated and those in the other sample are negatively equicorrelated with the same correlation in absolute value then the distribution remains the same.

We now present some applications related to MANOVA problems. The following theorem is useful in characterizing the class of all covariance matrices $\mathbf{W}$ such that the distributions of various matrices of sum of squares and cross products remain invariant.

Theorem 4.7.4 *Let $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$ be symmetric and idempotent matrices of order $n \times n$ such that $\mathbf{A}_i \mathbf{A}_j = \mathbf{O}$ for all $i \neq j$. Let $\mathbf{A} = \sum_{i=1}^{k} \mathbf{A}_i$ and $\mathbf{B} = \sum_{i=1}^{k} c_i \mathbf{A}_i$ where $c_i > 0$ for $1 \leq i \leq k$. Let $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ where $\mathbf{X}$, $\mathbf{M}$, $\mathbf{W}$ and $\mathbf{V}$ are as defined in Theorem 4.6.1. Let $\mathbf{Q}_i(\mathbf{X}) = \mathbf{X} \mathbf{A}_i \mathbf{X}'$ for $1 \leq i \leq k$. Then the multivariate quadratic forms $\mathbf{Q}_i(\mathbf{X})$'s are pairwise independent and distributed as $c_i W_p\left(r(\mathbf{A}_i), \mathbf{V}; \frac{1}{c_i} \mathbf{Q}_i(\mathbf{M})\right)$ for $1 \leq i \leq k$ if and only if $\mathbf{A} \mathbf{W} \mathbf{A} = \mathbf{B}$.*

Proof: From Example 4.7.2 and Theorem 4.6.2, we have $\mathbf{Q}_i(\mathbf{X})$'s are pairwise independent and distributed as $c_i W_p\left(r(\mathbf{A}_i), \mathbf{V}; \frac{1}{c_i} \mathbf{Q}_i(\mathbf{M})\right)$ for $1 \leq i \leq k$ if and only if

$$\mathbf{A}_i \mathbf{W} \mathbf{A}_j = \begin{cases} c_i \mathbf{A}_i & \text{if } i = j \\ \mathbf{O} & \text{if } i \neq j. \end{cases} \tag{4.7.39}$$

The theorem now follows from Lemma 3.2.2.

Remark 4.7.3 In the above theorem, for a given matrix $\mathbf{A}$, the class of all n.n.d. $\mathbf{W}$'s satisfying (4.7.39) is given by Theorem 3.2.2 and it contains the class of covariance matrices considered by Pavur (1987) in Corollary 1.

In all MANOVA problems the total corrected sum of squares and cross products matrix is decomposed into two or more orthogonal sum of squares and cross products matrices. Typically, for testing a sequence of k orthogonal hypotheses H_{oi} ($i = 1, \ldots, k$), we get the following decomposition of the total corrected sum of squares and cross products matrix

$$\mathbf{X}(\mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}')\mathbf{X}' = \mathbf{H}_1 + \cdots + \mathbf{H}_k + \mathbf{E} \tag{4.7.40}$$

where $\mathbf{H}_i$ represents the sum of squares and cross products matrix used to test H_{oi} and $\mathbf{E}$ is the residual or error sum of squares and cross products matrix. It is well known that the underlying matrices of the quadratic forms $\mathbf{H}_i$'s and $\mathbf{E}$ are idempotent and mutually orthogonal. Therefore, the invariance property of the distributions of quadratic forms in a MANOVA table follows from the theorem given below.

Theorem 4.7.5 *Let $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$ be symmetric and idempotent matrices of order $n \times n$ such that $\mathbf{A}_i\mathbf{A}_j = \mathbf{O}$ for all $i \neq j$. Let $\sum_{i=1}^{k} \mathbf{A}_i = \mathbf{A}^*$ and $\mathbf{B} = \sum_{i=1}^{k} c_i\,\mathbf{A}_i$ where $\mathbf{A}^*$ is the centering matrix and $c_i > 0$ for $1 \leq i \leq k$. Let $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ where $\mathbf{X}$, $\mathbf{M}$, $\mathbf{W}$ and $\mathbf{V}$ are as defined in Theorem 4.6.1. Let $\mathbf{Q}_i(\mathbf{X}) = \mathbf{X}\,\mathbf{A}_i\,\mathbf{X}'$ for $1 \leq i \leq k$. Then the multivariate quadratic forms $\mathbf{Q}_i(\mathbf{X})$'s are pairwise independent* *and distributed as* $c_i W_p \left(r(\mathbf{A}_i), \mathbf{V}; \frac{1}{c_i}\mathbf{Q}_i(\mathbf{M})\right)$ *for* $1 \le i \le k$ *if and only if* $\mathbf{W}$ *is of the form* (3.2.23) *where* $\mathbf{a}$ *is an arbitrary vector satisfying* (3.2.24).

Proof: It follows Theorem 4.7.4 that the multivariate quadratic forms $\mathbf{Q}_i(\mathbf{X})$'s are pairwise independent and distributed as $c_i W_p \left(r(\mathbf{A}_i), \mathbf{V}; \frac{1}{c_i}\mathbf{Q}_i(\mathbf{M})\right)$ for $1 \le i \le k$ if and only if $\mathbf{A}^* \mathbf{W} \mathbf{A}^* = \mathbf{B}$. The desired result follows from Theorem 3.2.3.

We now give some interesting results related to one way MANOVA problems. Let $\mathbf{X}_k = \left[\mathbf{x}_{k_1}, \mathbf{x}_{k_2}, \ldots, \mathbf{x}_{k_{n_k}}\right]$ be the kth observation matrix of order $p \times n_k$ from kth population such that $\mathbf{X}_k \sim N_{p,n_k}\left(\boldsymbol{\mu}_k \mathbf{e}'_{n_k}, \mathbf{W}_k \otimes \mathbf{V}\right)$ where $\boldsymbol{\mu}_k$, $\mathbf{W}_k$ and $\mathbf{V}$ are as defined in Theorem 4.7.3 for $k = 1, \ldots, g$ and $n = \sum_{k=1}^{g} n_k$. Let $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_g]$ be the combined observation matrix of order $p \times n$ then $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{W} \otimes \mathbf{V})$ where $\mathbf{M} = \left[\boldsymbol{\mu}_1 \mathbf{e}'_{n_1}, \ldots, \boldsymbol{\mu}_g \mathbf{e}'_{n_g}\right]$ and $\mathbf{W}$ is of order $n \times n$. In the usual one way MANOVA $\mathbf{W}_k = \mathbf{I}_{n_k}$ for all k and samples from different populations are also independent hence $\mathbf{W} = \mathbf{I}$. If samples from different populations are independent but correlated among themselves then $\mathbf{W} = \bigoplus_{k=1}^{g} \mathbf{W}_k$ where $\bigoplus$ denotes the direct sum of $\mathbf{W}_k$'s. If all the n observations are correlated then $\mathbf{W}$ is some n.n.d. matrix with $\mathbf{W}_k$ as the kth block diagonal matrix.

Our goal is to test a hypothesis that all populations have the same mean vector. The usual one way MANOVA model to compare g population mean vectors is as follows:

$$\mathbf{x}_{kj} = \boldsymbol{\mu}_k + \varepsilon_{kj}, \quad k = 1, \ldots, g \text{ and } j = 1, \ldots, n_k \tag{4.7.41}$$

where ε_{kj}'s are i.i.d. $N_p(\mathbf{0}, \mathbf{V})$ vectors of random errors. We can also write the above model in terms of $\mathbf{X}$ as given below

$$\mathbf{X} = \mathbf{M} + \mathcal{U} \tag{4.7.42}$$

$$= \left[\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_g\right] \bigoplus_{k=1}^{g} \mathbf{e}'_{n_k} + \mathcal{U}. \tag{4.7.43}$$

Let $\mathcal{U} = \left[\varepsilon_{11}, \ldots, \varepsilon_{1n_1}, \ldots, \varepsilon_{g1}, \ldots, \varepsilon_{gn_g}\right]$ then $\mathcal{U} \sim N_{p,n}(\mathbf{O}, \mathbf{I} \otimes \mathbf{V})$ and $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{I} \otimes \mathbf{V})$.

In order to test a hypothesis of equal population means, the total corrected sum of squares and cross products matrix is partitioned into the treatment sum of squares and cross products matrix $\mathbf{H}$ and residual sum of squares and cross products matrix $\mathbf{E}$. The matrices $\mathbf{H}$ and $\mathbf{E}$ are defined below

$$\mathbf{H} = \sum_{k=1}^{g} n_k \left(\bar{\mathbf{x}}_k - \bar{\mathbf{x}}\right)\left(\bar{\mathbf{x}}_k - \bar{\mathbf{x}}\right)' \tag{4.7.44}$$

and

$$\mathbf{E} = \sum_{k=1}^{g} \sum_{j=1}^{n_k} \left(\mathbf{x}_{kj} - \bar{\mathbf{x}}_k\right)\left(\mathbf{x}_{kj} - \bar{\mathbf{x}}_k\right)' \tag{4.7.45}$$

where $\bar{\mathbf{x}}_k = \frac{1}{n_k} \mathbf{X}_k \mathbf{e}_{n_k}$ and $\bar{\mathbf{x}} = \frac{1}{n} \mathbf{X}\,\mathbf{e}$. The total corrected sum of squares and cross products matrix is given by

$$\mathbf{X}\left(\mathbf{I} - \frac{1}{n} \mathbf{e}\,\mathbf{e}'\right)\mathbf{X}' = \sum_{k=1}^{g} \sum_{j=1}^{n_k} \left(\mathbf{x}_{kj} - \bar{\mathbf{x}}\right)\left(\mathbf{x}_{kj} - \bar{\mathbf{x}}\right)' = \mathbf{H} + \mathbf{E}. \tag{4.7.46}$$

We can also represent $\mathbf{H}$ and $\mathbf{E}$ in terms of the combined observation matrix $\mathbf{X}$. Let

$$\mathbf{P} = \bigoplus_{k=1}^{g} \frac{1}{n_k} \mathbf{e}_{n_k} \mathbf{e}'_{n_k} \tag{4.7.47}$$

then

$$\mathbf{H} = \mathbf{X}(\mathbf{P} - \frac{1}{n}\mathbf{e}\,\mathbf{e}')\mathbf{X}' \text{ and } \mathbf{E} = \mathbf{X}(\mathbf{I} - \mathbf{P})\mathbf{X}'. \tag{4.7.48}$$

It is well known that if the hypothesis of equal population means is true, that is, $\mathbf{M} = \boldsymbol{\mu}\mathbf{e}'$ where $\boldsymbol{\mu}$ represents the common mean vector and $\mathcal{U} \sim N_{p,n}(\mathbf{O}, \mathbf{I} \otimes \mathbf{V})$ then $\mathbf{H} \sim W_p(g-1, \mathbf{V})$ and $\mathbf{E} \sim W_p(n-g, \mathbf{V})$. In the results given below, we first assume that $\mathcal{U} \sim N_{p,n}(\mathbf{O}, \mathbf{W} \otimes \mathbf{V})$ and then characterize the class of all n.n.d. $\mathbf{W}$'s such that distributions of $\mathbf{H}$ and $\mathbf{E}$ remain invariant. Later, we derive parallel results assuming that $\mathbf{W} = \bigoplus_{k=1}^{g} \mathbf{W}_k$.

Theorem 4.7.6 *Consider the one way MANOVA model*

$$\mathbf{x}_{kj} = \boldsymbol{\mu}_k + \boldsymbol{\varepsilon}_{kj}, \quad k = 1, \ldots, g \;\; \textit{and} \;\; j = 1, \ldots, n_k. \tag{4.7.49}$$

Let $\mathcal{U} = \left[\boldsymbol{\varepsilon}_{11}, \ldots, \boldsymbol{\varepsilon}_{1n_1}, \ldots, \boldsymbol{\varepsilon}_{g1}, \ldots, \boldsymbol{\varepsilon}_{gn_g}\right]$ *be a matrix of random errors of order* $p \times n$ *where* $n = \sum_{k=1}^{g} n_k$. *Let* $\mathcal{U} \sim N_{p,n}(\mathbf{O}, \mathbf{W} \otimes \mathbf{V})$ *where* $\mathbf{W}$ *and* $\mathbf{V}$ *are n.n.d. matrices of order* $n \times n$ *and* $p \times p$*, respectively. If the hypothesis of equal population means,* $\boldsymbol{\mu}_k = \boldsymbol{\mu}$ *for* $k = 1, \ldots, g$*, is true then*

(1) $\mathbf{H} \sim c_1 W_p(g-1, \mathbf{V})$

(2) $\mathbf{E} \sim c_2 W_p(n-g, \mathbf{V})$

(3) $\mathbf{H}$ *is independent of* $\mathbf{E}$

if and only if

$$\mathbf{W} = c_2\,\mathbf{I} + (c_1 - c_2)\,\mathbf{P} + \frac{1}{n}(\mathbf{e}\,\mathbf{a}' + \mathbf{a}\,\mathbf{e}') - \frac{\bar{a} + c_1}{n}\,\mathbf{e}\,\mathbf{e}' \tag{4.7.50}$$

where $\mathbf{a}$ *is an arbitrary vector satisfying*

$$\frac{\mathbf{a}'\,(\mathbf{P} - \frac{1}{n}\mathbf{e}\,\mathbf{e}')\,\mathbf{a}}{n\,c_1} + \frac{\mathbf{a}'\,(\mathbf{I} - \mathbf{P})\,\mathbf{a}}{n\,c_2} \le \bar{a}, \tag{4.7.51}$$

$\mathbf{H}$, $\mathbf{E}$ *are defined in* (4.7.48), $\mathbf{P}$ *is defined in* (4.7.47) *and* $c_1 > 0$, $c_2 > 0$.

Proof: Let $\mathbf{X} = \left[\mathbf{x}_{11}, \ldots, \mathbf{x}_{1n_1}, \ldots, \mathbf{x}_{g1}, \ldots, \mathbf{x}_{gn_g}\right]$ and $\boldsymbol{\mu}_k = \boldsymbol{\mu}$ for $k = 1, \ldots, g$ then $\mathbf{X} \sim N_{p,n}(\boldsymbol{\mu}\,\mathbf{e}', \mathbf{W} \otimes \mathbf{V})$. It is clear from (4.7.46) and (4.7.48) that $\mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}' = (\mathbf{P} - \frac{1}{n}\mathbf{e}\,\mathbf{e}') + (\mathbf{I} - \mathbf{P})$. Letting $\mathbf{A}_1 = \mathbf{P} - \frac{1}{n}\mathbf{e}\,\mathbf{e}'$ and $\mathbf{A}_2 = \mathbf{I} - \mathbf{P}$ in Theorem 4.7.5, we have (1), (2) and (3) hold if and only if $\mathbf{W}$ is of the form (4.7.50) with $\mathbf{a}$ satisfying (4.7.51).

Remark 4.7.4 In the above theorem we have shown that if all the observations from all the populations are correlated with the covariance structure $\mathbf{W}$ given in (4.7.50) then distributions of $\mathbf{H}$ and $\mathbf{E}$ remain invariant except for a constant factor.

Corollary 4.7.1 *Under the assumptions of Theorem* 4.7.6, *we have*

(1) $\mathbf{H} \sim d\,W_p(g - 1, \mathbf{V})$

(2) $\mathbf{E} \sim d\,W_p(n - g, \mathbf{V})$

(3) $\mathbf{H}$ *is independent of* $\mathbf{E}$

if and only if $\mathbf{W} \in \mathcal{G}_{d,n}$ *for some constant* $d > 0$.

Proof: Putting $c_1 = c_2 = d$ in Theorem 4.7.6, we get from Remark 3.2.3 that (1), (2) and (3) hold if and only if $\mathbf{W} \in \mathcal{G}_{d,n}$ for some constant $d > 0$.

In some applications it is reasonable to assume that samples from different populations are independent but correlated among themselves, that is, $\mathbf{W} = \bigoplus_{k=1}^{g} \mathbf{W}_k$. With this assumption in the next theorem we characterize the class of all n.n.d. $\mathbf{W}$'s such that the distributions of $\mathbf{H}$ and $\mathbf{E}$ remain invariant except for a constant factor.

Theorem 4.7.7 *Consider the one way MANOVA model*

$$\mathbf{x}_{kj} = \boldsymbol{\mu}_k + \boldsymbol{\varepsilon}_{kj}, \quad k = 1, \ldots, g \ \text{ and } \ j = 1, \ldots, n_k. \tag{4.7.52}$$

Let $\mathcal{U}_k = \left[\boldsymbol{\varepsilon}_{k1}, \ldots, \boldsymbol{\varepsilon}_{kn_k}\right]$ *be a matrix of random errors from the kth population of order* $p \times n_k$ *and* $n = \sum_{k=1}^{g} n_k$. *Let* $\mathcal{U}_k \sim N_{p,n_k}(\mathbf{O}, \mathbf{W}_k \otimes \mathbf{V})$ *where* $\mathbf{W}_k$ *and* $\mathbf{V}$ *are n.n.d. matrices of order* $n_k \times n_k$ *and* $p \times p$, *respectively for* $k = 1, \ldots, g$. *Assume that* $\mathcal{U}_k$*'s are independent and* $\mathcal{U} = [\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_g]$ *be the combined random error matrix from g populations, that is,* $\mathcal{U} \sim N_{p,n}(\mathbf{O}, \mathbf{W} \otimes \mathbf{V})$ *where* $\mathbf{W} = \bigoplus_{k=1}^{g} \mathbf{W}_k$. *If the hypothesis of equal population means,* $\boldsymbol{\mu}_k = \boldsymbol{\mu}$ *for* $k = 1, \ldots, g$, *is true and* $g \geq 3$ *then* (1), (2) *and* (3) *of Theorem* 4.7.6 *hold if and only if*

$$\mathbf{W} = \bigoplus_{k=1}^{g} \left(c_2 \mathbf{I}_{n_k} + \frac{(c_1 - c_2)}{n_k} \mathbf{e}_{n_k} \mathbf{e}_{n_k}' \right) \tag{4.7.53}$$

where $c_1 > 0$ *and* $c_2 > 0$.

Proof: $\mathbf{X} = \left[\mathbf{x}_{11}, \ldots, \mathbf{x}_{1n_1}, \ldots, \mathbf{x}_{g1}, \ldots, \mathbf{x}_{gn_g}\right]$ and $\boldsymbol{\mu}_k = \boldsymbol{\mu}$ for $k = 1, \ldots, g$ then $\mathbf{X} \sim N_{p,n}(\boldsymbol{\mu}\,\mathbf{e}', \mathbf{W} \otimes \mathbf{V})$ where $\mathbf{W} = \bigoplus_{k=1}^{g} \mathbf{W}_k$. It follows from Theorem 4.7.6 that (1), (2) and (3) hold if and only if

$$\mathbf{W} = c_2\,\mathbf{I} + (c_1 - c_2)\,\mathbf{P} + \frac{1}{n}(\mathbf{e}\,\mathbf{a}' + \mathbf{a}\,\mathbf{e}') - \frac{\overline{a} + c_1}{n}\,\mathbf{e}\,\mathbf{e}' \tag{4.7.54}$$

for some vector $\mathbf{a}$ satisfying (4.7.51). Let w_{ij} denote the (i, j)th element of $\mathbf{W}$. Since $\mathcal{U}_k$'s are independent we also require that

$$w_{ij} = a_i + a_j - (\overline{a} + c_1) = 0 \text{ for } 1 \le i \le n_1,\ n_1 + 1 \le j \le n \tag{4.7.55}$$

and

$$w_{ij} = a_i + a_j - (\overline{a} + c_1) = 0 \text{ for } \begin{array}{l} i = n_1 + 1, \ldots, n_1 + n_2, \\ j = 1, \ldots, n_1, n_1 + n_2 + 1, \ldots, n. \end{array} \tag{4.7.56}$$

Since, $g \ge 3$ we have $n > n_1 + n_2$ and it is easy to check that (4.7.55) and (4.7.56) hold if and only if $\mathbf{a} = c_1\,\mathbf{e}$. Therefore from (4.7.54), we get

$$\mathbf{W} = c_2\,\mathbf{I} + (c_1 - c_2)\,\mathbf{P} \tag{4.7.57}$$

where $c_1 > 0$ and $c_2 > 0$. The result now follows from (4.7.47).

Remark 4.7.5 It is easy to note that if $n_k = m$ for $k = 1, \ldots, g$ and $c_1 = 1 + (m-1)\rho$, $c_2 = 1 - \rho$ where $-1/(m-1) < \rho < 1$ then $\mathbf{W}_k = (1-\rho)\,\mathbf{I}_m + \rho\,\mathbf{e}_m\,\mathbf{e}_m'$ for $k = 1, \ldots, g$. Hence for each sample, if the observations are positively equicorrelated ($\rho > 0$) or negatively equicorrelated $(-1/(m-1) < \rho < 0)$ then the distributions of $\mathbf{H}$ and $\mathbf{E}$ remain invariant except for a constant factor with different constant factors for $\mathbf{H}$ and $\mathbf{E}$. If $\rho = 0$ then $\mathbf{W} = \mathbf{I}$ and we get exact distributions in one way MANOVA table. Hence, to have exact distributions of $\mathbf{H}$ and $\mathbf{E}$ it is necessary and sufficient that $\rho = 0$.

Remark 4.7.6 For unbalanced one way MANOVA problem, we see from (4.7.53) that if observations in each sample are positively equicorrelated $(c_1 > c_2)$ or negatively equicorrelated $(c_1 < c_2)$ then the distributions of $\mathbf{H}$ and $\mathbf{E}$ remain invariant except for a constant factor with different constant factors for $\mathbf{H}$ and $\mathbf{E}$. The case, when $c_1 = c_2$ is discussed in the following corollary.

Corollary 4.7.2 *Under the assumptions of Theorem* 4.7.7, *we have*

(1) $\mathbf{H} \sim d W_p(g-1, \mathbf{V})$

(2) $\mathbf{E} \sim d W_p(n-g, \mathbf{V})$

(3) $\mathbf{H}$ *is independent of* $\mathbf{E}$

if and only if $\mathbf{W} = d\mathbf{I}$ *for some* $d > 0$.

Proof: Proof follows from Theorem 4.7.7 by letting $c_1 = c_2 = d$ where $d > 0$.

Remark 4.7.7 From the above corollary it is clear that in order to have exact distributions of $\mathbf{H}$ and $\mathbf{E}$ in one way MANOVA table it is necessary and sufficient that all the observations should be uncorrelated, that is, $\mathbf{W} = d\mathbf{I}$. It is interesting to note that if $g = 2$, we can allow some dependence as shown in Theorem 4.7.3 and still the distribution of the test statistic remains exact whereas for $g = 1$ and $g \geq 3$ all the observations must be uncorrelated in order to have exact distributions of the test statistics.

Chapter 5

Wishartness and Independence of Quadratic Forms Under A General Covariance Structure

5.1 Introduction

Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$ be an observation matrix of order $p \times n$ such that $\mathbf{X} \sim N_{p,n}(\mathrm{M}, \boldsymbol{\Sigma})$. In Chapter 4, under the assumption that $\boldsymbol{\Sigma} = \mathbf{W} \otimes \mathbf{V}$, we have characterized the class of all n.n.d. $\mathbf{W}$'s such that the distributions of certain multivariate test statistics remain invariant. In this chapter, we derive analogous results for an arbitrary covariance matrix $\boldsymbol{\Sigma}$.

Pavur (1987) considered the case when $\mathbf{M} = \mathbf{O}$ and obtained necessary and sufficient conditions for multivariate quadratic forms to be mutually independent and to have central and nonsingular Wishart distribution. We generalize this result for non-central and singular Wishart distribution, see Theorems 5.3.1 and 5.3.2. Pavur also studied the distributional properties of certain quadratic forms for special covariance structures $\boldsymbol{\Sigma}$. Some of the results of Pavur contain minor errors, that is, some classes of Σ's given by Pavur contain matrices that are not n.n.d. We give counter examples to point out those errors. In this chapter, we also give a complete and elegant characterization of covariance structures for the problems considered by Pavur. Our characterization of the class of Σ's is complete and elegant in the sense that it does not contain matrices that are not n.n.d. and the class is generated by an arbitrary matrix Q of order $np \times p$ as opposed to matrix $\mathbf{H}$ of order $np \times np$ used by Pavur in Theorem 2.

The organization of this chapter is as follows: Some results in matrix theory which are analogous to the results presented in Section 3.2 are given in Section 5.2. A version of the Cochran's theorem for an arbitrary covariance matrix Σ is given in Section 5.3. The statistical applications are presented in Section 5.4.

5.2 Results in Matrix Theory

The results presented here are generalizations of some of the results presented in Chapter 3. The following lemma is a straightforward generalization of Lemma 3.2.2.

Lemma 5.2.1 *Let $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$ be symmetric and idempotent matrices of order $n \times n$ such that $\mathbf{A}_i\mathbf{A}_j = \mathbf{O}$ for all $i \neq j$. Let $\mathbf{A} = \sum_{i=1}^{k} \mathbf{A}_i$ and $\mathbf{B} = \sum_{i=1}^{k} c_i\mathbf{A}_i$ where $c_i > 0$ for $1 \leq i \leq k$. Let $\mathbf{V}$ be a positive definite matrix of order $p \times p$ then $(\mathbf{A} \otimes \mathbf{I}_p)\Sigma(\mathbf{A} \otimes \mathbf{I}_p) = \mathbf{B} \otimes \mathbf{V}$ if and only if*

$$(\mathbf{A}_i \otimes \mathbf{I}_p)\Sigma(\mathbf{A}_j \otimes \mathbf{I}_p) = \begin{cases} c_i\mathbf{A}_i \otimes \mathbf{V} & \text{if } i = j \\ \mathbf{O} & \text{if } i \neq j. \end{cases} \tag{5.2.1}$$

Further, $\mathcal{M}(\mathbf{A} \otimes \mathbf{I}_p) = \mathcal{M}(\mathbf{B} \otimes \mathbf{V})$.

Proof: The proof of the first assertion is easy. We proceed to show $\mathcal{M}(\mathbf{A} \otimes \mathbf{I}_p) = \mathcal{M}(\mathbf{B} \otimes \mathbf{V})$. Since $\mathbf{AB} = \mathbf{B}$, we have $(\mathbf{A} \otimes \mathbf{I}_p)(\mathbf{B} \otimes \mathbf{V}) = (\mathbf{B} \otimes \mathbf{V})$. Hence $\mathcal{M}(\mathbf{B} \otimes \mathbf{V}) \subseteq \mathcal{M}(\mathbf{A} \otimes \mathbf{I}_p)$. Suppose $\mathbf{x} = (\mathbf{A} \otimes \mathbf{I}_p)\,\mathbf{v}$ and let $\mathbf{v}^* = \left(\sum_{i=1}^{k} \frac{1}{c_i} \mathbf{A}_i \otimes \mathbf{V}^{-1}\right) \mathbf{v}$. It is easy to verify that $\mathbf{x} = (\mathbf{B} \otimes \mathbf{V})\mathbf{v}^*$ and hence $\mathcal{M}(\mathbf{A} \otimes \mathbf{I}_p) \subseteq \mathcal{M}(\mathbf{B} \otimes \mathbf{V})$. This completes the proof of the lemma.

Remark 5.2.1 The above theorem is also true if $\mathbf{V}$ is an n.n.d. matrix except that in this case, we have $\mathcal{M}(\mathbf{B} \otimes \mathbf{V}) \subseteq \mathcal{M}(\mathbf{A} \otimes \mathbf{I}_p)$.

In the following theorem we find the class of all n.n.d. solutions to the matrix equation $(\mathbf{A} \otimes \mathbf{I}_p)\,\boldsymbol{\Sigma}\,(\mathbf{A} \otimes \mathbf{I}_p) = \mathbf{B} \otimes \mathbf{V}$ given in the above lemma.

Theorem 5.2.1 *Let* $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$, $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{V}$ *be as defined in Lemma* 5.2.1. *Then the class of all n.n.d. solutions for the matrix equation*

$$(\mathbf{A} \otimes \mathbf{I}_p)\,\boldsymbol{\Sigma}\,(\mathbf{A} \otimes \mathbf{I}_p) = \mathbf{B} \otimes \mathbf{V} \tag{5.2.2}$$

is given by

$$\boldsymbol{\Sigma} = (\mathbf{B} \otimes \mathbf{V}) + \mathbf{J}\,\mathbf{C} + \mathbf{C}\,\mathbf{J} - \mathbf{J}\,\mathbf{C}\,\mathbf{J} \tag{5.2.3}$$

where $\mathbf{J} = ((\mathbf{I} - \mathbf{A}) \otimes \mathbf{I}_p)$ *and* $\mathbf{C}$ *is a symmetric matrix of order* $np \times np$ *such that*

$$\mathbf{D} \stackrel{\text{def}}{=} \mathbf{J}\mathbf{C}\mathbf{J} - \mathbf{J}\mathbf{C}\left(\mathbf{B}^- \otimes \mathbf{V}^{-1}\right)\mathbf{C}\mathbf{J} \tag{5.2.4}$$

is an n.n.d. matrix. If $\mathbf{J}\,\mathbf{C}\,\mathbf{J}$ *is a null matrix then* $\boldsymbol{\Sigma}$ *given by* (5.2.3) *is an n.n.d. solution for* (5.2.2) *if and only if* $\mathbf{C}\,\mathbf{J} = \mathbf{O}$.

Proof: It is easy to see that (5.2.2) is a consistent equation since $(\mathbf{A} \otimes \mathbf{I}_p)(\mathbf{B} \otimes \mathbf{V})(\mathbf{A} \otimes \mathbf{I}_p) = \mathbf{B} \otimes \mathbf{V}$ by using the fact that $\mathbf{A}\,\mathbf{B} = \mathbf{B}$. The proof now follows easily from Theorem 3.2.2.

Remark 5.2.2 In the above theorem if $\mathbf{V}$ is an n.n.d. matrix then $\boldsymbol{\Sigma}$ given by (5.2.3) is also an n.n.d. solution for (5.2.2) where now $\mathbf{C}$ is a symmetric matrix such that $\mathbf{J}\,\mathbf{C}\,\mathbf{J} - \mathbf{J}\,\mathbf{C}\,(\mathbf{B}^- \otimes \mathbf{V}^-)\,\mathbf{C}\,\mathbf{J}$ is n.n.d. and $\mathcal{M}((\mathbf{A} \otimes \mathbf{I}_p)\,\mathbf{C}\,\mathbf{J}) \subseteq \mathcal{M}(\mathbf{B} \otimes \mathbf{V})$. Also, if $\mathbf{J}\,\mathbf{C}\,\mathbf{J}$ is a null matrix then $\boldsymbol{\Sigma}$ given by (5.2.3) is an n.n.d. solution for (5.2.2) if and only if $\mathbf{C}\,\mathbf{J} = \mathbf{O}$.

The next theorem is a special case of Theorem 5.2.1 where we assume $\mathbf{A} = \mathbf{A}^*$, the centering matrix. It plays an important role in characterizing the class of n.n.d. $\boldsymbol{\Sigma}$'s such that the distributions of certain test statistics in MANOVA problems remain invariant.

Theorem 5.2.2 *Let* $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$, $\mathbf{B}$ *and* $\mathbf{V}$ *be as defined in Lemma* 5.2.1. *Suppose that* $\sum_{i=1}^{k} \mathbf{A}_i = \mathbf{A}^*$ *where* $\mathbf{A}^*$ *is the centering matrix. Then the class of all n.n.d. solutions for the matrix equation*

$$(\mathbf{A}^* \otimes \mathbf{I}_p)\,\boldsymbol{\Sigma}\,(\mathbf{A}^* \otimes \mathbf{I}_p) = \mathbf{B} \otimes \mathbf{V} \tag{5.2.5}$$

is given by

$$\Sigma = (\mathbf{B} \otimes \mathbf{V}) + \frac{1}{n}\mathcal{E}\,\mathcal{Q}' + \frac{1}{n}\mathcal{Q}\,\mathcal{E}' - \frac{1}{n}\,\mathcal{E}\,\overline{\mathcal{Q}}\,\mathcal{E}' \tag{5.2.6}$$

where $\mathcal{Q}' = [\mathcal{Q}_1', \ldots, \mathcal{Q}_n']$ *is an arbitrary matrix of order* $p \times np$ *with* $\mathcal{Q}_i$*'s of order* $p \times p$ *for* $i = 1, \ldots, n$ *such that* $\overline{\mathcal{Q}} = \frac{1}{n}\sum_{i=1}^{n} \mathcal{Q}_i$ *is a symmetric matrix,*

$$\overline{\mathcal{Q}} - \frac{1}{n}\mathcal{Q}'\left(\sum_{i=1}^{k}\frac{1}{c_i}\mathbf{A}_i \otimes \mathbf{V}^{-1}\right)\mathcal{Q} \tag{5.2.7}$$

is n.n.d. and $\mathcal{E}' = [\mathbf{I}_p, \ldots, \mathbf{I}_p]$. *If* $\overline{\mathcal{Q}} = \mathbf{O}$ *then* Σ *given by* (5.2.6) *is an n.n.d. solution for* (5.2.5) *if and only if* $\mathcal{Q} = \mathbf{O}$.

Proof: We prove the theorem by simply applying the result given in Theorem 5.2.1. It is easy to see that equation (5.2.5) is consistent. The matrix $\mathbf{J}$ in Theorem 5.2.1 reduces to

$$\begin{aligned} \mathbf{J} &= \frac{1}{n}(\mathbf{e}\,\mathbf{e}' \otimes \mathbf{I}_p) \\ &= \frac{1}{n}\mathcal{E}\,\mathcal{E}' \end{aligned} \tag{5.2.8}$$

where $\mathcal{E}' = [\mathbf{I}_p, \ldots, \mathbf{I}_p]$. Let $\mathbf{C}$ be a symmetric matrix of order $np \times np$ and $\mathbf{C}\,\mathcal{E} = \mathcal{Q}$ then $\mathbf{J}\,\mathbf{C} = \frac{1}{n}\mathcal{E}\,\mathcal{Q}'$ and $\mathbf{J}\,\mathbf{C}\,\mathbf{J} = \frac{1}{n}\mathcal{E}\,\overline{\mathcal{Q}}\,\mathcal{E}'$ where $\overline{\mathcal{Q}} = \frac{1}{n}\mathcal{E}'\,\mathbf{C}\,\mathcal{E}$. It is not difficult to verify that if $\mathcal{Q}'$ is partitioned as $\mathcal{Q}' = [\mathcal{Q}_1', \ldots, \mathcal{Q}_n']$ with $\mathcal{Q}_i$ of order $p \times p$ for $i = 1, \ldots, n$ then $\mathcal{E}'\,\mathbf{C}\,\mathcal{E} = \sum_{i=1}^{n} \mathcal{Q}_i$. Also note that $\overline{\mathcal{Q}}$ is a symmetric matrix. Thus (5.2.3) reduces to (5.2.6). It is easy to verify $\mathbf{B}^- = \sum_{i=1}^{k}\frac{1}{c_i}\mathbf{A}_i$. From (5.2.4), we have

Σ given by (5.2.6) is an n.n.d. solution for (5.2.5) if and only if

$$\begin{aligned} \mathbf{D} &= \mathbf{J}\,\mathbf{C}\,\mathbf{J} - \mathbf{J}\,\mathbf{C}(\mathbf{B}^- \otimes \mathbf{V}^{-1})\,\mathbf{C}\,\mathbf{J} \\ &= \frac{1}{n}\mathcal{E}\,\overline{\mathcal{Q}}\,\mathcal{E}' - \frac{1}{n^2}\mathcal{E}\,\mathcal{Q}'(\mathbf{B}^- \otimes \mathbf{V}^{-1})\,\mathcal{Q}\,\mathcal{E}' \\ &= \frac{1}{n}\mathcal{E}\left(\overline{\mathcal{Q}} - \frac{1}{n}\mathcal{Q}'(\mathbf{B}^- \otimes \mathbf{V}^{-1})\,\mathcal{Q}\right)\mathcal{E}' \\ &= \frac{1}{n}\mathcal{E}\left(\overline{\mathcal{Q}} - \frac{1}{n}\mathcal{Q}'\left(\sum_{i=1}^{k}\frac{1}{c_i}\mathbf{A}_i \otimes \mathbf{V}^{-1}\right)\mathcal{Q}\right)\mathcal{E}' \qquad (5.2.9) \end{aligned}$$

is n.n.d. It is clear from Lemma 2.2.1 that $\mathbf{D}$ is n.n.d. if and only if the matrix (5.2.7) is n.n.d., since $\mathcal{E}'\,\mathcal{E} = \mathbf{I}_p$. The last assertion follows easily by noting that $\mathbf{J}\,\mathbf{C}\,\mathbf{J} = \mathbf{O}$ if and only $\overline{\mathcal{Q}} = \mathbf{O}$ and $\mathbf{C}\,\mathbf{J} = \mathbf{O}$ if and only if $\mathcal{Q} = \mathbf{O}$. In this case, we get from (5.2.6) that $\Sigma = \mathbf{B} \otimes \mathbf{V}$.

Remark 5.2.3 If $\mathbf{V}$ is an n.n.d. matrix then Σ given by (5.2.6) is also a solution to (5.2.5). In this case Σ is n.n.d. if and only if $\mathcal{Q}$ is such that $\mathcal{M}((\mathbf{A}^* \otimes \mathbf{I}_p)\,\mathcal{Q}\,\mathcal{E}') \subseteq \mathcal{M}(\mathbf{B} \otimes \mathbf{V})$ and $\overline{\mathcal{Q}} - \frac{1}{n}\mathcal{Q}'(\mathbf{B}^- \otimes \mathbf{V}^-)\,\mathcal{Q}$ is n.n.d. If $\overline{\mathcal{Q}} = \mathbf{O}$ then Σ given by (5.2.6) is n.n.d. if and only if $\mathcal{Q} = \mathbf{O}$.

Remark 5.2.4 It is interesting to note that the order of matrix given in (5.2.7) depends only on p but not on n. Therefore, Theorem 5.2.2 reduces the problem of verifying the nonnegative definiteness of Σ which is of order $np \times np$ to a simpler problem of verifying the nonnegative definiteness of the $p \times p$ matrix given in (5.2.7).

Remark 5.2.5 The class of covariance matrices given in Theorem 2 of Pavur (1987) has lot of redundancy in the sense that it also contains matrices which are not n.n.d.

To see this, we present an example. Let $p = 1$, $c_i = 1$ for all i and

$$\mathbf{H} = \begin{bmatrix} \frac{-d}{n} & \frac{-d}{n} & \cdots & \frac{-d}{n} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \tag{5.2.10}$$

where $d > 0$. Note that $r(\mathbf{H}) = 1$ and hence the matrix $\mathbf{H}$ satisfies the rank condition ($r(\mathbf{H}) \leq p$) of Pavur. Let $\mathbf{\Sigma} = 1$ in Theorem 2 of Pavur then, we get a matrix, $\left(\mathbf{I} - \frac{1+2d}{n}\mathbf{e}\,\mathbf{e}'\right)$, which is not n.n.d. Whereas, if we take $\mathcal{Q}' = \left[\frac{-d}{n}, \ldots, \frac{-d}{n}\right]$ then the condition (5.2.7) is violated and this choice of $\mathcal{Q}'$ is ruled out for generating a covariance matrix. Also, for the same $\mathbf{H}$ given above, we can show that the covariance matrix $\mathbf{V}$ given by Pavur in Theorem 4(a) is not n.n.d.

We now present a simple corollary of Theorem 5.2.2. The following corollary is used to characterize the class of $\mathbf{\Sigma}$'s such that sample covariance matrix $\mathbf{S}$ has a Wishart distribution.

Corollary 5.2.1 *Let $\mathbf{A}^*$ be the centering matrix of order $n \times n$ and $\mathbf{V}$ be a positive definite matrix of order $p \times p$. The class of all n.n.d. solutions, $C_{np}(\mathbf{V})$, for the equation*

$$(\mathbf{A}^* \otimes \mathbf{I}_p)\,\mathbf{\Sigma}\,(\mathbf{A}^* \otimes \mathbf{I}_p) = \mathbf{A}^* \otimes \mathbf{V} \tag{5.2.11}$$

is given by

$$\mathbf{\Sigma} = (\mathbf{A}^* \otimes \mathbf{V}) + \frac{1}{n}\,\mathcal{E}\,\mathcal{Q}' + \frac{1}{n}\,\mathcal{Q}\,\mathcal{E}' - \frac{1}{n}\,\mathcal{E}\,\overline{\mathcal{Q}}\,\mathcal{E}' \tag{5.2.12}$$

where $\mathcal{Q}' = [\mathcal{Q}_1', \ldots, \mathcal{Q}_n']$ *is an arbitrary matrix of order* $p \times np$ *with* $\mathcal{Q}_i$*'s of order* $p \times p$ *for* $i = 1, \ldots, n$ *such that* $\overline{\mathcal{Q}} = \frac{1}{n}\sum_{i=1}^{n} \mathcal{Q}_i$ *is a symmetric matrix,*

$$\overline{\mathcal{Q}} - \frac{1}{n}\mathcal{Q}' \left(\mathbf{A}^* \otimes \mathbf{V}^{-1}\right) \mathcal{Q} \tag{5.2.13}$$

is n.n.d. and $\mathcal{E}' = [\mathbf{I}_p, \ldots, \mathbf{I}_p]$. *If* $\overline{\mathcal{Q}} = \mathbf{O}$ *then* $\mathbf{\Sigma}$ *given by* (5.2.12) *is an n.n.d. solution for* (5.2.11) *if and only if* $\mathcal{Q} = \mathbf{O}$.

Proof: The result follows from Theorem 5.2.2 if we choose $c_i = 1$ for $i = 1, \ldots, k$.

Remark 5.2.6 Note that the class $\mathcal{C}_{np}(\mathbf{V})$ is also the class of all n.n.d. g-inverses of the matrix $\mathbf{A}^* \otimes \mathbf{V}^{-1}$. Clearly, $\mathcal{C}_{np}(\mathbf{I}_p)$ is the class of all n.n.d. g-inverses of the matrix $\mathbf{A}^* \otimes \mathbf{I}_p$.

Remark 5.2.7 Note that for $d > 0$, the class of all n.n.d. solutions for the equation

$$(\mathbf{A}^* \otimes \mathbf{I}_p)\,\mathbf{\Sigma}\,(\mathbf{A}^* \otimes \mathbf{I}_p) = d\,(\mathbf{A}^* \otimes \mathbf{V}) \tag{5.2.14}$$

is simply given by $\mathcal{C}_{d,np}(\mathbf{V}) = \{d\,\mathbf{\Sigma} : \ \mathbf{\Sigma} \in \mathcal{C}_{np}(\mathbf{V})\}$.

Remark 5.2.8 If $\mathbf{V}$ is an n.n.d. matrix then the class of all solutions, $\bar{\mathcal{C}}_{np}(\mathbf{V})$, for the equation (5.2.11) is given by (5.2.12) where $\mathcal{Q}$ is such that the matrix $\overline{\mathcal{Q}} - \frac{1}{n}\mathcal{Q}' \left(\mathbf{A}^* \otimes \mathbf{V}^-\right) \mathcal{Q}$ is n.n.d. and $\mathcal{M}((\mathbf{A}^* \otimes \mathbf{I}_p)\,\mathcal{Q}\,\mathcal{E}') \subseteq \mathcal{M}(\mathbf{A}^* \otimes \mathbf{V})$. Also, if $\overline{\mathcal{Q}} = \mathbf{O}$ then $\mathbf{\Sigma}$ given by (5.2.12) is n.n.d. if and only if $\mathcal{Q} = \mathbf{O}$. Similarly, the class $\bar{\mathcal{C}}_{d,np}(\mathbf{V}) = \left\{d\,\mathbf{\Sigma} : \ \mathbf{\Sigma} \in \bar{\mathcal{C}}_{np}(\mathbf{V})\right\}$ contains all n.n.d. solutions for the equation (5.2.14) when $\mathbf{V}$ is n.n.d.

We now present the main theorems of this chapter.

5.3 Wishartness and Independence of Quadratic Forms

In this section, we present some results related to the distributions of multivariate quadratic forms and also give a necessary and sufficient condition for the mutual independence of a pair of quadratic forms when all the observations are correlated with a general covariance matrix $\mathbf{\Sigma}$. At the end of this section, we present another version of the Cochran's theorem. In the following theorem, we generalize Theorem 1 of Pavur (1987) for the case of singular and noncentral Wishart distribution.

Theorem 5.3.1 *Let* $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$ *be an observation matrix of order* $p \times n$ *such that* $\mathbf{x}_j \sim N_p(\boldsymbol{\mu}_j, \Sigma_j)$ *for* $j = 1, \ldots, n$. *Let* $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{\Sigma})$ *where* $\mathbf{M} = [\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_n]$ *is of order* $p \times n$ *and* $\mathbf{\Sigma} = V(vec(\mathbf{X}))$ *is an n.n.d. matrix of order* $np \times np$. *Let* $\mathbf{A}$ *be n.n.d. of order* $n \times n$ *then* $\mathbf{Q}(\mathbf{X}) = \mathbf{X}\,\mathbf{A}\,\mathbf{X}' \sim W_p(r(\mathbf{A}), \mathbf{V}; \mathbf{Q}(\mathbf{M}))$ *if and only if*

$$(\mathbf{A} \otimes \mathbf{I}_p)\,\mathbf{\Sigma}\,(\mathbf{A} \otimes \mathbf{I}_p) = \mathbf{A} \otimes \mathbf{V} \tag{5.3.1}$$

where $\mathbf{V}$ *is n.n.d. of order* $p \times p$.

Proof: Let $\mathbf{A} = \mathbf{K}\,\mathbf{K}'$ and $\mathbf{V} = \mathbf{U}\,\mathbf{U}'$ where $\mathbf{K}$ and $\mathbf{U}$ are matrices of full column rank and of order $n \times m$ and $p \times s$, respectively. Let (5.3.1) be given and $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{\Sigma})$. It is clear that (5.3.1) is equivalent to

$$(\mathbf{K}' \otimes \mathbf{I}_p)\,\mathbf{\Sigma}\,(\mathbf{K} \otimes \mathbf{I}_p) = \mathbf{I}_m \otimes \mathbf{V}. \tag{5.3.2}$$

Let $\mathbf{X} = \mathbf{Y} + \mathbf{M}$ where $\mathbf{Y} \sim N_{p,n}(\mathbf{O}, \boldsymbol{\Sigma})$ then

$$\begin{aligned} \mathbf{X}\mathbf{A}\mathbf{X}' &= (\mathbf{Y}+\mathbf{M})\mathbf{K}\mathbf{K}'(\mathbf{Y}+\mathbf{M})' \\ &= (\mathbf{Y}\mathbf{K}+\mathbf{M}\mathbf{K})(\mathbf{Y}\mathbf{K}+\mathbf{M}\mathbf{K})'. \end{aligned} \tag{5.3.3}$$

It follows from the properties of *vec* operator and (5.3.2) that

$$vec(\mathbf{Y}\mathbf{K}) = (\mathbf{K}' \otimes \mathbf{I}_p)\, vec(\mathbf{Y}) \sim N_{pm}(\mathbf{0}, \mathbf{I}_m \otimes \mathbf{V}). \tag{5.3.4}$$

Hence $\mathbf{Y}\mathbf{K} + \mathbf{M}\mathbf{K} \sim N_{p,m}(\mathbf{M}\mathbf{K}, \mathbf{I}_m \otimes \mathbf{V})$. Now, from Remark 4.4.3, it is clear that $\mathbf{Q}(\mathbf{X}) \sim W_p(r(\mathbf{A}), \mathbf{V}; \mathbf{Q}(\mathbf{M}))$. To prove the converse, let $\mathbf{C} = (\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}'$ and $\mathbf{Q}(\mathbf{X}) \sim W_p(r(\mathbf{A}), \mathbf{V}; \mathbf{Q}(\mathbf{M}))$ then as shown in Theorem 4.6.1, we get

$$tr(\mathbf{C}\,\mathbf{Q}(\mathbf{X})\mathbf{C}') = vec(\mathbf{X})'(\mathbf{A} \otimes \mathbf{C}'\mathbf{C})\, vec(\mathbf{X}) \sim \chi^2(s\; r(\mathbf{A}); \delta) \tag{5.3.5}$$

where $\delta = vec(\mathbf{M})'(\mathbf{A} \otimes \mathbf{C}'\mathbf{C})\, vec(\mathbf{M})$. Since $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \boldsymbol{\Sigma})$, we have $vec(\mathbf{X}) \sim N_{pn}(vec(\mathbf{M}), \boldsymbol{\Sigma})$. From Corollary 2s.1 of Searle (1971) and Lemma 4.6.1 for $p = 1$, we get from (5.3.5)

$$(\mathbf{A} \otimes \mathbf{C}'\mathbf{C})\boldsymbol{\Sigma}(\mathbf{A} \otimes \mathbf{C}'\mathbf{C}) = \mathbf{A} \otimes \mathbf{C}'\mathbf{C}. \tag{5.3.6}$$

Let $\mathbf{P}_\mathbf{V} = \mathbf{U}(\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}'$ then from (5.3.6), we get

$$(\mathbf{A} \otimes \mathbf{P}_\mathbf{V})\boldsymbol{\Sigma}(\mathbf{A} \otimes \mathbf{P}_\mathbf{V}) = \mathbf{A} \otimes \mathbf{V}. \tag{5.3.7}$$

Since $\mathbf{Q}(\mathbf{X}) \sim W_p(r(\mathbf{A}), \mathbf{V}; \mathbf{Q}(\mathbf{M}))$, from Remark 4.4.6, we have

$$(\mathbf{I}_p - \mathbf{P}_\mathbf{V})\,\mathbf{Q}(\mathbf{X})\,(\mathbf{I}_p - \mathbf{P}_\mathbf{V}) \sim W_p(r(\mathbf{A}), \mathbf{O}; \Omega) \tag{5.3.8}$$

where $\Omega = (\mathbf{I}_p - \mathbf{P_V})\,\mathbf{Q}(\mathbf{M})\,(\mathbf{I}_p - \mathbf{P_V})$. Now it is clear from (5.3.8) that

$$P\left[(\mathbf{I}_p - \mathbf{P_V})\,\mathbf{Q}(\mathbf{X})\,(\mathbf{I}_p - \mathbf{P_V}) = \mathbf{O}\right] = 1 \tag{5.3.9}$$

and hence

$$P\left[(\mathbf{I}_p - \mathbf{P_V})\,\mathbf{X}\,\mathbf{K} = \mathbf{O}\right] = 1. \tag{5.3.10}$$

Since $vec((\mathbf{I}_p - \mathbf{P_V})\,\mathbf{X}\,\mathbf{K}) = (\mathbf{K}' \otimes (\mathbf{I}_p - \mathbf{P_V}))\,vec(\mathbf{X})$, from (5.3.10), we get

$$P\left[(\mathbf{K}' \otimes (\mathbf{I}_p - \mathbf{P_V}))\,vec(\mathbf{X}) = \mathbf{0}\right] = 1. \tag{5.3.11}$$

Hence, we have

$$(\mathbf{K}' \otimes (\mathbf{I}_p - \mathbf{P_V}))\,\boldsymbol{\Sigma}\,(\mathbf{K} \otimes (\mathbf{I}_p - \mathbf{P_V})) = \mathbf{O} \tag{5.3.12}$$

which is equivalent to

$$(\mathbf{A} \otimes (\mathbf{I}_p - \mathbf{P_V}))\,\boldsymbol{\Sigma}\,(\mathbf{A} \otimes (\mathbf{I}_p - \mathbf{P_V})) = \mathbf{O}. \tag{5.3.13}$$

It follows from the following lemma that (5.3.7) and (5.3.13) are equivalent to $(\mathbf{A} \otimes \mathbf{I}_p)\,\boldsymbol{\Sigma}\,(\mathbf{A} \otimes \mathbf{I}_p) = \mathbf{A} \otimes \mathbf{V}$. This completes the proof of the theorem.

Lemma 5.3.1 *Let* $\mathbf{P_V}$ *be as defined in the above theorem then*

$$(\mathbf{A} \otimes \mathbf{I}_p)\,\boldsymbol{\Sigma}\,(\mathbf{A} \otimes \mathbf{I}_p) = \mathbf{A} \otimes \mathbf{V} \tag{5.3.14}$$

if and only if

$$(\mathbf{A} \otimes \mathbf{P_V})\,\boldsymbol{\Sigma}\,(\mathbf{A} \otimes \mathbf{P_V}) = \mathbf{A} \otimes \mathbf{V} \tag{5.3.15}$$

and

$$(\mathbf{A} \otimes (\mathbf{I}_p - \mathbf{P_V}))\, \boldsymbol{\Sigma}\, (\mathbf{A} \otimes (\mathbf{I}_p - \mathbf{P_V})) = \mathbf{O}. \tag{5.3.16}$$

Proof: Since $\mathbf{V}\,\mathbf{P_V} = \mathbf{P_V}\,\mathbf{V} = \mathbf{V}$, it is clear that (5.3.14) implies (5.3.15) and (5.3.16). To show the converse, let (5.3.15) and (5.3.16) be given, then

$$(\mathbf{A} \otimes \mathbf{I}_p)\, \boldsymbol{\Sigma}\, (\mathbf{A} \otimes \mathbf{I}_p) = (\mathbf{A} \otimes (\mathbf{I}_p - \mathbf{P_V} + \mathbf{P_V}))\, \boldsymbol{\Sigma}\, (\mathbf{A} \otimes (\mathbf{I}_p - \mathbf{P_V} + \mathbf{P_V})). \tag{5.3.17}$$

Since $\boldsymbol{\Sigma}$ is n.n.d., (5.3.16) implies $\boldsymbol{\Sigma}\,(\mathbf{A} \otimes (\mathbf{I}_p - \mathbf{P_V})) = \mathbf{O} = (\mathbf{A} \otimes (\mathbf{I}_p - \mathbf{P_V}))\,\boldsymbol{\Sigma}$. Using properties of the Kronecker product, from (5.3.15) and (5.3.17), we get (5.3.14).

Remark 5.3.1 In the above theorem $\mathbf{Q}(\mathbf{X}) \sim d\,W_p(r(\mathbf{A}), \mathbf{V}; \Omega)$ if and only if $(\mathbf{A} \otimes \mathbf{I}_p)\, \boldsymbol{\Sigma}\, (\mathbf{A} \otimes \mathbf{I}_p) = d\,(\mathbf{A} \otimes \mathbf{V})$ where $d > 0$ and $\Omega = \frac{1}{d}\mathbf{Q}(\mathbf{M})$.

In the following theorem, we give a necessary and sufficient condition for a pair of multivariate quadratic forms to be independently distributed.

Theorem 5.3.2 *Let* $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$ *be an observation matrix of order* $p \times n$ *such that* $\mathbf{x}_j \sim N_p(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$ *for* $j = 1, \ldots, n$. *Let* $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \boldsymbol{\Sigma})$ *where* $\mathbf{M} = [\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_n]$ *is of order* $p \times n$ *and* $\boldsymbol{\Sigma} = V(vec(\mathbf{X}))$ *is an n.n.d. matrix of order* $np \times np$. *Let* $\mathbf{Q}_1(\mathbf{X}) = \mathbf{X}\,\mathbf{A}\,\mathbf{X}'$ *and* $\mathbf{Q}_2(\mathbf{X}) = \mathbf{X}\,\mathbf{B}\,\mathbf{X}'$ *then* $\mathbf{Q}_1(\mathbf{X})$ *and* $\mathbf{Q}_2(\mathbf{X})$ *are independently distributed if and only if*

$$(\mathbf{A} \otimes \mathbf{I}_p)\, \boldsymbol{\Sigma}\, (\mathbf{B} \otimes \mathbf{I}_p) = \mathbf{O} \tag{5.3.18}$$

where $\mathbf{A}$ *and* $\mathbf{B}$ *are n.n.d. matrices of order* $n \times n$.

Proof: Since $\mathbf{A}$ and $\mathbf{B}$ are n.n.d. matrices, we have $\mathbf{A} = \mathbf{K}\,\mathbf{K}'$ and $\mathbf{B} = \mathbf{L}\,\mathbf{L}'$ where $\mathbf{K}$ and $\mathbf{L}$ are matrices of full column rank and of order $n \times m$ and $n \times s$, respectively. Let (5.3.18) be given. It can be shown that (5.3.18) is equivalent to

$$(\mathbf{K}' \otimes \mathbf{I}_p)\,\mathbf{\Sigma}\,(\mathbf{L} \otimes \mathbf{I}_p) = \mathbf{O}. \tag{5.3.19}$$

Let $\mathbf{X}_1 = \mathbf{X}\,\mathbf{K}$ and $\mathbf{X}_2 = \mathbf{X}\,\mathbf{L}$. Using the properties of vec operator, we get $vec(\mathbf{X}_1) = (\mathbf{K}' \otimes \mathbf{I}_p)\,vec(\mathbf{X})$ and $vec(\mathbf{X}_2) = (\mathbf{L}' \otimes \mathbf{I}_p)\,vec(\mathbf{X})$. Using the condition (5.3.19), we have

$$\begin{aligned} cov\,(vec(\mathbf{X}_1), vec(\mathbf{X}_2)) &= (\mathbf{K}' \otimes \mathbf{I}_p)\,\mathbf{\Sigma}\,(\mathbf{L} \otimes \mathbf{I}_p) && (5.3.20) \\ &= \mathbf{O} && (5.3.21) \end{aligned}$$

since $vec(\mathbf{X}) \sim N_{pn}(vec(\mathbf{M}), \mathbf{\Sigma})$. Hence $\mathbf{X}_1$ and $\mathbf{X}_2$ are independently distributed. By using this fact, $\mathbf{X}_1\,\mathbf{X}_1'$ is independent of $\mathbf{X}_2\,\mathbf{X}_2'$. Hence $\mathbf{Q}_1(\mathbf{X})$ and $\mathbf{Q}_2(\mathbf{X})$ are independently distributed.

To show the converse, let $\mathbf{Q}_1(\mathbf{X})$ and $\mathbf{Q}_2(\mathbf{X})$ be independently distributed which implies $tr(\mathbf{Q}_1(\mathbf{X}))$ and $tr(\mathbf{Q}_2(\mathbf{X}))$ are also independently distributed. Again using the properties of vec operator, we get

$$tr(\mathbf{Q}_1(\mathbf{X})) = vec(\mathbf{X})'\,(\mathbf{A} \otimes \mathbf{I}_p)\,vec(\mathbf{X}) \tag{5.3.22}$$

and

$$tr(\mathbf{Q}_2(\mathbf{X})) = vec(\mathbf{X})'\,(\mathbf{B} \otimes \mathbf{I}_p)\,vec(\mathbf{X}). \tag{5.3.23}$$

Since $vec(\mathbf{X}) \sim N_{pn}(vec(\mathbf{M}), \mathbf{\Sigma})$; $\mathbf{A} \otimes \mathbf{I}_p$ and $\mathbf{B} \otimes \mathbf{I}_p$ are n.n.d. matrices, on using Theorem 4s of Searle (1971) and Lemma 4.6.2 for $p = 1$, we get

$$(\mathbf{A} \otimes \mathbf{I}_p)\, \mathbf{\Sigma}\, (\mathbf{B} \otimes \mathbf{I}_p) = \mathbf{O}. \tag{5.3.24}$$

This completes the proof of the theorem.

We now state a simple and useful corollary to the above theorem.

Corollary 5.3.1 *Let $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{\Sigma})$ where $\mathbf{X}$, $\mathbf{M}$ and $\mathbf{\Sigma}$ are as defined in Theorem 5.3.2. Then $\mathbf{X}\,\mathbf{A}\,\mathbf{X}'$ and $\mathbf{X}\,\mathbf{L}'$ are independently distributed if and only if $(\mathbf{A} \otimes \mathbf{I}_p)\, \mathbf{\Sigma}\, (\mathbf{L}' \otimes \mathbf{I}_p) = \mathbf{O}$ where $\mathbf{A}$ is an n.n.d. matrix of order $n \times n$ and $\mathbf{L}$ is of order $p \times n$.*

In the following theorem we present another version of the Cochran's theorem for an arbitrary covariance matrix $\mathbf{\Sigma}$.

Theorem 5.3.3 *Let $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{\Sigma})$ where $\mathbf{X}$, $\mathbf{M}$ and $\mathbf{\Sigma}$ are as defined in Theorem 5.3.2. Let $\mathbf{A}_i$ $(i = 1, \ldots, k)$ and $\mathbf{A}$ be symmetric matrices of order $n \times n$ such that $\mathbf{A} = \sum_{i=1}^{k} \mathbf{A}_i$. Let $\mathbf{V}$ be a positive definite matrix of order $p \times p$. Consider the following conditions:*

(a_1) $\mathbf{X}\,\mathbf{A}_i\,\mathbf{X}' \sim W_p(r(\mathbf{A}_i), \mathbf{V}; \Omega_i)$ *where* $\Omega_i = \mathbf{M}\,\mathbf{A}_i\,\mathbf{M}'$ *for* $i = 1, \ldots, k$

(a_2) $\mathbf{X}\,\mathbf{A}_i\,\mathbf{X}'$ *and* $\mathbf{X}\,\mathbf{A}_j\,\mathbf{X}'$ *are mutually independent for* $i \neq j = 1, \ldots, k$

(a_3) $\mathbf{X}\,\mathbf{A}\,\mathbf{X}' \sim W_p(r(\mathbf{A}), \mathbf{V}; \Omega)$ *where* $\Omega = \mathbf{M}\,\mathbf{A}\,\mathbf{M}'$

(b_1) $(\mathbf{A}_i \otimes \mathbf{I}_p)\,\boldsymbol{\Sigma}\,(\mathbf{A}_i \otimes \mathbf{I}_p) = \mathbf{A}_i \otimes \mathbf{V}$ *for* $i = 1, \ldots, k$

(b_2) $(\mathbf{A}_i \otimes \mathbf{I}_p)\,\boldsymbol{\Sigma}\,(\mathbf{A}_j \otimes \mathbf{I}_p) = \mathbf{O}$ *for* $i \neq j = 1, \ldots, k$

(b_3) $(\mathbf{A} \otimes \mathbf{I}_p)\,\boldsymbol{\Sigma}\,(\mathbf{A} \otimes \mathbf{I}_p) = \mathbf{A} \otimes \mathbf{V}$

(b_4) $\sum_{i=1}^{k} r(\mathbf{A}_i) = r(\mathbf{A})$.

Then

(1) *any two of the three conditions* (a_1), (a_2), (a_3) *or*

(2) *any two of the three conditions* (b_1), (b_2), (b_3) *or*

(3) *any two conditions* (a_i) *and* (b_j) *for* $i \neq j = 1, 2, 3$ *or*

(4) (b_3) *and* (b_4) *or*

(5) (a_3) *and* (b_4)

are necessary and sufficient for all the remaining conditions: (a_1) - (b_4).

Proof: The proof is similar to the proof of Theorem 4.6.3.

5.4 Statistical Applications

Let $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \boldsymbol{\Sigma})$ where $\mathbf{X}$, $\mathbf{M}$ and $\boldsymbol{\Sigma}$ are as defined in Theorem 5.3.2. The statistical applications in this section are analogous to the applications contained in Section 4.7. Note that, $\boldsymbol{\Sigma}$ is an arbitrary covariance matrix and not necessarily of the form $\mathbf{W} \otimes \mathbf{V}$. In Theorem 5.4.1, we characterize the class of all covariance matrices $\boldsymbol{\Sigma}$, see (5.4.3), such that the sample mean vector $\bar{\mathbf{x}}$ and the sample covariance matrix $\mathbf{S}$ are independently distributed and the distribution of $\mathbf{S}$ remains invariant. It is important to note, see Remark 5.4.3, that the distribution of one sample Hotelling's T^2 statistic defined in Theorem 4.7.2 does not remain invariant with $\mathbf{\Sigma}$ given in (5.4.3). The covariance structures for which the distribution of one sample Hotelling's T^2 statistic remains invariant except for a scale factor are of the type $\mathbf{\Sigma} = \mathbf{W} \otimes \mathbf{V}$ where $\mathbf{W}$ is given in Theorem 4.7.2. However, for all the other multivariate test statistics we do obtain invariance properties similar to the results in Section 4.7.

In the first example, we characterize the class of $\mathbf{\Sigma}$'s such that for a given n.n.d. matrix $\mathbf{A}$ the distribution of $\mathbf{X}\,\mathbf{A}\,\mathbf{X}'$ remains invariant.

Example 5.4.1 Let $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{\Sigma})$ where $\mathbf{X}$, $\mathbf{M}$ and $\mathbf{\Sigma}$ are as defined in Theorem 5.3.2. Let $\mathbf{A}$ and $\mathbf{V}$ be n.n.d. matrices of order $n \times n$ and $p \times p$, respectively. Then $\mathbf{X}\,\mathbf{A}\,\mathbf{X}' \sim W_p(r(\mathbf{A}), \mathbf{V}; \mathbf{Q}(\mathbf{M}))$ if and only if

$$\mathbf{\Sigma} = (\mathbf{A}^+ \otimes \mathbf{V}) + \mathbf{J}\,\mathbf{C} + \mathbf{C}\,\mathbf{J} - \mathbf{J}\,\mathbf{C}\,\mathbf{J} \qquad (5.4.1)$$

where $\mathbf{J} = ((\mathbf{I} - \mathbf{A}^+\,\mathbf{A}) \otimes \mathbf{I}_p)$ and $\mathbf{C}$ is a symmetric matrix satisfying the following two conditions:

(a) $\mathcal{M}((\mathbf{A} \otimes \mathbf{I}_p)\,\mathbf{C}\,\mathbf{J}) \subseteq \mathcal{M}(\mathbf{A} \otimes \mathbf{V})$

(b) $\mathbf{J}\,\mathbf{C}\,\mathbf{J} - \mathbf{J}\,\mathbf{C}\,(\mathbf{A} \otimes \mathbf{V}^-)\,\mathbf{C}\,\mathbf{J}$ is an n.n.d. matrix.

Also, if $\mathbf{J}\,\mathbf{C}\,\mathbf{J}$ is a null matrix then $\mathbf{\Sigma}$ given by (5.4.1) is n.n.d. if and only if $\mathbf{C}\,\mathbf{J} = \mathbf{O}$.

From Theorem 5.3.1, we have $\mathbf{X}\,\mathbf{A}\,\mathbf{X}' \sim W_p(r(\mathbf{A}), \mathbf{V}; \mathbf{Q}(\mathbf{M}))$ if and only if

$$(\mathbf{A} \otimes \mathbf{I}_p)\,\mathbf{\Sigma}\,(\mathbf{A} \otimes \mathbf{I}_p) = \mathbf{A} \otimes \mathbf{V} \tag{5.4.2}$$

which is equivalent to characterizing the class of all n.n.d. $\mathbf{\Sigma}$'s satisfying the equation (5.4.2) and such a class can be constructed by using Theorem 3.2.2. Also, if $\mathbf{V}$ is a positive definite matrix then condition (a) given above is trivially satisfied and in that case, we have $\mathbf{V}^{-1}$ instead of $\mathbf{V}^{-}$ in condition (b).

Similarly, using Remark 5.3.1 and Theorem 3.2.2, we can also characterize the class of all $\mathbf{\Sigma}$'s such that $\mathbf{X}\,\mathbf{A}\,\mathbf{X}' \sim d\,W_p(r(\mathbf{A}), \mathbf{V};\, \Omega)$. We now characterize the class of all $\mathbf{\Sigma}$'s such that the sample covariance matrix $\mathbf{S}$ has a Wishart distribution.

Example 5.4.2 Let $\mathbf{X} \sim N_{p,n}\,(\mathbf{M}, \mathbf{\Sigma})$ where $\mathbf{X}$, $\mathbf{M}$ and $\mathbf{\Sigma}$ are as defined in Theorem 5.3.2. Let $\mathbf{V}$ be an n.n.d. matrix of order $p \times p$. Then for $d > 0$, we have $(n-1)\,\mathbf{S} \sim d\,W_p(n-1, \mathbf{V};\, \Omega)$ if and only if $\mathbf{\Sigma} \in \bar{\mathcal{C}}_{d,np}(\mathbf{V})$ where the class $\bar{\mathcal{C}}_{d,np}(\mathbf{V})$ is defined in Remark 5.2.8, $\mathbf{S}$ is defined in (4.3.3), $\Omega = \frac{1}{d}\mathbf{M}\,\mathbf{A}^*\,\mathbf{M}'$ and $\mathbf{A}^*$ is the centering matrix.

From Remark 5.3.1, we get $(n-1)\,\mathbf{S} \sim d\,W_p(n-1, \mathbf{V};\, \Omega)$ if and only if $(\mathbf{A}^* \otimes \mathbf{I}_p)\,\mathbf{\Sigma}\,(\mathbf{A}^* \otimes \mathbf{I}_p) = d\,(\mathbf{A}^* \otimes \mathbf{V})$ which is true if and only if $\mathbf{\Sigma} \in \bar{\mathcal{C}}_{d,np}(\mathbf{V})$. Similarly, if $\mathbf{V}$ is a positive definite matrix then $(n-1)\,\mathbf{S} \sim d\,W_p(n-1, \mathbf{V};\, \Omega)$ if and only if $\mathbf{\Sigma} \in \mathcal{C}_{d,np}(\mathbf{V})$ where $\mathcal{C}_{d,np}(\mathbf{V})$ is defined in Remark 5.2.7.

We now characterize the class of all covariance matrices $\mathbf{\Sigma}$ such that the sample mean vector and covariance matrix are independently distributed.

Theorem 5.4.1 *Let* $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \boldsymbol{\Sigma})$ *where* $\mathbf{X}$, $\mathbf{M}$ *and* $\boldsymbol{\Sigma}$ *are as defined in Theorem* 5.3.2. *Let* $\mathbf{V}$ *be an n.n.d. matrix of order* $p \times p$. *Then* $(n-1)\,\mathbf{S} \sim d\,W_p(n-1, \mathbf{V};\, \Omega)$ *where* Ω *is as given in Example* 5.4.2 *and* $\bar{\mathbf{x}}$ *is independent of* $\mathbf{S}$ *if and only if*

$$\boldsymbol{\Sigma} = d\left((\mathbf{I} \otimes \mathbf{V}) - (\mathbf{e}\,\mathbf{e}' \otimes \frac{(\mathbf{V} - \Psi)}{n})\right) \tag{5.4.3}$$

where Ψ *is an n.n.d. matrix of order* $p \times p$ *and* $d > 0$.

Proof: From Example 5.4.2 for $d > 0$, we have $(n-1)\,\mathbf{S} \sim d\,W_p(n-1, \mathbf{V};\, \Omega)$ if and only if $\boldsymbol{\Sigma} \in \bar{\mathcal{C}}_{d,np}(\mathbf{V})$. Also, from Corollary 5.3.1, $\bar{\mathbf{x}}$ is independent of $\mathbf{S}$ if and only if

$$(\mathbf{A}^* \otimes \mathbf{I}_p)\,\boldsymbol{\Sigma}\,\left(\frac{\mathbf{e}}{n} \otimes \mathbf{I}_p\right) = \mathbf{O}. \tag{5.4.4}$$

It is easy to verify that $\boldsymbol{\Sigma} \in \bar{\mathcal{C}}_{d,np}(\mathbf{V})$ and satisfies (5.4.4) if and only if

$$(\mathbf{A}^* \otimes \mathbf{I}_p)\,\mathcal{Q} = \mathbf{O} \tag{5.4.5}$$

which holds if and only if $\mathcal{Q} = \mathcal{E}\,\Psi$ where $\mathcal{Q}$ and $\mathcal{E}$ are as defined in Corollary 5.2.1 and Ψ is a symmetric matrix of order $p \times p$. For $\boldsymbol{\Sigma} \in \bar{\mathcal{C}}_{d,np}(\mathbf{V})$ to be n.n.d., $\mathcal{Q} = \mathcal{E}\,\Psi$ must satisfy the following conditions:

(a) $\mathcal{M}(\mathbf{A}^* \otimes \mathbf{I}_p)\mathcal{Q}\,\mathcal{E}') \subseteq \mathcal{M}(\mathbf{A}^* \otimes \mathbf{V})$

(b) $\overline{\mathcal{Q}} - \frac{1}{n}\mathcal{Q}'\left(\mathbf{A}^* \otimes \mathbf{V}^-\right)\,\mathcal{Q}$ is an n.n.d. matrix.

It is clear that $\mathcal{Q} = \mathcal{E}\,\Psi$ trivially satisfies condition (a) since $\mathcal{Q} = \mathcal{E}\,\Psi = (\mathbf{e} \otimes \Psi)$. Also (b) holds for $\mathcal{Q} = \mathcal{E}\,\Psi$ if and only if Ψ is an n.n.d. matrix. Thus $(n-1)\,\mathbf{S} \sim$

$dW_p(n-1, \mathbf{V}; \Omega)$ and $\bar{\mathbf{x}}$ is independent of $\mathbf{S}$ if and only if

$$\Sigma = d\left((\mathbf{A}^* \otimes \mathbf{V}) + \frac{1}{n}\mathcal{E}\Psi\mathcal{E}'\right) \tag{5.4.6}$$

$$= d\left(((\mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}') \otimes \mathbf{V}) + \frac{1}{n}\mathcal{E}\Psi\mathcal{E}'\right) \tag{5.4.7}$$

$$= d\left((\mathbf{I} \otimes \mathbf{V}) - (\mathbf{e}\,\mathbf{e}' \otimes \frac{(\mathbf{V} - \Psi)}{n})\right) \tag{5.4.8}$$

since $\mathcal{E}\Psi\mathcal{E}' = (\mathbf{e}\,\mathbf{e}') \otimes \Psi$. This completes the proof of the theorem.

Remark 5.4.1 In the above theorem, if $\Psi = c\mathbf{V}$ where $c \geq 0$ then $\Sigma = d\left(\mathbf{I} - \frac{(1-c)}{n}\mathbf{e}\,\mathbf{e}'\right) \otimes \mathbf{V}$ is the class of all covariance matrices we obtained in Theorem 4.7.1.

Remark 5.4.2 At this stage we point out another error in the covariance structure given by Pavur (1987) in Theorem 4(b). Let $\mathbf{W} = -\Sigma$ in Theorem 4(b) of Pavur. For this choice of $\mathbf{W}$, we get a matrix $\mathbf{V}$ which is not an n.n.d. matrix.

Remark 5.4.3 If $\mathbf{X} \sim N_{p,n}(\boldsymbol{\mu}\,\mathbf{e}', \Sigma)$ where Σ is as given in (5.4.3) then it is easy to verify that $\bar{\mathbf{x}} \sim N_p(\boldsymbol{\mu}, \frac{d}{n}\Psi)$. It is interesting to note that the covariance matrix of $\bar{\mathbf{x}}$ does not depend on $\mathbf{V}$. Due to this reason, the distribution of one sample Hotelling's T^2 statistic, defined in Theorem 4.7.2, does not remain invariant with Σ given in (5.4.3). Hence, the only covariance structure for which the distribution of one sample Hotelling's T^2 statistic remains invariant except for a scale factor are of the type $\Sigma = \mathbf{W} \otimes \mathbf{V}$ where $\mathbf{W}$ is given in Theorem 4.7.2. Now, we examine the invariance property of the two sample Hotelling's T^2 statistic defined in Theorem 4.7.3.

Theorem 5.4.2 *Let* $\mathbf{X}_k = \left[\mathbf{x}_{k_1}, \mathbf{x}_{k_2}, \ldots, \mathbf{x}_{k_{n_k}}\right]$ *be an observation matrix of order* $p \times n_k$ *such that* $\mathbf{X}_k \sim N_{p,n_k}\left(\boldsymbol{\mu}_k \mathbf{e}'_{n_k}, \boldsymbol{\Sigma}_k\right)$ *for* $k = 1, 2$. *Let* $\boldsymbol{\Sigma}_1$ *and* $\boldsymbol{\Sigma}_2$ *be n.n.d. matrices,* $\mathbf{V}$ *be a positive definite matrix of order* $p \times p$, $\mathbf{X}_1$ *and* $\mathbf{X}_2$ *be mutually independent and* $n_1 + n_2 - 2 \geq p$. *Let* $\bar{\mathbf{x}}_1$, $\bar{\mathbf{x}}_2$ *be the mean vectors and* $\mathbf{S}_1$, $\mathbf{S}_2$ *be the sample covariance matrices of* $\mathbf{X}_1$ *and* $\mathbf{X}_2$, *respectively. Also, let* $\mathbf{S}_p = [(n_1 - 1)\,\mathbf{S}_1 + (n_2 - 1)\,\mathbf{S}_2]/(n_1 + n_2 - 2)$ *and* $\boldsymbol{\theta} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$. *Then*

$$\frac{n_1 n_2}{n_1 + n_2} \frac{\left((\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) - \boldsymbol{\theta}\right)' \mathbf{S}_p^{-1} \left((\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) - \boldsymbol{\theta}\right)}{n_1 + n_2 - 2} \frac{(n_1 + n_2 - p - 1)}{p} \sim F(p, n_1 + n_2 - p - 1) \tag{5.4.9}$$

if

$$\boldsymbol{\Sigma}_1 = d\left((\mathbf{I}_{n_1} \otimes \mathbf{V}) + (\mathbf{e}_{n_1} \mathbf{e}'_{n_1} \otimes \boldsymbol{\Phi})\right) \tag{5.4.10}$$

and

$$\boldsymbol{\Sigma}_2 = d\left((\mathbf{I}_{n_2} \otimes \mathbf{V}) - (\mathbf{e}_{n_2} \mathbf{e}'_{n_2} \otimes \boldsymbol{\Phi})\right) \tag{5.4.11}$$

where $\boldsymbol{\Phi}$ *is a symmetric matrix of order* $p \times p$ *such that*

$$-1/n_1 < \lambda_p(\mathbf{V}^{-1}\boldsymbol{\Phi}) \leq \lambda_1(\mathbf{V}^{-1}\boldsymbol{\Phi}) < 1/n_2$$

with $\lambda_1(\mathbf{V}^{-1}\boldsymbol{\Phi})$ *and* $\lambda_p(\mathbf{V}^{-1}\boldsymbol{\Phi})$ *denoting the maximum and the minimum eigenvalues of* $\mathbf{V}^{-1}\boldsymbol{\Phi}$, *respectively.*

Proof: If $d > 0$, then it follows from Theorem 5.4.1 that $(n_k - 1)\,\mathbf{S}_k \sim d\,W_p(n_k - 1, \mathbf{V})$ and $\bar{\mathbf{x}}_k$ is independent of $\mathbf{S}_k$ if and only if

$$\boldsymbol{\Sigma}_k = d\left((\mathbf{I}_{n_k} \otimes \mathbf{V}) - (\mathbf{e}_{n_k}\,\mathbf{e}'_{n_k} \otimes \frac{(\mathbf{V} - \Psi_k)}{n_k})\right) \tag{5.4.12}$$

where Ψ_k is an n.n.d. matrix for $k = 1, 2$. From Theorem 4.4.3, we have

$$\frac{n_1 + n_2 - 2}{d}\,\mathbf{S}_p \sim W_p(n_1 + n_2 - 2, \mathbf{V}). \tag{5.4.13}$$

It can be verified that

$$\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2 - \boldsymbol{\theta} \sim N_p\left(\mathbf{0}, d\,(\frac{\Psi_1}{n_1} + \frac{\Psi_2}{n_2})\right) \tag{5.4.14}$$

where $\boldsymbol{\theta} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$. Thus, with Ψ_1 and Ψ_2 positive definite, for (5.4.9) to hold, we must have

$$\frac{\Psi_1}{n_1} + \frac{\Psi_2}{n_2} = (\frac{1}{n_1} + \frac{1}{n_2})\,\mathbf{V} \tag{5.4.15}$$

or equivalently,

$$\frac{-(\mathbf{V} - \Psi_1)}{n_1} = \frac{(\mathbf{V} - \Psi_2)}{n_2} = \boldsymbol{\Phi} \quad \text{(say)} \tag{5.4.16}$$

where $\boldsymbol{\Phi}$ is a symmetric matrix of order $p \times p$. It follows from the following lemma that $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are positive definite matrices if and only if $-1/n_1 < \lambda_p(\mathbf{V}^{-1}\,\boldsymbol{\Phi}) \leq \lambda_1(\mathbf{V}^{-1}\,\boldsymbol{\Phi}) < 1/n_2$ with $\lambda_1(\mathbf{V}^{-1}\,\boldsymbol{\Phi})$ and $\lambda_p(\mathbf{V}^{-1}\,\boldsymbol{\Phi})$ denoting the maximum and the minimum eigenvalues of $\mathbf{V}^{-1}\,\boldsymbol{\Phi}$, respectively. From (5.4.16) and (5.4.12), we get (5.4.10) and (5.4.11). And in this case (5.4.9) also holds.

Lemma 5.4.1 *Let* (5.4.16) *be given, then* Ψ_1 *and* Ψ_2 *are positive definite matrices if and only if*

$$-1/n_1 < \lambda_p(\mathbf{V}^{-1}\,\mathbf{\Phi}) \leq \lambda_1(\mathbf{V}^{-1}\,\mathbf{\Phi}) < 1/n_2 \tag{5.4.17}$$

with $\lambda_1(\mathbf{V}^{-1}\,\mathbf{\Phi})$ *and* $\lambda_p(\mathbf{V}^{-1}\,\mathbf{\Phi})$ *denoting the maximum and the minimum eigenvalues of* $\mathbf{V}^{-1}\,\mathbf{\Phi}$*, respectively.*

Proof: From (5.4.16), we have $\Psi_1 = \mathbf{V} + n_1\,\mathbf{\Phi}$ and $\Psi_2 = \mathbf{V} - n_2\,\mathbf{\Phi}$. From Theorem 2.2.1, Ψ_2 is positive definite matrix if and only if $\lambda_1(\mathbf{V}^{-1}\,\mathbf{\Phi}) < 1/n_2$ and Ψ_1 is positive definite matrix if and only if $-1/n_1 < \lambda_p(\mathbf{V}^{-1}\,\mathbf{\Phi})$. Hence Ψ_1 and Ψ_2 are positive definite matrices if and only if (5.4.17) holds.

Remark 5.4.4 In the above theorem, let $\mathbf{\Phi} = \beta\,\mathbf{V}$ where β is a constant such that $-1/n_1 < \beta < 1/n_2$. Then the class of all covariance matrices we obtain in this case is the same as that in Theorem 4.7.3.

Remark 5.4.5 Basu et al. (1974) called the covariance structures (5.4.10) and (5.4.11) as equicorrelated covariance structures and pointed out that under the equicorrelated covariance structure the distributions of commonly used statistics are complicated and do not correspond to the well known distributions for which tables are given. We concur with their statement regarding the distribution of one sample Hotelling's T^2 statistic. However, we have shown that, under the equicorrelated covariance structure the distribution of the two sample Hotelling's T^2 statistic remains invariant.

We now examine the invariance properties of certain test statistics occurring in MANOVA problems.

Theorem 5.4.3 *Let* $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$ *be symmetric and idempotent matrices of order* $n \times n$ *such that* $\mathbf{A}_i \mathbf{A}_j = \mathbf{O}$ *for all* $i \neq j$. *Let* $\mathbf{A} = \sum_{i=1}^{k} \mathbf{A}_i$ *and* $\mathbf{B} = \sum_{i=1}^{k} c_i \mathbf{A}_i$ *where* $c_i > 0$ *for* $1 \leq i \leq k$. *Let* $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{\Sigma})$ *where* $\mathbf{X}$, $\mathbf{M}$ *and* $\mathbf{\Sigma}$ *are as defined in Theorem* 5.3.2. *Let* $\mathbf{V}$ *be a positive definite matrix of order* $p \times p$. *Let* $\mathbf{Q}_i(\mathbf{X}) = \mathbf{X}\mathbf{A}_i\mathbf{X}'$ *for* $1 \leq i \leq k$. *Then the multivariate quadratic forms* $\mathbf{Q}_i(\mathbf{X})$*'s are pairwise independent and distributed as* $c_i W_p\left(r(\mathbf{A}_i), \mathbf{V}; \frac{1}{c_i}\mathbf{Q}_i(\mathbf{M})\right)$ *for* $1 \leq i \leq k$ *if and only if* $(\mathbf{A} \otimes \mathbf{I}_p)\mathbf{\Sigma}(\mathbf{A} \otimes \mathbf{I}_p) = \mathbf{B} \otimes \mathbf{V}$.

Proof: From Remark 5.3.1 and Theorem 5.3.2, we have $\mathbf{Q}_i(\mathbf{X})$'s are pairwise independent and distributed as $c_i W_p\left(r(\mathbf{A}_i), \mathbf{V}; \frac{1}{c_i}\mathbf{Q}_i(\mathbf{M})\right)$ for $1 \leq i \leq k$ if and only if

$$(\mathbf{A}_i \otimes \mathbf{I}_p)\mathbf{\Sigma}(\mathbf{A}_j \otimes \mathbf{I}_p) = \begin{cases} c_i \mathbf{A}_i \otimes \mathbf{V} & \text{if } i = j \\ \mathbf{O} & \text{if } i \neq j. \end{cases} \tag{5.4.18}$$

The theorem now follows from Lemma 5.2.1.

Remark 5.4.6 In the above theorem, for a given matrix $\mathbf{A}$, the class of all n.n.d. $\mathbf{\Sigma}$'s satisfying (5.4.18) is given by Theorem 5.2.1.

As we have mentioned in Chapter 4 that in all MANOVA problems the total corrected sum of squares and cross products matrix is decomposed into two or more orthogonal sum of squares and cross products matrices. In this case, the matrix $\mathbf{A}$

defined in the above theorem equals the centering matrix $\mathbf{A}^*$. Hence we obtain the following special case of the above theorem.

Theorem 5.4.4 *Let $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$ be symmetric and idempotent matrices of order $n \times n$ such that $\mathbf{A}_i\mathbf{A}_j = \mathbf{O}$ for all $i \neq j$. Let $\sum_{i=1}^{k} \mathbf{A}_i = \mathbf{A}^*$ and $\mathbf{B} = \sum_{i=1}^{k} c_i \mathbf{A}_i$ where $\mathbf{A}^*$ is the centering matrix and $c_i > 0$ for $1 \leq i \leq k$. Let $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{\Sigma})$ where $\mathbf{X}$, $\mathbf{M}$ and $\mathbf{\Sigma}$ are as defined in Theorem 5.3.2. Let $\mathbf{V}$ be a positive definite matrix of order $p \times p$. Let $\mathbf{Q}_i(\mathbf{X}) = \mathbf{X}\mathbf{A}_i\mathbf{X}'$ for $1 \leq i \leq k$. Then the multivariate quadratic forms $\mathbf{Q}_i(\mathbf{X})$'s are pairwise independent and distributed as $c_i W_p\left(r(\mathbf{A}_i), \mathbf{V}; \frac{1}{c_i}\mathbf{Q}_i(\mathbf{M})\right)$ for $1 \leq i \leq k$ if and only if $\mathbf{\Sigma}$ is of the form (5.2.6) where Q is an arbitrary matrix of order $np \times p$ satisfying (5.2.7).*

Proof: It follows from the Theorem 5.4.3 that the multivariate quadratic forms $\mathbf{Q}_i(\mathbf{X})$'s are pairwise independent and distributed as $c_i W_p\left(r(\mathbf{A}_i), \mathbf{V}; \frac{1}{c_i}\mathbf{Q}_i(\mathbf{M})\right)$ for $1 \leq i \leq k$ if and only if $(\mathbf{A}^* \otimes \mathbf{I}_p)\mathbf{\Sigma}(\mathbf{A}^* \otimes \mathbf{I}_p) = \mathbf{B} \otimes \mathbf{V}$. The desired result now follows from Theorem 5.2.2.

In the following theorem we characterize the class of all covariance matrices $\mathbf{\Sigma}$ such that the distributions of the error sum of squares and cross products matrix $\mathbf{E}$ and the treatment sum of squares and cross products matrix $\mathbf{H}$ remain invariant except for a constant factor. For the following theorem we use the same notations as introduced after Theorem 4.7.5.

Theorem 5.4.5 *Consider the one way MANOVA model*

$$\mathbf{x}_{kj} = \boldsymbol{\mu}_k + \boldsymbol{\varepsilon}_{kj}, \quad k = 1, \ldots, g \ \ and \ \ j = 1, \ldots, n_k. \tag{5.4.19}$$

Let $\mathcal{U} = \left[\boldsymbol{\varepsilon}_{11}, \ldots, \boldsymbol{\varepsilon}_{1n_1}, \ldots, \boldsymbol{\varepsilon}_{g1}, \ldots, \boldsymbol{\varepsilon}_{gn_g}\right]$ *be a matrix of random errors of order* $p \times n$ *where* $n = \sum_{k=1}^{g} n_k$. *Let* $\mathcal{U} \sim N_{p,n}(\mathbf{O}, \boldsymbol{\Sigma})$ *where* $\boldsymbol{\Sigma}$ *is an n.n.d. matrix. Let* $\mathbf{V}$ *be a positive definite matrix of order* $p \times p$. *If the hypothesis of equal population means,* $\boldsymbol{\mu}_k = \boldsymbol{\mu}$ *for* $k = 1, \ldots, g$, *is true then*

(1) $\mathbf{H} \sim d\,W_p(g-1, \mathbf{V})$

(2) $\mathbf{E} \sim d\,W_p(n-g, \mathbf{V})$

(3) $\mathbf{H}$ *is independent of* $\mathbf{E}$

if and only if $\boldsymbol{\Sigma} \in \mathcal{C}_{d,np}(\mathbf{V})$.

Proof: Let $\mathbf{X} = \left[\mathbf{x}_{11}, \ldots, \mathbf{x}_{1n_1}, \ldots, \mathbf{x}_{g1}, \ldots, \mathbf{x}_{gn_g}\right]$ and $\boldsymbol{\mu}_k = \boldsymbol{\mu}$ for $k = 1, \ldots, g$ then $\mathbf{X} \sim N_{p,n}(\boldsymbol{\mu}\,\mathbf{e}', \boldsymbol{\Sigma})$. It is clear from (4.7.46) and (4.7.48) that $\mathbf{I} - \frac{1}{n}\mathbf{e}\,\mathbf{e}' = (\mathbf{P} - \frac{1}{n}\mathbf{e}\,\mathbf{e}') + (\mathbf{I} - \mathbf{P})$. From Theorem 5.4.4 with $\mathbf{A}_1 = \mathbf{P} - \frac{1}{n}\mathbf{e}\,\mathbf{e}'$ and $\mathbf{A}_2 = \mathbf{I} - \mathbf{P}$, we have (1), (2) and (3) hold if and only if $\boldsymbol{\Sigma} \in \mathcal{C}_{d,np}(\mathbf{V})$.

Usually, samples from different populations are independent but correlated among themselves, that is, $\boldsymbol{\Sigma} = \bigoplus_{k=1}^{g} \boldsymbol{\Sigma}_k$ where $\boldsymbol{\Sigma}_k$ denotes the covariance matrix of the kth population. In this case, if $g \geq 3$ then (1), (2) and (3) of Theorem 5.4.5 hold if and only if $\boldsymbol{\Sigma} = \mathbf{I}_{np}$. We skip the proof since it is similar to the proof of Theorem 4.7.7 and Corollary 4.7.2.

5.5 Summary and Conclusions

The most widely used statistical methods are concerned with drawing inference for the parameters of the normal populations. In these problems the distributions of the test statistics are derived under the assumption that the observations are independent and identically distributed. While the independence assumption may be approximately valid, due to the choice of the experimental designs, exploring the problem of dependence between the observations is of practical as well as aesthetic interest. In this dissertation, we have characterized the class of covariance matrices such that the distributions of the common test statistics remain invariant, that is, the distributions remain the same except for a scale factor. We have shown that in most cases the covariance structure need necessarily be equicorrelated for the distributions of the test statistics to remain invariant. We have achieved this by first obtaining an elegant characterization of the class of all nonnegative definite solutions to a matrix equation that occurs in statistics.

Bibliography

[1] Albers, W. (1978). Testing the mean of a normal population under dependence. *Ann. Stat.*, **6**, 1337-1344.

[2] Ali, M. (1973). Distribution of Student's ratio for correlated observations. *Stat. Hefte*, **14**, 350-356.

[3] Anderson, T. W. (1984). *An Introduction to Multivariate Analysis*, 2nd ed. John Wiley and Sons, New York.

[4] Baldessari, B. (1965). Remarque sur le rapport de combinaisons lineaires de χ^2. *Publications de l'Institut de Statistique de l'Université Paris*, **14**, 393-406.

[5] Banerjee, K. S. (1964). A note on idempotent matrices. *Ann. Math. Stat.*, **35**, 880-882.

[6] Basu, J. P., Odell, P. L. and Lewis, T. O. (1974). The effect of intraclass correlation on certain significance tests when sampling from multivariate normal population. *Comm. Stat.*, **3**, 899-908.

[7] Basu, S. and DasGupta, A. (1991). Robustness of standard confidence intervals for location parameters under departure from normality. *Technical Report* #91-47, Dept. of Stat., Purdue University.

[8] Bellman, R. (1970). *Introduction to Matrix Analysis*, 2nd ed. McGraw-Hill, New York.

[9] Bhat, B. R. (1962). On the distribution of certain quadratic forms in normal variables. *J. Roy. Stat. Soc., B*, **24**, 148-151.

[10] Bhimasankaram, P. and Majumdar, D. (1980). Hermitian and nonnegative definite solutions of some matrix equations connected with distribution of quadratic forms. *Sankhyā, Ser. A*, **42**, 272-282.

[11] Cochran, W. G. (1934). The distribution of quadratic forms in a normal system with application to the analysis of variance. *Proc. Cambridge Phil. Soc.*, **30**, 178-191.

[12] Craig, A. T. (1943). Note on the independence of certain quadratic forms. *Ann. Math. Stat.*, **14**, 195-197.

[13] Eaton, M. L. (1983). *Multivariate Statistics.* John Wiley and Sons, New York.

[14] Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, **10**, 507-521.

[15] Gastwirth, J. L. and Rubin, H. (1971). Effect of dependence on the level of some one-sample tests. *J. Amer. Stat. Assoc.,* **66**, 816-820.

[16] Good, I. J. (1966). On the independence of quadratic expressions: corrigenda. *J. Roy. Stat. Soc., B*, **28**, 584.

[17] Graybill, F. A. (1969). *Introduction to Matrices with Applications in Statistics.* Wadsworth, Belmont, CA.

[18] Graybill, F. A. and Marsaglia, G. (1957). Idempotent matrices and quadratic forms in general linear hypothesis. *Ann. Math. Stat.*, **28**, 678-686.

[19] Henderson, H. V., Pukelsheim, F. and Searle, S. R. (1983). On the history of the Kronecker product. *Lin. and Multilinear Algebra*, **14**, 113-120.

[20] Henderson, H. V. and Searle, S. R. (1979). Vec and Vech operators for matrices with some uses in Jacobians and multivariate statistics. *Can. J. Stat.*, **7**, 65-81.

[21] Henderson, H. V. and Searle, S. R. (1981). The vec-permutation matrix, the vec operator and Kronecker products: a review. *Lin. and Multilinear Algebra*, **9**, 271-288.

[22] Hsu, P. L. (1939a). On the distribution of the roots of certain determinantal equations. Ann. Eugenics, **9**, 250-258.

[23] Huber, P. J. (1981). *Robust Statistics.* John Wiley and Sons, New York.

[24] James, A. T. (1954). Normal multivariate analysis and orthogonal groups. *Ann. Math. Stat.*, **25**, 40-75.

[25] Jensen, D. R. (1989a). Misspecified location and scale-I. Student's test. *J. Stat. Computation and Simulation*, **32**, 115-118.

[26] Jensen, D. R. (1989b). Misspecified location and scale-II. Independence of $\overline{y}$ and s^2. *J. Stat. Computation and Simulation*, **32**, 119-121.

[27] Jensen, D. R. (1989c). Misspecified location and scale-III. The use of s^2. *J. Stat. Computation and Simulation*, **32**, 121-124.

[28] Khatri, C. G. (1959). On conditions for the forms of the type: $\mathbf{X}\mathbf{A}\mathbf{X}'$ to be distributed independently or to obey Wishart distribution. *Calcutta Stat. Assoc. Bull.*, **8**, 162-168.

[29] Khatri, C. G. (1962). Conditions for Wishartness and independence of second degree polynomials in normal vector. *Ann. Math. Stat.*, **33**, 1002-1007.

[30] Khatri, C. G. (1963). Further contributions to Wishartness and independence of second degree polynomials in normal vectors. *J. Indian Stat. Assoc.*, **1**, 61-70.

[31] Khatri, C. G. (1968). Some results for the singular normal multivariate regression models. *Sankhyā, Ser. A*, **30**, 267-280.

[32] Khatri, C. G. and Mitra, S. K. (1976). Hermitian and nonnegative definite solutions of linear matrix equations. *SIAM J. Appl. Math.*, **31**, 579-585.

[33] Lancaster, H. O. (1954). Traces and cumulants of quadratic forms in normal variables. *J. Roy. Stat. Soc., B*, **16**, 247-254.

[34] Lehmann, E. L. (1983). *Theory of Point Estimation.* John Wiley and Sons, New York.

[35] Loynes, R. M. (1966). On idempotent matrices. *Ann. Math. Stat.*, **37**, 295-296.

[36] Madow, W. G. (1938). Contributions to the theory of multivariate statistical analysis. *Transactions Amer. Math. Soc.*, **44**, 454-495.

[37] Magnus, J. R. and Neudecker, H. (1979). The commutation matrix: some properties and applications. *Ann. Stat.*, **7**, 381-394.

[38] Marshall, A. W. and Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications.* Academic Press, New York.

[39] Mirsky, L. (1955). *An Introduction to Linear Algebra.* Oxford Univ. Press, London, UK.

[40] Nair, K. R. (1956). A note on the estimation of mean change. *J. Indian Soc. of Agri. Stat.*, **8**, 122-124.

[41] Neudecker, H. (1969). Some theorems on matrix differentiation with special reference to Kronecker matrix products. *J. Amer. Stat. Assoc.*, **64**, 953-963.

[42] Olkin, I. (1992). A matrix formulation on how deviant an observation can be. *The Amer. Stat.*, **46(3)**, 205-209.

[43] Olkin, I. and Roy, S. N. (1954). On multivariate distribution theory. *Ann. Math. Stat.*, **25**, 329-339.

[44] Pavur, R. J. (1987). Distribution of multivariate quadratic forms under certain covariance structures. *Can. J. Stat.*, **15(2)**, 169-176.

[45] Pavur, R. J. and Lewis, T. O. (1983). Unbiased F test for factorial experiments for correlated data. *Comm. Stat.*, **12**, 829-840.

[46] Rao, C. R. (1973). *Linear Statistical Inference and Its Application*, 2nd ed. John Wiley and Sons, New York.

[47] Rao, C. R. and Mitra, S. K. (1971). *Generalized Inverse of Matrices and Its Applications.* John Wiley and Sons, New York.

[48] Rayner, A. A. and Livingstone, D. (1965). On the distribution of quadratic forms in singular normal variables. *S. African J. Agri. Sci.*, **8**, 357-370.

[49] Roy, S. N. and Gnanadesikan, R. (1959). Some contributions to ANOVA in one or more dimensions: II. *Ann. Math. Stat.*, **30**, 318-340.

[50] Samuelson, P. A. (1968). How deviant can you be? *J. Amer. Stat. Assoc.*, **63**, 1522-1525.

[51] Scariano, S. M. and Davenport, J. M. (1987). The effect of violations of independence assumptions in the one-way ANOVA. *The Amer. Stat.*, **41(2)**, 123-129.

[52] Searle, S. R. (1971). *Linear Models.* John Wiley and Sons, New York.

[53] Searle, S. R. (1982). *Matrix Algebra Useful For Statistics.* John Wiley and Sons, New York.

[54] Seber, G. A. F. (1977). *Linear Regression Analysis.* John Wiley and Sons, New York.

[55] Shanbhag, D. N. (1966). On the independence of quadratic forms. *J. Roy. Stat. Soc., B*, **28**, 582-583.

[56] Shanbhag, D. N. (1968). Some remarks concerning Khatri's results on quadratic forms. *Biometrika*, **55**, 593-595.

[57] Siotani, M., Hayakawa, T. and Fujikoshi, Y. (1985). *Modern Multivariate Statistical Analysis: A Graduate Course and Handbook.* American Sciences Press, Columbus, Ohio.

[58] Smith, J. H. and Lewis, T. O. (1980). Determining the effect of intraclass correlation on factorial experiments. *Comm. Stat. A - Theory and Methods*, **9(13)**, 1353-1364.

[59] Styan, G. P. H. (1969). Notes on the distribution of quadratic forms in singular normal variables. *Technical Report* # 122, University of Minnesota, Minneapolis.

[60] Thomson, G . W. (1955). Bounds for the ratio of range to standard deviation. *Biometrika*, **42**, 268-269.

[61] Tukey, J. W. (1960). A survey of sampling from contaminated distributions. *Contri. to Prob. and Stat.*, Ed. I. Olkin, Stanford University Press, Stanford, CA.

[62] Walsh, J. E. (1947). Concerning the effect of intraclass correlation on certain significance tests. *Ann. Math. Stat.,* **18**, 88-96.

[63] Wishart, J. (1928). The generalized product moment distribution in samples from a normal multivariate population. *Biometrika*, **20**, 32-52.

[64] Wishart, J. and Bartlett, M. S. (1933). The generalized product moment distribution in a normal system. *Proc. Cambridge Phil. Soc.*, **29**, 260-270.

[65] Wolkowicz, H. and Styan, G. P. H. (1979). Extensions of Samuelson's inequality. *The Amer. Stat.*, **33(3)**, 143-144.

AUTOBIOGRAPHICAL STATEMENT

Place of Birth	Kanpur, India
Date of Birth	January 4, 1964
Education	M.Sc. in Statistics, May 1989 Indian Institute of Technology Kanpur, India
Publication	"A New Non-Fourier Tomographic Filter for Image Reconstruction," Rathore, R. K. S. and Vaish, A. K. et al., *Nuclear Technology,* Vol. 85, June 1989.
Employment	Visiting Assistant Professor, 1994-1995 Department of Mathematics University of North Carolina Charlotte, NC
	Lecturer, 1993 - 1994 Department of Mathematics University of Wisconsin Oshkosh, WI
	Graduate Teaching Assistant, 1989 - 1993 Department of Mathematics and Statistics Old Dominion University Norfolk, VA
Awards and Honors	Special Doctoral Research Assistantship Old Dominion University, 1992 - 1993
	Merit Scholarship Indian Institute of Technology, 1984 - 1989
	Phi Kappa Phi honor society