PITMAN ESTIMATION FOR ENSEMBLES AND MIXTURES

by

Shankar S. Srinivasan

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Statistics

APPROVED:

Donald R. Jensen

Robert V. Foutz

Raymond H. Myers

Eric P. Smith

George R. Terrell

October 1995

Blacksburg, Virginia

Key words: Elliptical Symmetry, Mixtures, Pitman Estimation, Repeated Measures.

PITMAN ESTIMATION FOR ENSEMBLES AND MIXTURES

by

Shankar S. Srinivasan

Donald R. Jensen, Chair

Statistics

(ABSTRACT)

This dissertation considers minimal risk equivariant (MRE) estimation of a location scalar μ in ensembles and mixtures of translation families having structured dispersion matrices Σ. The principal focus is the preservation of Pitman's solutions across classes of distributions.

To these ends the cone S_n^+ of positive definite matrices is partitioned into various equivalence classes. The classes $\aleph_C$ are indexed through matrices $\mathbf{C}$ from a class $\mathscr{C}(n)$ comprising positive semidefinite (n×n) matrices with one-dimensional subspace spanned by the unit vector $\mathbf{1}_n' = [1, 1, \ldots, 1]$. Here $\Sigma \in \aleph_C$ has the structure $\Sigma(\gamma) = \mathbf{C} + \gamma \mathbf{1}_n' + \mathbf{1}_n \gamma' - \bar{\gamma} \mathbf{1}_n \mathbf{1}_n'$, for some vector γ such that $\gamma' \mathbf{C}^{-1}_{\mathscr{C}(n)} \gamma < \bar{\gamma}$, where $\mathbf{C}^{-1}_{\mathscr{C}(n)}$ is the Moore-Penrose inverse of $\mathbf{C}$. Of particular interest is the class $\Gamma(n) = \aleph_{\mathbf{B}}$ with $\mathbf{B} = [\mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n']$. In addition, the equivalence classes $\Lambda(\mathbf{w})$ in S_n^+ are indexed through elements of $\mathscr{W}(n)$ containing n-dimensional vectors $\mathbf{w}$ such that $\sum_{i=1}^{n} w_i = 1$, where $\mathbf{w}'\Sigma = c\mathbf{1}_n'$ for some scalar c>0. Of interest is the class $\Omega(n) = \Lambda(n^{-1}\mathbf{1}_n)$, containing equicorrelation matrices in the intersection $\Gamma(n) \cap \Omega(n)$. Ensembles of elliptically contoured distributions having dispersion matrices in the foregoing classes, and mixtures over these, are considered further with regard to Pitman estimation of μ.

For elliptical random vectors $\mathbf{X}$ the Pitman estimator continues to take the generalized least squares form. Further, ensembles of elliptically symmetric distributions having dispersion matrices in $\Omega(n)$ preserve the equivariant admissibility of the sample average $\overline{X}$ under squared error loss. For dispersion matrices Σ in each class $\aleph_C$ the estimator is obtained as a correction of $\overline{X}$ taking the form $\delta_{\Sigma}(\mathbf{X}) = \overline{X} - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\mathbf{X}$, with γ as in the expansion for Σ. This simplifies when $\Sigma \in \Gamma(n)$ to $\delta_{\Sigma}(\mathbf{X}) = \overline{X} - \gamma'\mathbf{e}$, where $\mathbf{e}$ is the vector of residuals $\{e_i = x_i - \bar{x};\ i = 1, 2, ..., n\}$. These results carry over to dispersion mixtures of elliptically symmetric distributions when the mixing measure $\mathbf{G}$ is restricted to the corresponding subsets of S_n^+. The estimators are now given through a dispersion matrix Ψ which is the expectation of Σ over $\mathbf{G}$. For mixing measures over S_n^+, for which each conditional expectation for Σ given $\mathbf{C} \in \mathscr{C}(n)$ is in $\Omega(n)$, $\overline{X}$ is the Pitman estimator for μ for the corresponding mixture distribution. Similar results apply for each linear estimator. In both elliptical ensembles and mixtures over these, the Pitman estimator is shown to be linear and unbiased.

ACKNOWLEDGMENTS

I am indebted to my committee chair, Dr. Donald R. Jensen, for training me in mathematical statistics, patiently helping me out of many deadlocks, and providing encouragement and support throughout my research. I thank professors Foutz, Myers, Smith and Terrell for serving on my committee and for their comments and ideas. Michele Marini helped scan my graphics and helped out with numerous other computer applications. I thank faculty members at the statistics department for maintaining high standards of teaching and consulting in difficult times. Graphics and intuitive explanations provided by the faculty in classrooms and consulting sessions contributed significantly to developments in this dissertation.

I thank the department secretaries Linda Breeding, Dinah Easterly, Betty Higginbotham, Cindy Watson, and Linda Seawell for their able administrative support and assistance to my consulting, teaching and research activities as a graduate student. The geniality of the janitors Mary Bradley, Elaine Cook and Doris Minnik, during their cleaning rounds at the graduate carrels, always lifted up my spirits. I thank the statistics graduate students for allowing me to use the gateway computers at the consulting laboratory for long stretches of time and for being considerate in other ways.

I thank my colleague Shanthi Sethuraman for her friendship in stressful times. I am also grateful to her for her help with running computer software, for her constructive criticism of my professional presentations, and for discussions on topics in statistics through which I learnt a lot. Finally I thank family members in India and in the USA for their love and support throughout my graduate studies at Virginia Tech.

To

Amma and Appa.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

1.1 Background. 1
1.2 Motivation. 4

CHAPTER 2: CONCEPTUAL BACKGROUND

2.1 Introduction: A Characterization of Normality. 7
2.2 Pitman estimation. 11
2.3 Elliptically Symmetric Distributions. 16
2.4 Mixtures. 18
2.5 Overview. 20

CHAPTER 3: MATRICES OF SCALE PARAMETERS

3.1 Introduction. 22
3.2 Helmert Matrices. 25
3.3 Matrices in the Classes $\Xi(\mathbf{A},\mathbf{H})$. 27
3.4 Matrices in the Class $\Gamma(n)$. 45
3.5 Matrices with Equal Linear Functions on Columns. 49
3.6 Matrices with Equal Columns Sums. 59
3.7 Overview. 63

CHAPTER 4: PITMAN ESTIMATION IN ENSEMBLES AND MIXTURES

4.1 Introduction. 66
4.2 The Pitman Estimator for Elliptically Symmetric Distributions. 67
4.3 Results for the Classes $\aleph_C$ and $\Lambda(\mathbf{w})$. 72
4.4 Pitman Estimation: General Results for Mixtures. 79
4.5 Estimation for Mixtures over the Classes $\aleph_C$ and $\Lambda(\mathbf{w})$. 87
4.6 Conditions for the Admissibility of the Sample Average. 98
4.7 Retrospective. 103

CHAPTER 5: CONCLUSIONS AND FURTHER STUDY

5.1 Conclusions. 106
5.2 Further Study. 112

GLOSSARY 115
REFERENCES 117

LIST OF TABLES

Table 5.1	Characterization of classes of positive definite matrices.	106
Table 5.2	Pitman estimators for elliptical ensembles.	106
Table 5.3	Pitman estimates for mixture distributions.	108
Table 5.4	Summary of numerical examples of dispersion structures and the corresponding Pitman estimator for location.	111

LIST OF FIGURES

Figure 2.1	Bivariate normal plot illustrating the Pitman ratio of integrals.	14
Figure 2.2	Density of a mixture of bivariate normals.	19
Figure 3.1	Admissible vectors γ for positive definiteness of $\Sigma = \mathbf{H'AH} + \mathbf{M}(\gamma)$.	33
Figure 3.2	Admissible vectors for positive definiteness of $\Sigma = \Sigma(\gamma)$ for n=2.	48
Figure 3.3	Constraints on γ in $\Re^2$ to generate matrices in $\beta(\mathbf{w},\mathbf{C})$.	55
Figure 3.4	Partition of the set S_n^+ of positive definite matrices.	64
Figure 4.1	Mixture of bivariate normals with unconstrained mixing measure.	85
Figure 4.2	Mixture of bivariate normals with $\mathrm{Dom}(\mathbf{G}) = \Omega(n)$.	90
Figure 4.3	Mixture of bivariate normals with $\mathrm{Dom}(\mathbf{G}) = \Gamma(n)$.	94
Figure 4.4	Mixture of bivariate normals with $\mathrm{Dom}(\mathbf{G}) = \Gamma(n)$ and $\Psi \in \Omega(n)$.	96
Figure 4.5	Illustration of conditions for the sample average to be MRE.	102

Chapter 1

INTRODUCTION

1.1 BACKGROUND

This dissertation has a dual emphasis. Firstly, a classification within the class of positive definite matrices is obtained. Secondly, this classification is applied to obtain conditions for the equivariant admissibility of the sample average. We examine equivariant estimation with emphasis on squared error loss. Theory concerning equivariant estimation is drawn from Pitman (1948) and Lehmann (1983). These results are in Section 2.2. Equivariant estimation is applicable when a random vector has a distribution in a location family. An n-dimensional random vector **X** with such a distribution has distribution function taking the form

$$J(\mathbf{x}; \mu) = F(x_1-\mu, x_2-\mu, \ldots, x_n-\mu).$$

Kagan, Linnik and Rao (1973) have examined linear estimation in this context. They assume that the components of the random vector **X** are independent and that $n>2$. Under these assumptions they find that the unbiased admissibility of the linear estimator for μ characterizes normality. Their results are reviewed in Section 2.1. This context is distinct from that of this dissertation. Independence is relaxed through (i) elliptically symmetric distributional assumptions and through (ii) distributional assumptions in a class of mixtures of elliptically symmetric distributions. Details about these two classes of distributions are in Sections 2.3 and 2.4, respectively.

These mixture distributions are obtained on mixing over the class S_n^+ of positive definite matrices. Mixing measures over subsets of S_n^+ are of particular interest. This motivates the development of a two-way partition of S_n^+ in Chapter 3. The partition is indexed by elements of classes denoted as $\mathscr{W}(n)$ and $\mathscr{C}(n)$. Let $\mathbf{1}_n' = [1, \ 1, \ \ldots, \ 1]$ denote the unit vector in $\Re^n$. Then the class $\mathscr{W}(n)$ contains all n-dimensional vectors $\mathbf{w}$ such that $\mathbf{1}_n'\mathbf{w} = 1$. The class $\mathscr{C}(n)$ contains all n-dimensional real symmetric matrices $\mathbf{C}$ with one-dimensional null-space spanned by the unit vector. Classes $\Lambda(\mathbf{w})$ and $\aleph_C$ consisting of positive definite matrices are defined in terms of elements of $\mathscr{C}(n)$ and $\mathscr{W}(n)$. Figure 3.4 summarizes results concerning the partition of S_n^+. As $\mathbf{C}$ varies over $\mathscr{C}(n)$, the $\aleph_C$ classes partition S_n^+. These classes are studied in Section 3.3. They are depicted as horizontal strips in Figure 3.4. As $\mathbf{w}$ varies over $\mathscr{W}(n)$, the $\Lambda(\mathbf{w})$ classes partition S_n^+. These classes are studied in Section 3.5. They are depicted as vertical strips in Figure 3.4. The two-way classification of positive definite matrices contains subsets of particular interest in statistics. Consider the horizontal strip labeled $\Gamma(n)$ in Figure 3.4. This represents the class of matrices having the structure of Huynh and Feldt (1970); see also Rouanet and Le'pine (1970). Dispersion structures from within this class validate the use of F-ratios in the context of repeated measurement designs. Details pertaining to $\Gamma(n)$ are in Section 3.4. Another class of particular interest is the central vertical strip $\Omega(n)$. This class of matrices, developed in Jensen (1989a), contains positive definite matrices having equal column sums. For elliptically symmetric random vectors, a dispersion structure in $\Omega(n)$ yields admissibility for the sample average. Details pertaining to $\Omega(n)$ are in Section 3.6. Also pertinent is the class of equicorrelation matrices. Elements of these are seen to be scale multiples of matrices in the intersection of $\Gamma(n)$ and $\Omega(n)$. Section 3.7 summarizes results concerning the partition of the set of positive definite matrices as undertaken here.

Admissibility in this dissertation, unless otherwise stated, pertains to admissibility under squared error loss in the class of equivariant estimators . In Chapter 4 admissibility of an estimator for location is examined for the two classes of distributions described

1.2: MOTIVATION

In this section we discuss the relevance of developments in this dissertation to calibration and to computations involving matrices. Consider measuring instruments calibrated through a controlled experiment. Let the model relating the reading R of a measuring instrument to an actual measurement Z be given as

$$\hat{R} = \hat{\alpha} + \hat{\beta} Z$$

where $\hat{\alpha}$ and $\hat{\beta}$ are estimators of the intercept and slope respectively. Readings $\{R_1, R_2, \ldots, R_n\}$ obtained experimentally from the device would then be converted to Z through

$$Y_i = \frac{(R_i - \hat{\alpha})}{\hat{\beta}}.$$

For fixed $\hat{\beta} = \beta$, suppose that $\{R_1, R_2, \ldots, R_n\}$ are mutually uncorrelated and independent of the controlled experiment which generated the calibration. Then we note that

$$\begin{aligned} \mathrm{Cov}(Y_i, Y_j) &= \beta^{-2}\left[\mathrm{Cov}(R_i, R_j) - \mathrm{Cov}(R_i, \hat{\alpha}) - \mathrm{Cov}(R_j, \hat{\alpha}) + \mathrm{Cov}(\hat{\alpha}, \hat{\alpha})\right] \\ &= \beta^{-2}\mathrm{Var}(\hat{\alpha}). \end{aligned}$$

Similarly for $\{Y_i,\ i = 1, 2, \ldots, n.\}$

$$\begin{aligned} \mathrm{Var}(Y_i) &= \beta^{-2}\left[\mathrm{Var}(R_i) + \mathrm{Var}(\hat{\alpha}) - 2\mathrm{Cov}(R_i, \hat{\alpha})\right] \\ &= \beta^{-2}\left[\sigma^2 + \mathrm{Var}(\hat{\alpha})\right] \end{aligned}$$

Thus, the dispersion matrix for $[Y_1, Y_2, \ldots, Y_n]$ is

$$\beta^{-2}\begin{bmatrix} \sigma^2 + \mathrm{Var}(\hat{\alpha}) & \mathrm{Var}(\hat{\alpha}) & \ldots & \mathrm{Var}(\hat{\alpha}) \\ \mathrm{Var}(\hat{\alpha}) & \sigma^2 + \mathrm{Var}(\hat{\alpha}) & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \mathrm{Var}(\hat{\alpha}) & \ldots & \ldots & \sigma^2 + \mathrm{Var}(\hat{\alpha}) \end{bmatrix}.$$

Different controlled experiments to calibrate the device can generate different values of $\hat{\beta}$ and $\mathrm{Var}(\hat{\alpha})$ through differing experimental designs. This provides a motivation for the distributions studied. The Distribution for $\mathbf{Y}$ could be considered as an elliptically symmetric distribution generated as a scale mixture of spherical distributions for $\{R_1, R_2, \ldots, R_n\}$ as $\hat{\beta}$ varies. Alternately this distribution can be modeled as the mixture of elliptically symmetric distributions over the dispersion matrices obtained on varying $\mathrm{Var}(\hat{\alpha})$. For such classes of distributions we study estimation of the common mean.

A possible matrix application is in computations involving positive definite matrices. This exploits properties preserved within the subsets of positive definite matrices studied in Chapter 3. Consider the subsets $\Lambda(\mathbf{w})$ indexed through n-dimensional vectors $\mathbf{w}$ such that $\mathbf{1}_n'\mathbf{w} = 1$. For every element Σ in $\Lambda(\mathbf{w})$ the vector $(\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n)^{-1}\mathbf{1}_n'\Sigma^{-1} = \mathbf{w}$, and every positive definite matrix of the form $\Sigma + \lambda(\mathbf{1}_n\mathbf{1}_n')$ for scalar λ is contained in $\Lambda(\mathbf{w})$. Thus one may choose a positive definite matrix $\mathbf{S}_2$ instead of $\mathbf{S}_1$ when computing $\mathbf{w} = (\mathbf{1}_n'\mathbf{S}_1^{-1}\mathbf{1}_n)^{-1}\mathbf{1}_n'\mathbf{S}_1^{-1}$ provided that $\mathbf{S}_2 = \mathbf{S}_1 + \lambda(\mathbf{1}_n\mathbf{1}_n')$. Such a replacement may increase accuracy in computations when $\mathbf{S}_2$ is less ill-conditioned than $\mathbf{S}_1$. Consider $\mathbf{S}_1 = \begin{bmatrix} 5 & 3 & 2 \\ 3 & 3 & 1 \\ 2 & 1 & 1 \end{bmatrix}$ and $\mathbf{S}_2 = \mathbf{S}_1 + 2(\mathbf{1}_n\mathbf{1}_n') = \begin{bmatrix} 7 & 5 & 4 \\ 5 & 5 & 3 \\ 4 & 3 & 3 \end{bmatrix}$. To gauge conditioning we compute the condition number

$$C(\Sigma) = \left[\sum_{i=1}^{n} \xi_i^2\right]\left[\sum_{i=1}^{n}\left(1/\xi_i^2\right)\right]$$

where $\{\xi_i;$ for $i = 1,2, \ldots, n\}$ are the n eigenvalues of the positive definite matrix Σ. The function $C(\Sigma)$ is from a class of functions invariant under orthogonal transformations as in Marshall and Olkin (1965). For $\mathbf{S}_1$ the ordered eigenvalues are $\{0.127, 1.0, 7.873\}$, and for $\mathbf{S}_2$ the ordered eigenvalues are $\{0.5193, 1.0, 13.4807\}$. These yield $C(\mathbf{S}_1) = 3968.59$, which is much larger that $C(\mathbf{S}_2) = 862.603$. The matrix $\mathbf{S}_1$ is ill-conditioned as compared to $\mathbf{S}_2$. Finding better conditioned matrices may improve accuracy in computations when they need to be computed at low precision possibly due to software/hardware constraints. The study of subsets of positive definite matrices in this dissertation may provide ways of increasing accuracy in such contexts.

We apply developments about subsets of positive definite matrices primarily to problems in statistical inference. It is shown that the appropriateness of linear estimation for location extends well beyond normality. The Aitkin (1934) estimator, which is linear, is found to be minimal risk equivariant (MRE) under squared error loss in the wider class of elliptically symmetric distributions. Certain nonelliptical classes of distributions also validate linear estimation for location. These distributions arise as mixtures over the subsets of matrices studied. The next chapter reviews MRE estimation and the distributions to which it is applied.

Chapter 2

CONCEPTUAL BACKGROUND

2.1 INTRODUCTION: A CHARACTERIZATION OF NORMALITY

This dissertation develops a classification for positive definite matrices. This classification is applied in turn to a problem in mathematical statistics concerning linear estimation in certain classes of nonstandard distributions. Relevant theories on estimation and on these distributions are excerpted in this chapter. First concepts in estimation theory are reviewed, as developed in Lehmann (1983), Rohatgi (1976), Kagan, Linnik, and Rao (1973), and elsewhere. The definitions in this section are based on those references, beginning with the loss in estimating an unknown parameter or parametric function. The location parameter μ is to be estimated in certain location families of n-dimensional distributions. The space of n-dimensional vectors is $\mathfrak{R}^n$. For $\mathbf{x} \in \mathfrak{R}^n$, $\delta(\mathbf{x})$ denotes an estimator. The loss in using $\delta(\mathbf{x})$ to estimate μ is $L[\delta(\mathbf{x}),\mu]$. Loss functions examined in this dissertation are the squared error and Laplace loss functions, given by

$$L[\delta(\mathbf{x})-\mu] = [\delta(\mathbf{x})-\mu]^2,$$

and

$$L[\delta(\mathbf{x})-\mu] = |\delta(\mathbf{x})-\mu|$$

respectively.

The risk of an estimator is its expected loss as given by

$$R(\delta,\mu) = \mathscr{E}\{L[\delta(\mathbf{x})-\mu]\},$$

whenever the expectation is well defined. Estimation theory seeks to determine whether an estimator with minimal risk exists within a given class of estimators. Such an estimator is called admissible. A definition of admissibility is in the following.

Definition 2.1: Let $\mathscr{A}$ denote a class of estimators of a parameter μ; and let δ be an estimator contained in $\mathscr{A}$. Then the estimator δ is admissible within the class $\mathscr{A}$ for a given loss function, if there exists no estimator δ' in $\mathscr{A}$ for which

$$R(\delta',\mu) \leq R(\delta,\mu) \text{ for all } \mu$$

with strict equality for at least one μ.

Of primary interest here is the class of equivariant estimators. If a sample vector $\mathbf{x}$ is translated element-wise, then an equivariant estimate is translated by the same amount. This is stated in Definition 2.2, where $\mathbf{1}_n$ again denotes the n-dimensional unit vector.

Definition 2.2: An estimator δ is contained in the class $\mathscr{I}$ of equivariant estimators for a scalar parameter μ if and only if

$$\delta(\mathbf{x} + a\mathbf{1}_n) = \delta(\mathbf{x}) + a$$

for all scalars a.

A procedure for obtaining a minimal risk equivariant estimator is described in the next section. Equivariant estimation uses closure of a location family of distributions under scalar translations. The definition of a location family is in the following.

Definition 2.3: For an n-dimensional random vector X let the cumulative distribution function $F(\mathbf{x};\mu)$ take the form

$$F(\mathbf{x};\mu) = F(x_1-\mu, x_2-\mu, \ldots ,x_n-\mu).$$

Then for each such distribution function, the totality of distributions obtained as μ varies from $-\infty$ to $+\infty$ constitutes an n-dimensional location family.

When a density exists for a member of a location family, it takes the form

$$f(\mathbf{x};\mu) = f(x_1-\mu, x_2-\mu, \dots ,x_n-\mu),$$

where μ is the location parameter. Each distribution considered in this dissertation is in a location family for which linear estimators continue to be admissible among equivariant estimators. In contrast to Kagan, Linnik, and Rao (1973), who assume independence, components of random vectors considered here are dependent. The findings under independence are reviewed in the next subsection.

A Characterization of Normality

Kagan, Linnik, and Rao (1973) assume that the components $[X_1, X_2, .. , X_n]$ of an n-dimensional random vector **X** are independent having distribution functions $\{F_1(x_1-\mu), F_2(x_2-\mu), \dots ,F_n(x_n-\mu)\}$ in one-dimensional location families with common location μ. This implies that joint distributions for **X** take the form

$$F(\mathbf{x};\mu) = \prod_{i=1}^{n} F_i(x_i - \mu).$$

These are members of n-dimensional location families as in Definition 2.4. Further assumptions are the existence of finite first and second moments for each component X_i of **X**. When the second central moment for the ith component is σ_i^2, then the optimal linear estimate for μ is

$$\hat{L} = \sum_{i=1}^{n} c_i x_i \text{ for } c_i = \frac{1/\sigma_i^2}{\sum_{i=1}^{n} 1/\sigma_i^2} \qquad (2.1.1)$$

Conditions for the admissibility of this estimate within the larger class of unbiased estimators are given in Theorem 7.4.1 of Kagan, Linnik, and Rao (1973). That theorem applies to the class of equivariant as well as unbiased estimators. Theorem 2.1.1 restates their theorem for equivariant estimators.

Theorem 2.1.1: For $n \geq 3$ let the components $\{X_1, X_2, .. , X_n\}$ of a random vector $\mathbf{X}$ be mutually independent with distribution functions $\{F_i(x_i - \mu); \text{ for } i = 1, 2, .. ,n\}$. Assume $\mathcal{E}[X_i] = \mu$, for all i; that second order moments exist; and consider the optimal linear estimator $\hat{L}$ in (2.1.2). Then this estimator is admissible in the class of all equivariant estimators under squared error loss if and only if each X_i is normally distributed.

This result limits to linear estimation for a location parameter under squared error loss. For nonnormal n-dimensional location families, the independence assumption through Theorem 2.1.1 implies nonlinearity of an admissible estimator. Thus normality is necessary and sufficient for the equivariant admissibility of a linear estimator for location. However, relaxing independence reopens the question of admissibility of linear estimation for location. Here independence is relaxed in favor of elliptically symmetric distributions, and distributions obtained as mixtures of these. These classes are reviewed in Sections 2.3 and 2.4. The next section reviews results in equivariant estimation due to Pitman (1948) and Lehmann (1983).

2.2 PITMAN ESTIMATION:

The Pitman estimator is the minimum risk equivariant (MRE) estimator under squared error loss. Definition 2.2 states the condition for equivariance. When $\mathbf{x} \in \Re^n$ is subjected to a scalar translation, the equivariant estimator is similarly translated. The set of all scalar translations form a group, i.e., a collection that is closed under a binary operation ρ, acting associatively, containing a unique identity element, and a unique inverse. This and other definitions are as in Lehmann (1983). The set of all scalar translations with addition as the binary relation is the principal group to be considered here. The group of orthogonal matrices under matrix multiplication is used in Chapter 3. For a set $\mathcal{X}$ a group $\mathcal{G}$ provides a set of transformations of the form $g\mathbf{x}$ mapping back into $\mathcal{X}$ for $g \in \mathcal{G}$ and $\mathbf{x} \in \mathcal{X}$. The concept of an orbit is defined in the following.

Definition 2.4: Let G be a group acting on $\mathcal{X}$. Then the set

$$\mathcal{O}(\mathbf{x}_0) = \{\mathbf{x} \in \mathcal{X} | \mathbf{x} = g\mathbf{x}_0, g \in \mathcal{G}\}$$

obtained on successive transformations of $\mathbf{x}_0$ by the elements of $\mathcal{G}$, is an orbit containing $\mathbf{x}_0$.

An orbit under translations is shown in Figure 2.1, where scalar translations of the form $\mathbf{x} = \mathbf{x}_0 + a\mathbf{1}_n$, for $a \in \Re^1$ and $\mathbf{x}_0 = [5,4]'$, trace the straight line through [5,4] as the orbit containing [5,4]. Further comments on Figure 2.1 are in discussions pertaining to the Pitman representation later in this section. Functions which take the same value on each orbit are called invariant functions. A maximal invariant takes distinct values on distinct orbits, and hence the range of a maximal invariant can be used to index the orbits.

This is applied in Chapter 3 in the context of the real orthogonal group and the following definition.

Definition 2.5: A function ν is maximal invariant under $\mathbf{G}$ if and only if $\nu(\mathbf{x}_1) = \nu(\mathbf{x}_2)$ implies that $\mathbf{x}_1 = g\mathbf{x}_2$ for some element g of the group $\mathbf{G}$.

Under the group of scalar translations the (n-1)-dimensional vector $\mathbf{Y}$ with elements $\{Y_i = X_i - X_n \text{for} 1 = 1,2,\ldots,n-1\}$, is maximal invariant. The Pitman estimator under quadratic loss can be obtained as a correction to any chosen equivariant estimator $\delta_0(\mathbf{X})$. The correction is given by the conditional mean $\mathscr{E}_0[\delta_0(\mathbf{X})|\mathbf{Y} = \mathbf{y}]$ when $\mu = 0$. A similar expression holds for Laplace loss. These results are in Theorem 1.3 and Corollary 1.1 of Lehmann (1983, Chapter 3). They are restated in the following.

Theorem 2.2.1: (a) Let a random vector $\mathbf{X} = [X_1, X_2, .. X_n]'$ have a distribution in an n-dimensional location family; let $\mathbf{Y}$ be the (n-1)-dimensional vector of differences with $\{Y_i = X_i - X_n \text{for} 1 = 1,2,\ldots,n-1\}$; and assume that there exists an equivariant estimator δ_0 with finite risk under squared error loss. The minimal risk equivariant (MRE) estimator for the location parameter μ is given by

$$\delta(\mathbf{X};y) = \delta_0(\mathbf{X}) - \mathscr{E}_0[\delta_0(\mathbf{X})|\mathbf{Y} = \mathbf{y}]$$

where $\mathscr{E}_0[\delta_0(\mathbf{X})|\mathbf{Y} = \mathbf{y}]$ is the conditional mean of $\delta_0(\mathbf{X})$ when $\mu = 0$.

(b) Consider Laplace loss. When δ_0 has finite risk, then an MRE estimator for μ is given by

$$\delta(\mathbf{X};y) = \delta_0(\mathbf{X}) - \mathbf{M}_0[\delta_0(\mathbf{X})|\mathbf{Y} = \mathbf{y}]$$

where $\mathbf{M}_0[\delta_0(\mathbf{X})|\mathbf{Y} = \mathbf{y}]$ is any median of the condition distribution of $\delta_0(\mathbf{X})$ when $\mu = 0$.

A one-to-one relationship exists between any two maximal invariants. Hence the distribution of $\delta_0(\mathbf{X})$ conditional on any maximal invariant equals that obtained on conditioning on any other. In Chapter 4 the differences **y** in Theorem 2.2.1 are replaced by an (n-1)-dimensional vector **d**. For $\mathbf{x} \in \Re^n$, the vector **d** is obtained as a set of scaled contrasts on **x**. This is shown to be maximal invariant under scalar translations in Lemma 4.2.1. Pitman (1948) represented the MRE estimator for μ under squared error loss as a ratio of integrals as in the following theorem.

Theorem 2.2.2: Let a random vector **X** have a density of the form

$$f(\mathbf{x};\mu) = f(x_1-\mu, x_2-\mu, \dots ,x_n-\mu).$$

Then the MRE estimator for μ under squared error loss is given by

$$\delta(\mathbf{x}) = \frac{\int_{-\infty}^{\infty} t f(\mathbf{x} - t\mathbf{1}_n)dt}{\int_{-\infty}^{\infty} f(\mathbf{x} - t\mathbf{1}_n)dt}.$$

A proof to this theorem is given in Chapter 3 of Lehmann (1983). This result supports a simple graphical identification of the Pitman estimator as in Figure 2.1. The argument $(\mathbf{x} - t\mathbf{1}_n)$ of f(.) may be interpreted as a scale translation of (-t) on a sample vector $\mathbf{x} \in \Re^n$. Varying t traces the straight line parallel to the equiangular line which contains **x**. We evaluate $f(\mathbf{x} - t\mathbf{1}_n)$ as a function of t with **x** fixed. The function

$$g(t) = \frac{f(\mathbf{x} - t\mathbf{1}_n)}{\int_{-\infty}^{\infty} f(\mathbf{x} - t\mathbf{1}_n)dt}$$

scales $f(\mathbf{x} - t\mathbf{1}_n)$ to a density. The Pitman estimator in Theorem 2.2.2 can then be identified as the mean $\int_{-\infty}^{\infty} t g(t)dt$ for this density. Figure 2.1 depicts a bivariate normal distribution with $\Sigma = \begin{bmatrix} 11 & 7 \\ 7 & 5 \end{bmatrix}$. In this figure the line through [5,4] is the orbit containing

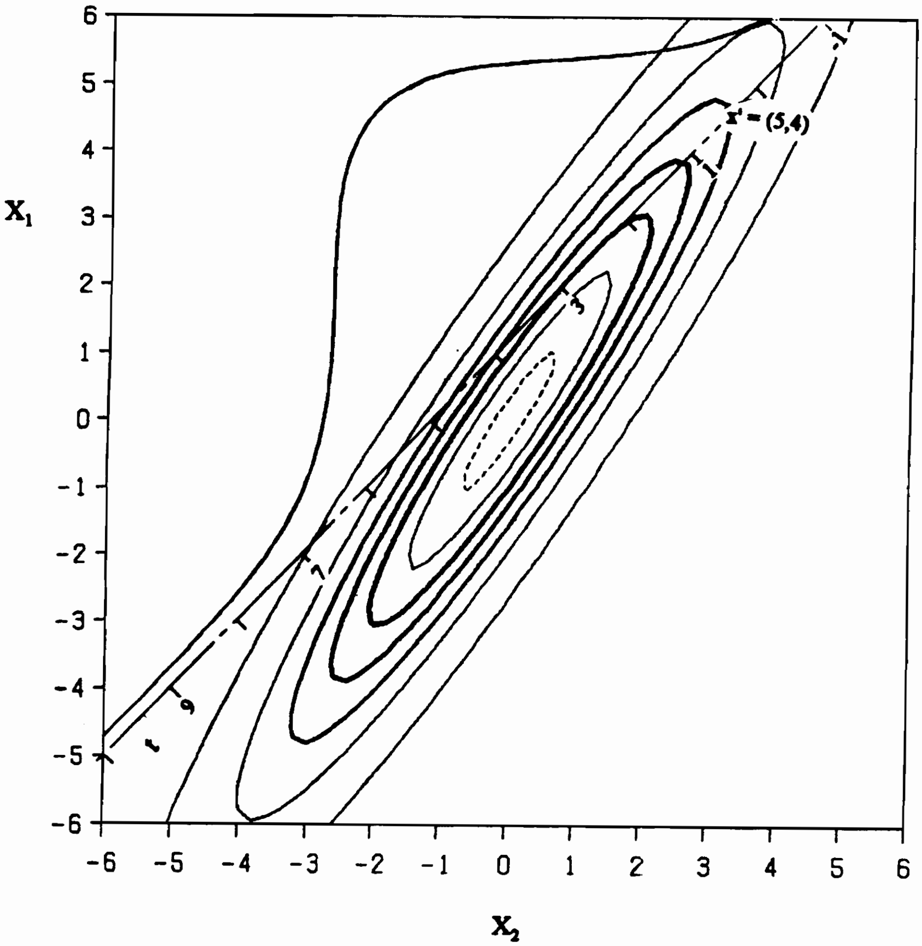

Figure 2.1: Bivariate normal distribution with dispersion matrix $\begin{bmatrix} 11 & 7 \\ 7 & 5 \end{bmatrix}$. The overlaid univariate plot is that for $g(t) = \dfrac{f(\mathbf{x} - t\mathbf{1}_n)}{\int_{-\infty}^{\infty} f(\mathbf{x} - t\mathbf{1}_n)dt}$. The expected value of t for this gives the Pitman estimate.

this point and the overlaid univariate plot is the function g(t). In later developments the Pitman estimate is given as the Aitkin (1934) estimate $\delta(\mathbf{x}) = (\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n)^{-1}\mathbf{1}_n'\Sigma^{-1}\mathbf{x}$, where $\mathbf{1}_n$ is the unit vector. For $\mathbf{x}' = \begin{bmatrix}5 & 4\end{bmatrix}$, $\delta(\mathbf{x}) = 3$. By inspecting Figure 2.1 we note that this is the mean value of g(t).

In Chapter 4 MRE estimators are obtained by using the corrections provided in Theorem 2.2.1. The equivariant estimator $\overline{X}$ is corrected by its conditional mean given the value taken by a maximal invariant. Examples illustrating the MRE estimator are presented for elliptically symmetric distributions and for bivariate mixture distributions. The Pitman(1948) representation in Theorem 2.2.2 can provide a quick graphical analog verifying the bivariate examples. The next section reviews elliptically symmetric distributions.

2.3 ELLIPTICALLY SYMMETRIC DISTRIBUTIONS

Elliptically symmetric distributions derive from spherical distributions by affine transformations. A particular distribution is specified as $E_n(\mu, \Sigma, \phi)$ through an n-dimensional location parameter μ, a positive semidefinite matrix Σ of order n and rank $k \leq n$, and a function ϕ from a class $\Phi(k)$ of admissible characteristic functions. Characteristic functions for elliptically symmetric distributions take the form

$$\mathbf{I_X}(\mathbf{t}) = e^{it'\mu}\phi(\mathbf{t}'\Sigma\mathbf{t}).$$

From this expression we see that the function ϕ is the characteristic function of the spherically symmetic distribution with location $\mathbf{0}$ from which the elliptically symmetric distribution derives on linear transformations. The normal distribution, which we denote by $N_n(\mu, \Sigma)$, is a member of the class with $\phi(.) = e^{-0.5(.)}$. The class $\Phi(k)$ derives from developments in Schoenberg (1938). Each function ϕ in $\Phi(k)$ is related to a radial distribution R on $[0, \infty)$ through the characteristic function $\tau(\|\mathbf{t}\|^2)$ of the k-dimensional random vector $\mathbf{U}$ uniformly distributed on a unit sphere in $\Re^k$. When R has a distribution function F then a function ϕ is a characteristic function in $\Phi(k)$ if and only if

$$\phi(u) = \int_0^\infty \tau(r^2 u)dF(r) \text{ for } u > 0.$$

In part (a) of the following theorem each distribution $E_n(\mu, \Sigma, \phi)$ is characterized through R, $\mathbf{U}$ and a matrix $\mathbf{B}$ of order (n×k) which is a rank factorization of Σ. Part (b) states properties of elliptical random vectors under linear transformations. Marginal distributions are in part (c). Conditional properties are in part (d). Part (e) gives expectations and variances for these random vectors. The results in parts (a), (d) and (e) are demonstrated in Cambanis, Huang, and Simons (1981). Proofs for results in parts (b) and (c) are in Fang and Zhang (1990). In these results *L* (.) means that the argument of *L* has the distribution specified.

Theorem 2.3.1: (a) Let Σ be of rank k. Then $L(\mathbf{X}) = E_n(\mu, \Sigma, \phi)$ if and only if

$$L(\mathbf{X}) = L(\mu + \mathrm{RUB}).$$

(b) Let $\mathbf{A}$ be a matrix of order $n \times m$ and let $\mathbf{a} \in \Re^m$, then

$$L(\mathbf{A'X} + \mathbf{a}) = E_m(\mathbf{A'}\mu + \mathbf{a}, \mathbf{A'}\Sigma\mathbf{A}, \phi).$$

(c) For $\mathbf{X}$ partitioned as $\mathbf{X} = [\mathbf{X}_1', \mathbf{X}_2']'$ with $\mathbf{X}_i$ $(n_i \times 1)$, $\mu = [\mu_1', \mu_2']'$ with μ_i $(n_i \times 1)$, and $\Sigma = [\Sigma_{ij}]$ with Σ_{ij} of order $(n_i \times n_j : i,j = 1,2)$, the marginal distributions are given by

$$L(\mathbf{X}_i) = E_{n_i}(\mu_i, \Sigma_{ii}, \phi).$$

(d) For $\mathbf{X}_1$ and $\mathbf{X}_2$, as defined above, the conditional distribution of $\mathbf{X}_1$, given that $\mathbf{X}_2$ takes the value $\mathbf{c}$, is given by;

$$L(\mathbf{X}_1 \mid \mathbf{X}_2 = \mathbf{c}) = E_{n_1}(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{c} - \mu_2), \Sigma_{11.2}, \phi^*)$$

where $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{21}\Sigma_{22}^{-1}\Sigma_{12}$, and $\phi^* = \int_0^\infty \tau(r^2 u)\,dF^*(r)$ with F^* given by the c.d.f. of the distribution

$$L(R^*) = L[(R^2 - (\mathbf{c}-\mu_2)\Sigma_{22}^{-1}(\mathbf{c}-\mu_2)' \mid X_2 = \mathbf{c})]$$

(e) If $E|X_i| < \infty$ $(i = 1, 2, \ldots, n)$, then $E(\mathbf{X}) = \mu$ and if $E(\mathbf{X'X}) < \infty$, then $\mathrm{Var}(\mathbf{X}) = \alpha\Sigma$ for some $\alpha = (k)^{-1}\mathcal{E}(R^2)$ when $\mathrm{rank}(\Sigma) = k \le n$. For the normal distribution $\alpha = 1$.

For these elliptically symmetric distributions, we henceforth assume that the scalar components of the random vector $\mathbf{X}$ have the same mean μ. Also Σ is assumed to have full rank. A member of this restricted class of distributions is denoted by

$$E_n(\mu\mathbf{1}_n, \Sigma, \phi)$$

The characteristic function of a distribution for this restricted set works out to

$$\mathbf{I}_X(\mathbf{t}) = e^{i\mathbf{t}'(\mu\mathbf{1}_n)}\phi(\mathbf{t}'\Sigma\mathbf{t}) = e^{i\left(\sum_{i=1}^{n} t_i\right)\mu}\phi(\mathbf{t}'\Sigma\mathbf{t}).$$

To see that each such elliptical random vector X is in a location family consider $\mathbf{Y}$ = $\mathbf{X}$+a$\mathbf{1}_n$. Using properties of characteristic functions it follows that

$$\mathbf{I}_{\mathbf{Y}}(\mathbf{t}) = e^{i\left(\sum_{i=1}^{n} t_i\right)\mu^*} \phi(\mathbf{t}'\Sigma\mathbf{t}),$$

where $\mu^* = \mu + a$. Thus each elliptically symmetric random vector is in a family of distributions which is closed under scalar translations, and MRE estimation is applicable.

Chapter 4 develops MRE estimation for ensembles of distributions obtained as Σ varies within subsets of positive definite matrices and for mixtures of these having absolutely continuous densities. For a distribution $L(\mathbf{X}) = E_n(\mu\mathbf{1}_n, \Sigma, \phi)$, when a density exists, it takes the form

$$k(\mathbf{x}; \mu\mathbf{1}_n, \Sigma, \phi) = f_\phi((\mathbf{x} - \mu\mathbf{1}_n)' \Sigma^{-1}(\mathbf{x} - \mu\mathbf{1}_n)),$$

where the function f_ϕ has a unique correspondence to ϕ. This density features in the mixture densities described in the next section.

2.4 MIXTURES

The class of distributions obtained as mixtures of elliptically symmetric distributions have densities taking the form

$$h(\mathbf{x}; \mu, \phi, \mathbf{G}) = \int_{S_n^+} f_\phi\left[(\mathbf{x} - \mu\mathbf{1}_n)'\Sigma^{-1}(\mathbf{x} - \mu\mathbf{1}_n)\right] d\mathbf{G}(\Sigma)$$

where $f_\phi(.)$ is the density of the elliptically symmetric components in the mixture; $\mathbf{G}$ is a probability measure over S_n^+; and μ and ϕ are as defined earlier. A random vector $\mathbf{X}$ with a mixture distribution is succinctly specified as $L(\mathbf{X}) = EM_n(\mu, \phi, \mathbf{G})$ through μ, ϕ, and $\mathbf{G}$. The density h(.) depends on $\mathbf{x}$ through differences of the form $\mathbf{x}-\mu\mathbf{1}_n$. Hence each mixture distribution belongs in an n-dimensional location family. Estimation for these mixtures is developed in Chapter 4 when the domain of $\mathbf{G}$ is restricted to subsets of S_n^+ described in Chapter 3. Figure 2.2 gives an example of a mixture distribution, where $\mathbf{G}$ is

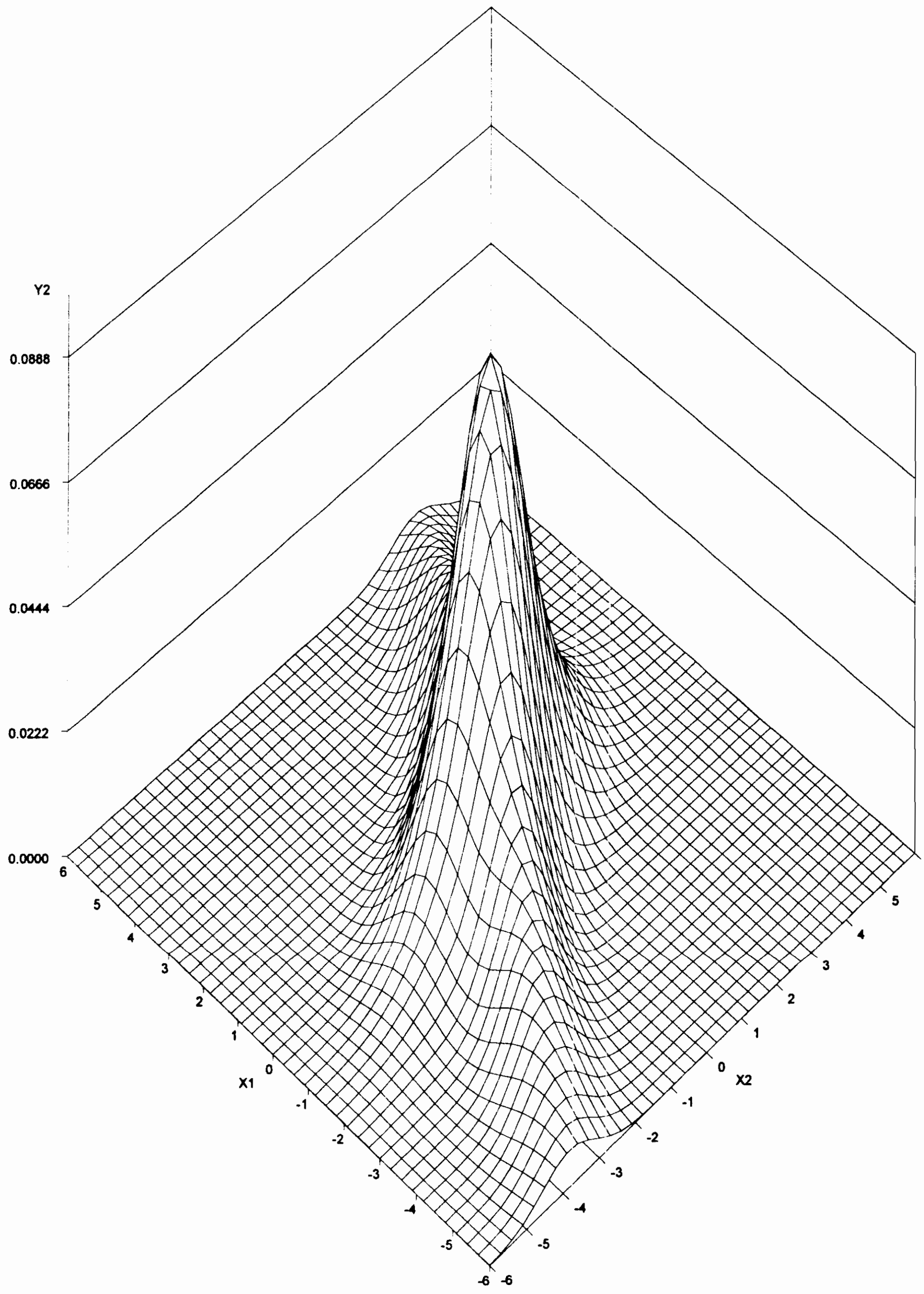

Figure 2.2: Mixture of bivariate normal distributions for which **G** assigns 0.5 probability to $\Sigma_1 = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}$ and to $\Sigma_2 = \begin{bmatrix} 11 & 7 \\ 7 & 5 \end{bmatrix}$.

a discrete distribution which gives equal probability of 0.5 over two dispersion matrices as identified in the figure.

2.5: OVERVIEW

This section summarizes concepts introduced in this chapter. Basic ideas are presented in Section 2.1. For a sample vector $\mathbf{x} \in \Re^n$, an equivariant estimator $\delta(\mathbf{x})$ for a location parameter μ is one for which

$$\delta(\mathbf{x}+t\mathbf{1}_n) = \delta(\mathbf{x}) + t,$$

for all scalar t. A family of distributions closed under scalar translations constitutes a location family. Equivariant estimation is applicable when distributions belong to location families. In later developments minimal risk equivariant (MRE) estimators are derived under nonstandard distributional assumptions. Under the assumption of independence, and for $n \geq 3$, Kagan, Linnik, and Rao (1973) have examined distributions in location families. Admissibility of a linear estimate under squared error is seen to characterize normality. In contrast we relax independence in favor of elliptically symmetric distributions and distributions obtained as mixtures of these.

Section 2.2 gives the procedure to be used to derive MRE estimators using a maximal invariant function of a sample vector under the group of scalar translations. The maximal invariant is the vector $\mathbf{d}$ containing the (n-1) scaled orthogonal contrasts on $\mathbf{x}$. The MRE estimator under squared error loss is obtained by correcting any equivariant estimator $\delta_0(\mathbf{x})$ by the conditional mean $\mathscr{E}_0[\delta_0(\mathbf{x})|\mathbf{D} = \mathbf{d}]$ when $\mu = 0$.

The classes of distributions for which we derive MRE estimators are in Sections 2.3 and 2.4. Applicable properties of elliptically symmetric distributions are those under affine transformations; and those for marginal and conditional distributions. These are similar to the multinormal, which is a special case. Mixtures are obtained by mixing elliptical distributions over the class S_n^+ of positive definite matrices. These distributions are studied under restrictions on the domain of the mixing measure to subsets of S_n^+. These subsets of positive definite matrices are developed in the next chapter.

Chapter 3

MATRICES OF SCALE PARAMETERS

3.1 INTRODUCTION

To motivate the utility of various classes of positive definite matrices, we preview the general form of the Pitman estimator for μ in the mixtures to be studied. More detailed definitions and properties then follow.

Recall that a mixture density h(.) follows on mixing over dispersion matrices as

$$h(\mathbf{x};\, \mu,\, \phi,\, \mathbf{G}) = \int_{S_n^+} f_\phi\left[(\mathbf{x}-\mu\mathbf{1}_n)'\Sigma^{-1}(\mathbf{x}-\mu\mathbf{1}_n)\right]d\mathbf{G}(\Sigma)$$

where $f_\phi(.)$ is the density of elliptically symmetric components in the mixture, and $\mathbf{G}$ is a probability measure over S_n^+. A typical member of the class of such distributions is denoted by $EM_n(\ \mu\mathbf{1}_n,\ \mathbf{G},\ \phi\ )$. For a sample vector $\mathbf{x}$ from a distribution in this class, the Pitman estimate for location is denoted as $\delta_{\mathbf{G},\phi}(\mathbf{x};\mathbf{d})$. This estimate is expressed in terms of matrices from the class $\mathscr{H}(n)$ as defined in Section 3.2, consisting of matrices of order $(n-1)\times n$ which are orthonormal completions of the unit vector. For any $\mathbf{H} \in \mathscr{H}(n)$, the Pitman estimate is obtained on extracting information provided by the maximal invariant $\mathbf{d} = \mathbf{Hx}$, as

$$\delta_{\mathbf{G},\phi}(\mathbf{x};\mathbf{d}) = \frac{\int_{S_n^+} \delta_\Sigma(\mathbf{x})g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d})d\mathbf{G}(\Sigma)}{b(\mathbf{d};\mathbf{G},\phi)}, \tag{3.1.1}$$

where $\delta_\Sigma(\mathbf{x})$ denotes the Pitman estimate for each elliptically symmetric distribution in the mixture; $g_\phi(.)$ is the density of $\mathbf{D} = \mathbf{HX}$ when $L(\mathbf{X}) = E_n(\mu\mathbf{1}_n, \Sigma, \phi)$ and b(.) is the corresponding density for the mixture. For unrestricted probability measures $\mathbf{G}$ on S_n^+ the Pitman estimate does depend on the maximal invariant $\mathbf{d}$ through the parameters $\mathbf{G}$ and ϕ.

Restrictions of the support of $\mathbf{G}$ to subsets of S_n^+, as developed in this chapter, are of particular interest in Chapter 4. Then the density $g_\phi(.)$ in (3.1.1) is invariant as Σ varies, yielding a simpler form for the Pitman estimate. This drives our interest in subsets of S_n^+ taking the form $\Xi(\mathbf{A}, \mathbf{H})$ as studied in Section 3.3, with conditions for membership in each class as in Definition 3.2. Under these restrictions on $\mathbf{G}$ the expression (3.1.1) for the Pitman estimate becomes

$$\delta_{\mathbf{G},\phi}(\mathbf{x};\mathbf{d}) = \int_{\Xi(\mathbf{A},\,\mathbf{H})} \delta_\Sigma(\mathbf{x})d\mathbf{G}(\Sigma).$$

Amongst subsets $\Xi(\mathbf{A}, \mathbf{H})$of S_n^+, a subset to be denoted by $\Gamma(n)$ is of special interest. Mixtures over this subset have stronger invariance properties than for other subsets $\Xi(\mathbf{A}, \mathbf{H})$ studied here, in that the distribution of the maximal invariant $\mathbf{D} = \mathbf{HX}$ is spherically symmetric in addition to the invariance of the density $g_\phi(.)$ in (3.1.1). Matrices in $\Gamma(n)$ are characterized in Definition 3.4.

The expression (3.1.1) motivates our interest in another kind of restriction on the domain of $\mathbf{G}$. For each elliptically symmetric component in the mixture, the Pitman estimator $\delta_\Sigma(\mathbf{x})$ is a weighted average of the elements of $\mathbf{x}$, i.e., $\delta_\Sigma(\mathbf{x}) = \mathbf{w}'\mathbf{x}$ for some n-dimensional vector $\mathbf{w}$ of constants such that $\sum_{i=1}^{n} w_i = 1$. Consider subsets of S_n^+ for which the Pitman estimate is $\delta_\Sigma(\mathbf{x}) = \mathbf{w}'\mathbf{x}$ for all elements of that subset with $\mathbf{w}$ fixed. Such classes of subsets are denoted as $\Lambda(\mathbf{w})$ and are characterized in Definition 3.5. When $\mathbf{G}$ is restricted to a subset $\Lambda(\mathbf{w})$ then $\delta_\Sigma(\mathbf{x})$ factors out of the integral in the numerator of (3.1.1), showing that $\delta_{\mathbf{G},\phi}(\mathbf{x};\mathbf{d}) = \mathbf{w}'\mathbf{x}$ is the Pitman estimate for all mixtures over $\Lambda(\mathbf{w})$.

Amongst subsets $\Lambda(\mathbf{w})$, a subset denoted by $\Omega(n)$ is of special interest. For mixtures of elliptically symmetric distributions over $\Omega(n)$, the Pitman estimate is the sample average. We show that the admissibility of the sample average under squared error loss functions is preserved for all mixtures over $\Omega(n)$. Matrices in $\Omega(n)$ are characterized in Definition 3.6. Results pertaining to $\Omega(n)$ and other subsets of positive definite matrices are developed using Helmert matrices as presented in the following section.

3.2 HELMERT MATRICES

Helmert matrices are denoted by **K** and take the form

$$\mathbf{K}_{(n\times n)} = \begin{bmatrix} (1/\sqrt{n})\mathbf{1}_n' \\ \mathbf{H} \end{bmatrix} \tag{3.2.1}$$

where **H** is any semiorthogonal completion of order (n-1)×n of the scaled unit vector, whose rows comprise coefficients for a set of (n-1) orthonormal contrasts. Applying **H** to a random vector $\mathbf{x} \in \Re^n$ maps it onto an (n-1)-dimensional space orthogonal to the equiangular line. The collection of such matrices is denoted by $\mathscr{H}$ as in Definition 3.1 that follows. An example of the matrix **K** for n=3 is given below.

Example 3.1: For $\mathbf{K} = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ -1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \end{bmatrix}$, $\mathbf{H} = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ -1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \end{bmatrix}$.

Definition 3.1: Let $\mathbf{I}_{(n-1)}$ be the identity matrix of order (n-1). A matrix **H** of order (n-1)×n belongs to the class $\mathscr{H}$ if and only if

(a) $\mathbf{HH}' = \mathbf{I}_{(n-1)}$, and

(b) $\mathbf{H1}_n = \mathbf{0}$, the zero vector.

Properties of matrices in $\mathscr{H}$ are given in the Lemma that follows. The proof for the result in (c) is provided.

Lemma 3.2.1: (a) For all $\mathbf{H} \in \mathscr{H}$, $\mathbf{H'H} = [\, \mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n']$ where $\mathbf{I}_n$ is the identity matrix of order n.

(b) For a vector $\mathbf{x}$ of observations it follows that $\mathbf{H'Hx} = \mathbf{e}$, where $\mathbf{e}$ is the vector of ordinary least squares (OLS) residuals with $e_i = x_i - \bar{x}$.

(c) If $\mathbf{H}_1$ belongs to $\mathscr{H}$, then the necessary and sufficient condition for $\mathbf{H}_2$ to be an element of $\mathscr{H}$ is that $\mathbf{H}_2 = \mathbf{OH}_1$ for some $\mathbf{O} \in \mathscr{O}(n-1)$, the set of orthogonal matrices of order $(n-1)$.

Proof to (c): Imposing the assumption $\mathbf{H}_2 = \mathbf{OH}_1$ we note that

$$\mathbf{H}_2\mathbf{H}_2' = \mathbf{OH}_1\mathbf{H}_1'\mathbf{O}' = \mathbf{O}\,\mathbf{I}_{(n-1)}\mathbf{O}' = \mathbf{I}_{(n-1)}.$$

Hence the matrix $\mathbf{H}_2$ satisfies the condition in part (a) of Definition 3.1. Part (b) implies that $\mathbf{H}_1\mathbf{1}_n = \mathbf{0}$. Hence $\mathbf{H}_2\mathbf{1}_n = \mathbf{OH}_1\mathbf{1}_n = \mathbf{0}$. Hence $\mathbf{H}_2 \in \mathscr{H}$.

Assuming $\mathbf{H}_2 \in \mathscr{H}$, we have by the result in part (a) that $\mathbf{H}_2'\mathbf{H}_2 = \mathbf{H}_1'\mathbf{H}_1 = [\, \mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n']$. Premultiplying by $\mathbf{H}_2$ we get

$$\mathbf{H}_2\mathbf{H}_2'\mathbf{H}_2 = \mathbf{H}_2\mathbf{H}_1'\mathbf{H}_1, \text{ or } \mathbf{H}_2 = \mathbf{H}_2\mathbf{H}_1'\mathbf{H}_1.$$

The matrix $\mathbf{H}_2\mathbf{H}_1'$ is an orthogonal matrix of order (n-1) as

$$[\mathbf{H}_2\mathbf{H}_1']'[\mathbf{H}_2\mathbf{H}_1'] = \mathbf{H}_1\mathbf{H}_2'\mathbf{H}_2\mathbf{H}_1' = \mathbf{H}_1[\, \mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n']\mathbf{H}_1' = \mathbf{I}_{(n-1)},$$

and therefore the transpose of $[\mathbf{H}_2\mathbf{H}_1']$ is equal to its inverse, giving the orthogonal matrix $\mathbf{O} = \mathbf{H}_2\mathbf{H}_1'$ through which $\mathbf{H}_2$ and $\mathbf{H}_1$ are related as $\mathbf{H}_2 = \mathbf{OH}_1$. This completes the proof.

□

As noted earlier, subclasses of mixtures arise on constraining the domain of $\mathbf{G}$ to include specified subsets of S_n^+. We will now use properties of matrices $\mathbf{H}$ in $\mathscr{H}$ to define and explore properties of subsets $\Xi(\mathbf{A}, \mathbf{H})$ of S_n^+.

3.3 MATRICES IN THE CLASS $\Xi(\mathbf{A},\mathbf{H})$

Let $\mathscr{H}$ be as before and let $S^+_{(n-1)}$ denote the set of positive definite matrices of order $(n-1)$. Then membership in a subset denoted as $\Xi(\mathbf{A}, \mathbf{H})$ of S^+_n is specified through matrices $(\mathbf{A}, \mathbf{H})$ chosen from the ordered set $\{S^+_{(n-1)}, \mathscr{H}\}$ as in the following.

Definition 3.2: A $(n\times n)$ matrix Σ belongs to the class $\Xi(\mathbf{A}, \mathbf{H})$, if and only if

(a) $\Sigma \in S^+_n$ and

(b) For every $\Sigma \in \Xi(\mathbf{A}, \mathbf{H})$ the identity $\mathbf{H}\Sigma\mathbf{H}' = \mathbf{A}$ holds for the specified $\mathbf{H} \in \mathscr{H}$.

To see the relevance of this class recall that elliptically symmetric random vectors have many properties similar to those of multivariate normal vectors, including the structure of marginal and conditional distributions, and closure under linear transformations. In particular, for each random vector $\mathbf{X}$ with a distribution $L(\mathbf{X}) = E_n(\mu\mathbf{1}_n, \Sigma, \phi)$, the distribution of the (n-1) dimensional vector $\mathbf{D} = \mathbf{HX}$ is $L(\mathbf{D}) = E_{n-1}(\mathbf{0}, \mathbf{H}\Sigma\mathbf{H}', \phi)$ through the definition $\mathbf{H}\Sigma\mathbf{H}' = \mathbf{A}$ for Σ in $\Xi(\mathbf{A}, \mathbf{H})$. This results in the invariance of the distribution of $\mathbf{D}$ to varying the dispersion matrix Σ within each class $\Xi(\mathbf{A}, \mathbf{H})$. Thus the density $g_\phi(.)$ of $\mathbf{D}$ featured in expression (3.1.1) for the Pitman estimate is invariant, yielding a simple form for the Pitman estimate in the context of mixture distributions. This connects the classes of matrices of Definition 3.2 with results to be developed in Chapter 4. Examples of matrices in a specified class $\Xi(\mathbf{A}, \mathbf{H})$ follow.

Example 3.2(a): Matrices belonging to the class $\Xi(\mathbf{A}_a, \mathbf{H}_a)$.

Let $\mathbf{A}_a = \begin{bmatrix} 11/6 & 5/\sqrt{12} \\ 5/\sqrt{12} & 11/2 \end{bmatrix}$ and $\mathbf{H}_a = \begin{bmatrix} -1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix}$. Then $\Sigma_1 = \begin{bmatrix} 5 & 3 & 1 \\ 3 & 4 & 2 \\ 1 & 2 & 8 \end{bmatrix}$ is positive definite as $|\Sigma_1| = 76 > 0$ and all principal minors of lower order are positive. Further note that

$$\mathbf{H}_a\Sigma_1\mathbf{H}_a' = \begin{bmatrix} -1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 5 & 3 & 1 \\ 3 & 4 & 2 \\ 1 & 2 & 8 \end{bmatrix} \begin{bmatrix} -1/\sqrt{6} & 1/\sqrt{2} \\ 2/\sqrt{6} & 0 \\ -1/\sqrt{6} & -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 11/6 & 5/\sqrt{12} \\ 5/\sqrt{12} & 11/2 \end{bmatrix} = \mathbf{A}_a.$$

Hence the matrix Σ_1 meets the conditions in Definition 3.3. Similar calculations show that the matrix $\Sigma_2 = \begin{bmatrix} 7 & 6 & 5 \\ 6 & 8 & 7 \\ 5 & 7 & 14 \end{bmatrix}$ belongs to the same class.

The theorem that follows characterizes dispersion matrices in each class $\Xi(\mathbf{A}, \mathbf{H}) \subset S_n^+$. In part (a) the matrices in each class are represented as a function of matrices in $\{S_{(n-1)}^+, \mathscr{H}\}$ defining the class, and of an n-dimensional vector γ of constants. From parts (b) to (e) we obtain the admissible set of vectors γ which yield positive definiteness for matrices in $\Xi(\mathbf{A}, \mathbf{H})$. From this theorem any mixing distribution over $\Xi(\mathbf{A}, \mathbf{H})$ is equivalent to a mixing distribution over the corresponding admissible set of vectors γ. The proof to part (a) essentially reproduces a proof due to Jensen (1989b). Proofs to (b) are in Jensen (1992). The proposition in part (e) follows immediately from results in Chaganty (1992). Conclusions identical to those in (e) except for minor parametric differences are drawn in Jensen (1992) in the context of the subset $\Gamma(n)$ of S_n^+ which will be defined later in this section.

Theorem 3.3.1: (a) Let the matrices $\mathbf{H}$ and $\mathbf{A}$ be as defined earlier. A matrix Σ meets the second condition in Definition 3.2 if and only if it can be expressed through a vector γ of constants as

$$\Sigma = \mathbf{H'AH} + \mathbf{M}(\gamma)$$

where $\mathbf{M}(\gamma) = \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n'$, and $\bar{\gamma}$ is the average of the elements of the vector γ.

(b) Let $\eta_1 = \sum_{i=1}^{n}\gamma_i$ and $\eta_2 = \sum_{i=1}^{n}(\gamma_i - \bar{\gamma})^2$. Matrices of the form $\mathbf{M}(\gamma)$ in (a) are indefinite matrices of rank 2 unless $\gamma = \bar{\gamma}\,\mathbf{1}_n$. The positive and negative eigenvalues of $\mathbf{M}(\gamma)$ are

$$\alpha_1 = [\eta_1 + (\eta_1^2 + 4n\eta_2)^{1/2}]/2, \text{ and } \alpha_2 = [\eta_1 - (\eta_1^2 + 4n\eta_2)^{1/2}]/2$$

respectively.

(c) The corresponding eigenvectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are given by

$$[e_{1j}] = \frac{\sqrt{n}(\gamma_j - \bar{\gamma})}{\sqrt{(n\eta_2 + \alpha_1^2)}} + \frac{\alpha_1}{\sqrt{(n^2\eta_2 + n\alpha_1^2)}} \quad \text{and} \quad [e_{2j}] = \frac{\sqrt{n}(\gamma_j - \bar{\gamma})}{\sqrt{(n\eta_2 + \alpha_2^2)}} + \frac{\alpha_2}{\sqrt{(n^2\eta_2 + n\alpha_2^2)}},$$

for $j = 1,2,..\ n$.

(d) The necessary and sufficient condition for $\mathbf{M}(\gamma)$ to be of rank 1 is that $\gamma = \bar{\gamma}\,\mathbf{1}_n$. The single non-zero eigenvalue of $\mathbf{M}(\gamma)$ is then given as $\alpha_1 = \eta_1$ and the corresponding eigenvector is $\mathbf{e}_1 = \left(1/\sqrt{n}\right)\mathbf{1}_n$.

(e) Let $\mathbf{H'AH} = \sum_{i=1}^{n-1}\xi_i\mathbf{g}_i\mathbf{g}_i'$ be the spectral decomposition of the matrix $\mathbf{H'AH}$, and let Σ take the form in part (a). The necessary and sufficient condition for the matrix Σ to be positive definite is that $\sum_{i=1}^{n-1}\frac{\gamma'\mathbf{g}_i\mathbf{g}_i'\gamma}{\xi_i} < \bar{\gamma}$.

Proof: To prove (a) we start by showing that solutions Σ to $\mathbf{H\Sigma H'} = \mathbf{A}$ take the required form. Note on premultiplying and postmultiplying the identity relation $\mathbf{H\Sigma H'} = \mathbf{A}$ by $\mathbf{H'}$ and $\mathbf{H}$ respectively, that

$$\mathbf{H'H\Sigma H'H} = \mathbf{H'AH}, \text{ or } \mathbf{B\Sigma B} = \mathbf{H'AH},$$

where $\mathbf{B} = [\,\mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n'\,]$ by part (a) of Lemma 3.2.1. Further the identity $\mathbf{B\Sigma B} = \mathbf{H'AH}$ leads to the condition $\mathbf{H\Sigma H'} = \mathbf{A}$ on expressing $\mathbf{B}$ as $\mathbf{H'H}$ and premultiplying $\mathbf{B\Sigma B} = \mathbf{H'AH}$ by $\mathbf{H}$ and $\mathbf{H'}$ respectively as in the following

$$\mathbf{HH'H\Sigma H'HH' = HH'AHH'}, \text{ or } \mathbf{H\Sigma H' = A}.$$

Thus the condition $\mathbf{H\Sigma H' = A}$ is equivalent to requiring that $\mathbf{B\Sigma B = H'AH}$. Solutions for Σ meeting this condition are obtained by applying Theorem 2.3.2 of Rao and Mitra (1971). They give solutions $\mathbf{X}$ to the equation $\mathbf{JXE = C}$ as

$$\mathbf{X = J^{-}CE^{-} + Z - J^{-}JZEE^{-}}$$

where $\mathbf{Z}$ is an arbitrary matrix and $\mathbf{E}^{-}$ is any generalized inverse of $\mathbf{E}$. The necessary and sufficient condition for a solution to exist is that $\mathbf{JJ^{-}CE^{-}E = C}$. In our case, using $\mathbf{I}_n$ as a g-inverse of $\mathbf{B}$ we have

$$\mathbf{BB^{-}H'AHB^{-}B = BH'AHB = H'HH'AHH'H = H'AH},$$

and the required condition holds. Hence solutions for Σ are given as

$$\Sigma = \mathbf{B^{-}H'AHB^{-} + S - BSB},$$

where symmetry for Σ requires that $\mathbf{S}$ be symmetric. Substituting $\mathbf{B} = [\,\mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n'\,]$ in the expression and using $\mathbf{I}_n$ as a g-inverse for $\mathbf{B}$ we get

$$\begin{aligned}\Sigma &= \mathbf{H'AH} + \mathbf{S} - [\,\mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n'\,]\mathbf{S}[\,\mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n'\,] \\ &= \mathbf{H'AH} + (1/n)\mathbf{S1}_n\mathbf{1}_n' + (1/n)\mathbf{1}_n\mathbf{1}_n'\mathbf{S} - (n^{-2})\mathbf{1}_n\mathbf{1}_n'\mathbf{S1}_n\mathbf{1}_n'.\end{aligned}$$

Using the vector of column means of $\mathbf{S}$, $\gamma = (1/n)\mathbf{S1}_n$ as the parameters for each solution, we obtain

$$\Sigma = \mathbf{H'AH} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\mathbf{1}_n\mathbf{1}_n' = \mathbf{H'AH} + \mathbf{M}(\gamma),$$

as required. To show the converse we assume that a matrix Σ is given by this expression for some vector γ of constants. Since $\mathbf{H} \in \mathscr{H}$, $\mathbf{H'H} = \mathbf{I}_{(n-1)}$ and $\mathbf{H}_1\mathbf{1}_n = \mathbf{0}$, we find that

$$\mathbf{H\Sigma H'} = \mathbf{H}[\mathbf{H'AH} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\mathbf{1}_n\mathbf{1}_n']\mathbf{H'} = \mathbf{A},$$

as in Definition 3.2 (b). This establishes the equivalence of the stated expression for matrices Σ and the condition $\mathbf{H\Sigma H' = A}$ in Definition 3.2 (b).

The eigenvectors of $\mathbf{M}(\gamma)$ in part (c) are obtained on solving for $\mathbf{e}_i$ in $\mathbf{M}(\gamma)\mathbf{e}_i = \alpha_i\mathbf{e}_i$. This leads to

$$(\gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n')\mathbf{e}_i = \gamma\mathbf{1}_n'\mathbf{e}_i + \mathbf{1}_n\gamma'\mathbf{e}_i - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n'\mathbf{e}_i = \alpha_i\mathbf{e}_i$$

$$\Rightarrow [e_{ij}] = (\alpha_i)^{-1}[\gamma_j\sum_{j=1}^{n} e_{ij} + \sum_{j=1}^{n}\gamma_j e_{ij} + \bar{\gamma}\sum_{j=1}^{n} e_{ij}] \text{ for } j = 1, \ldots, n. \tag{3.3.1}$$

$$\Rightarrow \sum_{j=1}^{n} e_{ij} = (\alpha_i)^{-1}[n\bar{\gamma}\sum_{j=1}^{n} e_{ij} + n\sum_{j=1}^{n}\gamma_j e_{ij} - n\bar{\gamma}\sum_{j=1}^{n} e_{ij}] = (\alpha_i)^{-1}n\sum_{j=1}^{n}\gamma_j e_{ij}.$$

Using $\sum_{j=1}^{n}\gamma_j e_{ij} = n^{-1}\alpha_i\sum_{j=1}^{n} e_{ij}$ in (3.3.1) yields

$$[e_{ij}] = (\alpha_i)^{-1}[\gamma_j\sum_{j=1}^{n} e_{ij} + n^{-1}\alpha_i\sum_{j=1}^{n} e_{ij} + \bar{\gamma}\sum_{j=1}^{n} e_{ij}]$$

$$= (\alpha_i)^{-1}(\gamma_j - \bar{\gamma})\sum_{j=1}^{n} e_{ij} + n^{-1}\sum_{j=1}^{n} e_{ij} \text{ for } j = 1, \ldots n. \tag{3.3.2}$$

Requiring in addition that $\sum_{j=1}^{n} e_{ij}^2 = 1$, we obtain the following using expression (3.3.2).

$$\sum_{j=1}^{n} e_{ij}^2 = 1 = \sum_{j=1}^{n}\left[(\alpha_i)^{-1}(\gamma_j - \bar{\gamma})\sum_{j=1}^{n} e_{ij} + n^{-1}\sum_{j=1}^{n} e_{ij}\right]^2$$

$$\Rightarrow 1 = \left[\sum_{j=1}^{n} e_{ij}\right]^2 \sum_{j=1}^{n}\left[(\alpha_i)^{-2}(\gamma_j - \bar{\gamma})^2 + 2n^{-1}(\alpha_i)^{-1}(\gamma_j - \bar{\gamma}) + n^{-2}\right]^2$$

$$\Rightarrow 1 = \left[\sum_{j=1}^{n} e_{ij}\right]^2 \left[\sum_{j=1}^{n}(\alpha_i)^{-2}(\gamma_j - \bar{\gamma})^2 + n^{-1}\right]$$

$$\Rightarrow \sum_{j=1}^{n} e_{ij} = \sqrt{\frac{\alpha_i^2 n}{n\eta_2 + \alpha_i^2}} \quad \text{for } \eta_2 = \sum_{i=1}^{n}(\gamma_i - \bar{\gamma})^2.$$

Substituting this in (3.3.2), we get

$$[e_{ij}] = \frac{\sqrt{n}(\gamma_i - \bar{\gamma})}{\sqrt{(n\eta_2 + \alpha_i^2)}} + \frac{\alpha_i}{\sqrt{(n^2\eta_2 + n\alpha_i^2)}},$$

as required.

The necessary and sufficient condition required in part (d) follows on noting as in Jensen (1992) that the matrix $\mathbf{M}(\gamma)$ takes the form

$$\mathbf{M}(\gamma) = \mathbf{1}_n\theta' + \theta\mathbf{1}_n' \text{ with } \theta = (\gamma - (1/2)\bar{\gamma}\,\mathbf{1}_n).$$

Thus $\mathbf{M}(\gamma)$ is of rank 1 if and only if $\theta = (\gamma - (1/2)\bar{\gamma}\,\mathbf{1}_n) = d\mathbf{1}_n$ for some scalar d, which is true if and only if $d = \gamma_i - (1/2)\bar{\gamma}$ for each i. Thus the condition $\gamma = \bar{\gamma}\,\mathbf{1}_n$ holds as required. The eigenvalue and eigenvector are obtained on imposing this condition to parts (b) and (c). Note that the second eigenvalue

$$\alpha_2 = [\eta_1 - (\eta_1^2 + 4n\eta_2)^{1/2}]/2 = 0, \text{ as } \eta_2 = \sum_{i=1}^{n}(\gamma_i - \bar{\gamma})^2 = 0.$$

The first eigenvalue

$$\alpha_1 = [\eta_1 + (\eta_1^2 + 4n\eta_2)^{1/2}]/2 = \eta_1$$

and the eigenvector corresponding to this eigenvalue is given as

$$[e_{1j}] = \frac{\sqrt{n}(\gamma_j - \bar{\gamma})}{\sqrt{(n\eta_2 + \alpha_1^2)}} + \frac{\alpha_1}{\sqrt{(n^2\eta_2 + n\alpha_1^2)}} = 0 + \frac{\eta_1}{\sqrt{n\eta_1^2}}\mathbf{1}_n, \text{ or } \mathbf{e}_1 = \left(1/\sqrt{n}\right)\mathbf{1}_n.$$

□

The preceding theorem helps in generating positive definite matrices meeting the condition of Definition 3.2. For a given choice of (**A**, **H**) in $\{S^+_{(n-1)}, \mathscr{H}\}$, the result in part (e) yields an admissible vector γ. This can then be used to compute a matrix in Ξ(**A**, **H**) by using the result in part (a). This shows that matrices in Ξ(**A**, **H**) are obtained by adding two matrices. The first is the matrix **H'AH** of rank (n-1) with the unit vector spanning its null space, and the second is the singular matrix **M**(γ) with eigenspace associated with non-zero eigenvalues in the plane containing the vector γ and the unit vector. From part (d) we see that the eigenspace of **M**(γ) collapses onto the equiangular line when $\gamma = \bar{\gamma}\mathbf{1}_n$. Part (e) deals with constraints on γ for positive definiteness of matrices obtained as Σ = **H'AH**+ **M**(γ). Figure 3.1 illustrates the constraint in part (e) for n = 3. The admissible vectors γ are inside a convex set about the equiangular line with elliptical sections perpendicular to the equiangular line. As noted, the proof to this result is in Chaganty (1992). We support that conclusion through the following comments. Positive definiteness for Σ requires

$$\mathbf{a}'\Sigma\mathbf{a} = \mathbf{a}'\mathbf{H}'\mathbf{A}\mathbf{H}\mathbf{a} + \mathbf{a}'\mathrm{M}(\gamma)\mathbf{a} > 0,$$

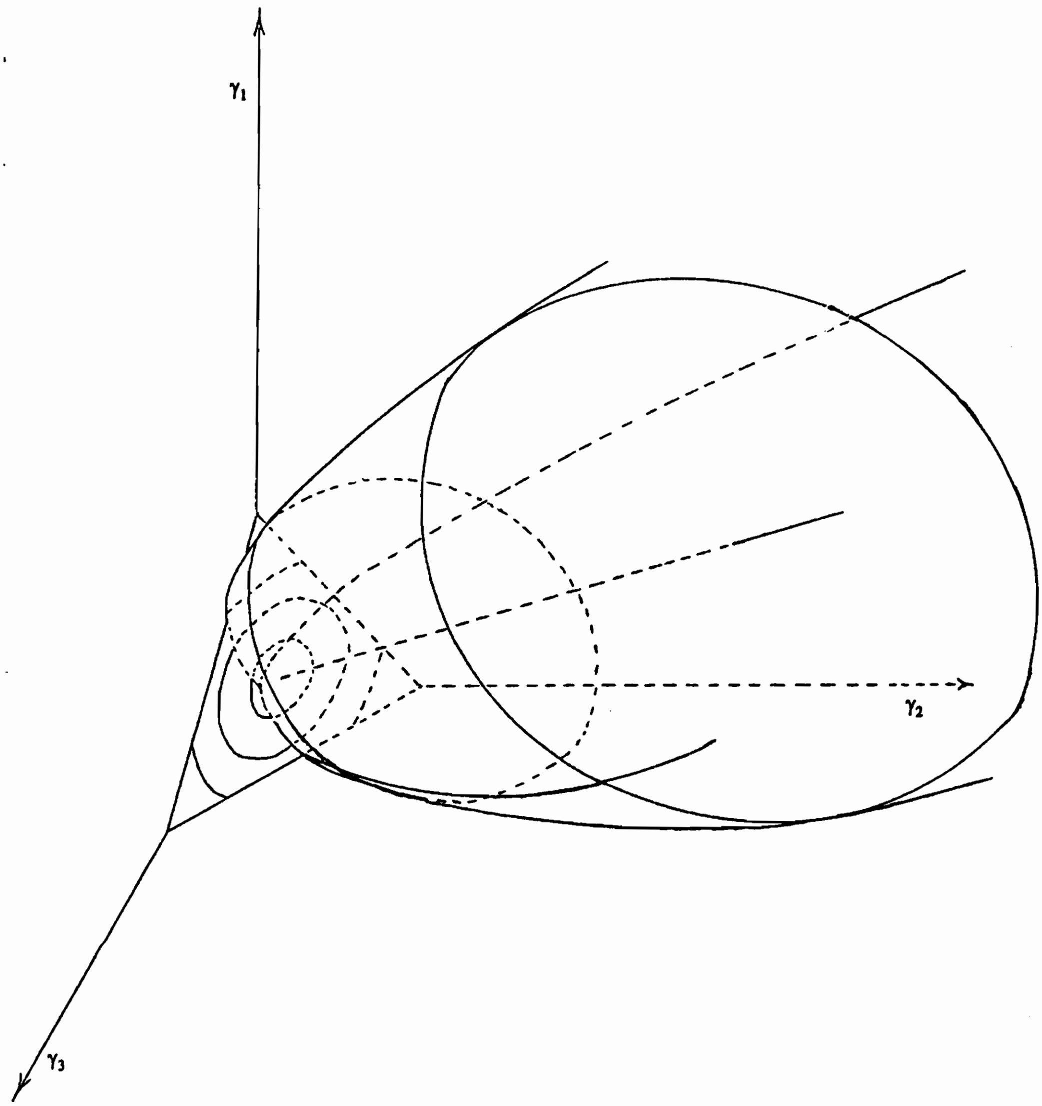

Figure 3.1: Admissible vectors γ for positive definiteness of $\Sigma = \mathbf{H'AH} + \mathbf{M}(\gamma)$ in S_3^+. The set of admissible vectors is a convex set with elliptical sections as shown. Level sets for the quadratic form $\mathbf{a'H'AHa}$ for $\mathbf{a} \in \Re^3$ are depicted on the triangular plane perpendicular to the equiangular line. This figure illustrates the relationship between these level sets and the elliptical sections of the convex set of admissible vectors.

for all $\mathbf{a} \in \Re^n$. This expression suggests constraints on γ depending on the spectral properties of $\mathbf{H'AH}$. Level sets for the function $\mathbf{a'H'AHa}$ and the corresponding constraints on γ for n=3 are depicted in Figure 3.1. For instance consider vectors $\mathbf{a} \in \Re^3$ spanning the direction perpendicular to and away from the equiangular line for which the function $\mathbf{a'H'AHa}$ has its maximum gradient. This direction corresponds to the minor axis of the elliptical level sets for $\mathbf{a'H'AHa}$. The constraints on γ in this direction are less restrictive as the major axis of the elliptical sections of the admissible set lie in this direction. Distributions over this set of admissible vectors generate mixing distributions over $\Xi(\mathbf{A}, \mathbf{H})$ for any $(\mathbf{A}, \mathbf{H})$ in $\{S^+_{(n-1)}, \mathscr{H}\}$. Later results for mixtures demonstrate the preservation of the sample average as the Pitman estimator when the expected vector $E(\gamma)$ under mixing lies on the equiangular line.

We observe from results in Theorem 3.3.1 that for any choice $(\mathbf{A}, \mathbf{H})$ in $\{S^+_{(n-1)}, \mathscr{H}\}$, the representation of $\Xi(\mathbf{A}, \mathbf{H})$ in (a), as well as the constraints on the admissible vectors γ in (e), depend solely on the function $T(\mathbf{A}, \mathbf{H}) = \mathbf{H'AH}$. This implies that if other ordered choices from $\{S^+_{(n-1)}, \mathscr{H}\}$ yield the same matrix as $T(.,.)$ they characterize the same subset of positive definite matrices. In Example 3.2(a) two positive definite matrices of order 3 were verified as belonging to $\Xi(\mathbf{A}, \mathbf{H})$ for matrices in $\{S^+_2, \mathscr{H}\}$ as specified. In Example 3.2(b) we demonstrate that the specification of the subset of S^+_3 to which both matrices belong is not unique, but uniqueness is clarified through results in Lemma 3.3.1.

Example 3.2(b): In Example 3.2(a) we had shown that the positive definite matrices $\Sigma_1 = \begin{bmatrix} 5 & 3 & 1 \\ 3 & 4 & 2 \\ 1 & 2 & 8 \end{bmatrix}$ and $\Sigma_2 = \begin{bmatrix} 7 & 6 & 5 \\ 6 & 8 & 7 \\ 5 & 7 & 14 \end{bmatrix}$ belonged to the class $\Xi(\mathbf{A}_a;\mathbf{H}_a)$ for $\mathbf{A}_a = \begin{bmatrix} 11/6 & 5/\sqrt{12} \\ 5/\sqrt{12} & 11/2 \end{bmatrix}$ and $\mathbf{H}_a = \begin{bmatrix} -1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix}$. In this example we note that the class to which both

Σ_1 and Σ_2 belong may be specified by using other matrices $\mathbf{H}_b \in \mathscr{H}$ and $\mathbf{A}_b \in S^+_{(n-1)}$, as given by $\mathbf{H}_b = \begin{bmatrix} 2/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$ and $\mathbf{A}_b = \begin{bmatrix} 20/6 & 8/\sqrt{12} \\ 8/\sqrt{12} & 4 \end{bmatrix}$. This follows on noting that

$$\mathbf{H}_b\Sigma_1\mathbf{H}_b' = \begin{bmatrix} 2/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}\begin{bmatrix} 5 & 3 & 1 \\ 3 & 4 & 2 \\ 1 & 2 & 8 \end{bmatrix}\begin{bmatrix} 2/\sqrt{6} & 0 \\ -1/\sqrt{6} & 1/\sqrt{2} \\ -1/\sqrt{6} & -1/\sqrt{2} \end{bmatrix}$$

$$= \begin{bmatrix} 20/6 & 8/\sqrt{12} \\ 8/\sqrt{12} & 4 \end{bmatrix} = \mathbf{A}_b,$$

and that

$$\mathbf{H}_b\Sigma_2\mathbf{H}_b' = \begin{bmatrix} 2/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}\begin{bmatrix} 7 & 6 & 5 \\ 6 & 8 & 7 \\ 5 & 7 & 14 \end{bmatrix}\begin{bmatrix} 2/\sqrt{6} & 0 \\ -1/\sqrt{6} & 1/\sqrt{2} \\ -1/\sqrt{6} & -1/\sqrt{2} \end{bmatrix}$$

$$= \begin{bmatrix} 20/6 & 8/\sqrt{12} \\ 8/\sqrt{12} & 4 \end{bmatrix} = \mathbf{A}_b.$$

The seeming ambiguity in specifying a subset of S^+_n through Definition 3.2, as illustrated in the preceding example, is clarified in the lemma that follows. For orthogonal matrices **O** of order $(n-1)$, we show in part (a) of the theorem that ordered pairs chosen from $\{S^+_{(n-1)}, \mathscr{H}\}$ specify identical subsets of S^+_n through Definition 3.2, if and only if they lie on the same orbit under the transformation group $g(\mathbf{A}, \mathbf{H}) = (\mathbf{OAO}', \mathbf{OH})$ applied to the elements of $\{S^+_{(n-1)}, \mathscr{H}\}$. In parts (b) and (c) we obtain maximal invariant functions under this group of transformations.

Lemma 3.3.1: (a) Let $\mathscr{O}(n-1)$ denote the set of orthogonal matrices of order $(n-1)$, and let $(\mathbf{A}_1, \mathbf{H}_1)$ and $(\mathbf{A}_2, \mathbf{H}_2)$ be two distinct choices of matrices from $\{S^+_{(n-1)}, \mathscr{H}\}$. Then the necessary and sufficient condition for the subset $\Xi(\mathbf{A}_1, \mathbf{H}_1)$ of S^+_n to coincide with the subset $\Xi(\mathbf{A}_2, \mathbf{H}_2)$ is that $\mathbf{H}_2 = \mathbf{OH}_1$ and $\mathbf{A}_2 = \mathbf{OA}_1\mathbf{O}'$ for some $\mathbf{O} \in \mathscr{O}(n-1)$.

(b) For $(\mathbf{A}, \mathbf{H}) \in \{S^+_{(n-1)}, \mathscr{H}\}$ consider the group of transformations

$$g(\mathbf{A}, \mathbf{H}) = (\mathbf{OAO}', \mathbf{OH}) \text{ for } \mathbf{O} \in \mathcal{O}(n-1).$$

A maximal invariant function under this group is

$$\mathbf{T}(\mathbf{A}, \mathbf{H}) = \mathbf{H'AH}.$$

(c) For the group of transformations in (b) another maximal invariant is given by

$$\mathbf{U}(\mathbf{A}, \mathbf{H}) = \mathbf{H'A^{-1}H}.$$

Proof: To prove (a) we note by part (c) of Lemma 3.2.1 that if matrices $\mathbf{H}_2$ and $\mathbf{H}_1$ are members of $\mathcal{H}$ then $\mathbf{H}_2 = \mathbf{OH}_1$ for some $\mathbf{O} \in \mathcal{O}(n-1)$. We now need to show that every Σ chosen from $\Xi(\mathbf{A}_1, \mathbf{H}_1)$ belongs to $\Xi(\mathbf{A}_2, \mathbf{H}_2)$, if and only if $\mathbf{A}_2 = \mathbf{OA}_1\mathbf{O}'$ for this $\mathbf{O} \in \mathcal{O}(n-1)$. On assuming that $\Sigma \in \Xi(\mathbf{A}_2, \mathbf{H}_2)$, we have by Definition 3.2 that $\mathbf{H}_2\Sigma\mathbf{H}_2' = \mathbf{A}_2$. Substituting $\mathbf{H}_2 = \mathbf{OH}_1$ in this expression and using $\mathbf{H}_1\Sigma\mathbf{H}_1' = \mathbf{A}_1$, we obtain

$$\mathbf{H}_2\Sigma\mathbf{H}_2' = \mathbf{OH}_1\Sigma\mathbf{H}_1'\mathbf{O}' = \mathbf{OA}_1\mathbf{O}' = \mathbf{A}_2.$$

Further on assuming $\mathbf{A}_2 = \mathbf{OA}_1\mathbf{O}'$ for a matrix $\mathbf{O} \in \mathcal{O}(n-1)$ such that $\mathbf{H}_2 = \mathbf{OH}_1$, we have that

$$\mathbf{H}_2\Sigma\mathbf{H}_2' = \mathbf{OH}_1\Sigma\mathbf{H}_1'\mathbf{O}' = \mathbf{OA}_1\mathbf{O}' = \mathbf{A}_2,$$

implying that $\Sigma \in \Xi(\mathbf{A}_2, \mathbf{H}_2)$.

To prove that the function in (b) is maximal invariant we need to show that

$$\mathbf{T}[\mathbf{A}, \mathbf{H}] = \mathbf{T}[g(\mathbf{A}, \mathbf{H})] \text{ for all } (\mathbf{A}, \mathbf{H}) \in \{S^{+}_{(n-1)}, \mathcal{H}\},$$

and that the condition $\mathbf{T}(\mathbf{A}_1, \mathbf{H}_1) = \mathbf{T}(\mathbf{A}_2, \mathbf{H}_2)$ implies that

$$\{\mathbf{A}_2, \mathbf{H}_2\} = \{\mathbf{OA}_1\mathbf{O}', \mathbf{OH}_1\} \text{ for some } \mathbf{O} \in \mathcal{O}(n-1).$$

First note that for $\mathbf{T}[\mathbf{A}, \mathbf{H}] = \mathbf{H'AH}$ we have

$$\mathbf{T}[g(\mathbf{A}, \mathbf{H})] = \mathbf{T}[\mathbf{OAO}', \mathbf{OH}] = \mathbf{H'O'OAO'OH} = \mathbf{H'AH} = \mathbf{T}[\mathbf{A}, \mathbf{H}],$$

so that $\mathbf{T}[\mathbf{A}, \mathbf{H}]$ is invariant. Further requiring that $\mathbf{T}(\mathbf{A}_1, \mathbf{H}_1) = \mathbf{T}(\mathbf{A}_2, \mathbf{H}_2)$, we have that $\mathbf{H}_1'\mathbf{A}_1\mathbf{H}_1 = \mathbf{H}_2'\mathbf{A}_2\mathbf{H}_2$. Premultiplying and postmultiplying by $\mathbf{H}_2$ and $\mathbf{H}_2'$, respectively, we get

$$\mathbf{H}_2\mathbf{H}_1'\mathbf{A}_1\mathbf{H}_1\mathbf{H}_2' = \mathbf{H}_2\mathbf{H}_2'\mathbf{A}_2\mathbf{H}_2\mathbf{H}_2'.$$

This expression yields $\mathbf{OA_1O'} = \mathbf{A_2}$ for $\mathbf{O} = \mathbf{H_2H_1'}$ as $\mathbf{H_2H_2'} = \mathbf{I}_{(n-1)}$ by definition of matrices in $\mathscr{H}$. As shown earlier in Lemma 3.2.1, $\mathbf{H_2H_1'}$ is an orthogonal matrix of order $(n-1)$. It is noted that for this orthogonal matrix

$$\mathbf{OH_1} = \mathbf{H_2H_1'H_1} = \mathbf{H_2}[\,\mathbf{I_n} - (1/n)\mathbf{1_n1_n'}] = \mathbf{H_2}.$$

Thus $\mathbf{T}(\mathbf{A_1}, \mathbf{H_1}) = \mathbf{T}(\mathbf{A_2}, \mathbf{H_2})$ implies that $\{\ \mathbf{A_2}, \mathbf{H_2}\ \} = \{\ \mathbf{OA_1O'}, \mathbf{OH_1}\ \}$, and therefore the function $\mathbf{T}(.,.)$ is a maximal invariant under the specified group of transformations.

The maximal invariance of $\mathbf{U}(.,.)$ in part (c) follows using arguments similar to those for invariance of $\mathbf{T}(.,.)$. □

The example that follows illustrates the results in parts (a) and (b) of Lemma 3.3.2. We show that the two choices of ordered pairs in $\{S_2^+ , \mathscr{H}\}$ in Examples 3.2(a) and 3.2(b) are related through an orthogonal matrix of order 2. This illustrates part (a) of Lemma 3.3.2. Further, since both these choices lie on the same orbit under the group $g(\mathbf{A}, \mathbf{H}) = (\mathbf{OA_1O'}, \mathbf{OH})$ applied to the elements of $\{S_2^+ , \mathscr{H}\}$, we verify the invariance of the maximal invariant $\mathbf{T}[\mathbf{A},\mathbf{H}] = \mathbf{H'AH}$. This demonstrates the result in part (b) of the lemma.

Example 3.2(c): Relationship between matrices in $\{S_2^+ , \mathscr{H}\}$ in parts (a) and (b) of Example 3.2.

In parts (a) and (b) of this example two alternate choices of matrices $(\mathbf{A_a}, \mathbf{H_a})$ and $(\mathbf{A_b}, \mathbf{H_b})$ were seen to specify the subset to which both $\Sigma_1 = \begin{bmatrix} 5 & 3 & 1 \\ 3 & 4 & 2 \\ 1 & 2 & 8 \end{bmatrix}$ and $\Sigma_2 = \begin{bmatrix} 7 & 6 & 5 \\ 6 & 8 & 7 \\ 5 & 7 & 14 \end{bmatrix}$ belonged. We now show that $(\mathbf{A_a}, \mathbf{H_a})$ are related to $(\mathbf{A_b}, \mathbf{H_b})$ through an orthogonal matrix $\mathbf{O}$ as $\mathbf{H_b} = \mathbf{OH_a}$ and $\mathbf{A_b} = \mathbf{OA_aO'}$. This orthogonal matrix is $\mathbf{O} = \begin{bmatrix} -1/2 & 3/\sqrt{12} \\ 3/\sqrt{12} & 1/2 \end{bmatrix}$, as

$$\mathbf{OH_a} = \begin{bmatrix} -1/2 & 3/\sqrt{12} \\ 3/\sqrt{12} & 1/2 \end{bmatrix} \begin{bmatrix} -1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 2/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} = \mathbf{H_b},$$

$$\text{and } \mathbf{OA_aO'} = \begin{bmatrix} -1/2 & 3/\sqrt{12} \\ 3/\sqrt{12} & 1/2 \end{bmatrix} \begin{bmatrix} 11/6 & 5/\sqrt{12} \\ 5/\sqrt{12} & 11/2 \end{bmatrix} \begin{bmatrix} -1/2 & 3/\sqrt{12} \\ 3/\sqrt{12} & 1/2 \end{bmatrix}$$

$$= \begin{bmatrix} 20/6 & 8/\sqrt{12} \\ 8/\sqrt{12} & 4 \end{bmatrix} = \mathbf{A_b}.$$

We verify the result in part (b) of Theorem 3.3.2 by noting that

$$\mathbf{H_a'A_aH_a} = \begin{bmatrix} -1/\sqrt{6} & 1/\sqrt{2} \\ 2/\sqrt{6} & 0 \\ -1/\sqrt{6} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 20/6 & 8/\sqrt{12} \\ 8/\sqrt{12} & 4 \end{bmatrix} \begin{bmatrix} -1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix}$$

$$= (1/9) \begin{bmatrix} 20 & 2 & -22 \\ 2 & 11 & -13 \\ -22 & -13 & 35 \end{bmatrix}$$

and that this is the same as

$$\mathbf{H_b'A_bH_b} = \begin{bmatrix} 2/\sqrt{6} & 0 \\ -1/\sqrt{6} & 1/\sqrt{2} \\ -1/\sqrt{6} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 20/6 & 8/\sqrt{12} \\ 8/\sqrt{12} & 4 \end{bmatrix} \begin{bmatrix} 2/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

$$= (1/9) \begin{bmatrix} 20 & 2 & -22 \\ 2 & 11 & -13 \\ -22 & -13 & 35 \end{bmatrix}.$$

This example and Lemma 3.3.1 suggest direct specification of the classes $\Xi(\mathbf{A}, \mathbf{H})$ in terms of elements in the range of $\mathbf{T}(.,.)$. In part (a) of Lemma 3.3.2 to follow, the function $\mathbf{T}(.,.)$ applied to elements of the ordered set $\{S^{+}_{(n-1)}, \mathscr{H}\}$ is shown to yield a class of matrices denoted as $\mathscr{C}(n)$. Membership in this class requires conditions noted in Definition 3.3 to follow. Part (b) of Lemma 3.3.2 obtains a unique minimum norm least squares g-inverse for matrices in $\mathscr{C}(n)$, to be used in Chapter 4.

Definition 3.3: A matrix $\mathbf{C}$ of order n×n belongs to the class $\mathscr{C}(n)$ if and only if the matrix is positive semidefinite having a one dimensional null space spanned by the unit vector.

Lemma 3.3.2: (a) For $(\mathbf{A}, \mathbf{H})$ in $\{S^{+}_{(n-1)}, \mathscr{H}\}$ let $\mathbf{R(T)}$ denote the range of the function $\mathbf{T}(\mathbf{A}, \mathbf{H}) = \mathbf{H'AH}$. Then $\mathbf{R(T)} = \mathscr{C}(n)$.

(b) Consider any $\mathbf{C} \in \mathscr{C}(n)$, and let $(\mathbf{A}, \mathbf{H})$ be any pair from the ordered set $\{S^{+}_{(n-1)}, \mathscr{H}\}$ for which $\mathbf{H'AH} = \mathbf{C}$. Then the matrix given by $\mathbf{H'A^{-1}H}$ is the minimum norm least squares inverse of $\mathbf{C}$.

Proof: To prove (a) note that whenever $\mathbf{C}$ is a matrix in $\mathbf{R(T)}$, then $\mathbf{C} = \mathbf{H'AH}$ for some $(\mathbf{A}, \mathbf{H})$ from $\{S^{+}_{(n-1)}, \mathscr{H}\}$. Hence $\mathbf{1}_n'\mathbf{C} = \mathbf{1}_n'\mathbf{H'AH} = \mathbf{0}$ as $\mathbf{1}_n'\mathbf{H'} = \mathbf{0}$ by definition. Thus the null space of every matrix $\mathbf{C}$ in $\mathbf{R(T)}$ is spanned by the unit vector. The matrix $\mathbf{C}$ is of rank (n-1) as the matrix $\mathbf{A}$ in the representation of $\mathbf{C}$ as $\mathbf{H'AH}$ is positive definite of rank (n-1). Hence $\mathbf{C} \in \mathbf{R(T)}$ implies that $\mathbf{C} \in \mathscr{C}(n)$. To prove the converse chose any $\mathbf{C} \in \mathscr{C}(n)$. Then the chosen matrix $\mathbf{C}$ has a spectral decomposition as

$$\mathbf{C} = \sum_{i=1}^{n-1} \xi_i \mathbf{g}_i \mathbf{g}_i', \tag{3.3.3}$$

where ξ_i, for i = 1, ... (n-1), are the eigenvalues of $\mathbf{C}$, with $\mathbf{g}_i$ as the corresponding eigenvectors. The vectors $\mathbf{g}_i$ are orthonormal to the unit vector as the unit vector spans the null space of the matrix $\mathbf{C}$. The expression (3.3.3) can be rewritten as

$$\mathbf{C} = \sum_{i=1}^{n-1} \xi_i \mathbf{g}_i \mathbf{g}_i' = \mathbf{H}_1'\mathbf{D}_1\mathbf{H}_1,$$

where $\mathbf{H}_1$ is the (n-1) × n matrix containing the (n-1) eigenvectors $\mathbf{g}_i$ of $\mathbf{C}$ as row vectors and $\mathbf{D}_1$ is the diagonal matrix with the corresponding eigenvalues. Since the $\mathbf{g}_i$ are normalized eigenvectors we have $\mathbf{H}_1\mathbf{H}_1' = \mathbf{I}_{(n-1)}$. Further $\mathbf{H}_1\mathbf{1}_n = \mathbf{0}$ as the $\mathbf{g}_i$ are orthonormal to the unit vector. Hence by Definition 3.1 $\mathbf{H}_1 \in \mathscr{H}$. In addition $\mathbf{D} \in S^{+}_{(n-1)}$ as it is a diagonal matrix containing positive eigenvalues. Thus every matrix from $\mathscr{C}(n)$ has a decomposition in terms of some pair $(\mathbf{A}, \mathbf{H})$ in $\{S^{+}_{(n-1)}, \mathscr{H}\}$ taking the form $\mathbf{C} = \mathbf{H'AH} = \mathbf{T}(\mathbf{A}, \mathbf{H})$. Hence $\mathbf{R(T)} = \mathscr{C}(n)$.

The result in (b) follows on using Theorem 3.3.1 in Rao and Mitra (1971). In that theorem the necessary and sufficient conditions for a matrix **G** to be the minimum norm least squares g-inverse of **A** are

(i) $\mathbf{AGA} = \mathbf{A}$, (ii) $(\mathbf{AG})' = \mathbf{AG}$, (iii) $\mathbf{GAG} = \mathbf{G}$ and, (iv) $(\mathbf{GA})' = \mathbf{GA}$.

These conditions apply to the matrix $\mathbf{H'A^{-1}H}$. Note that

$$\mathbf{H'AHH'A^{-1}HH'AH} = \mathbf{H'AH},$$

as $\mathbf{HH'} = \mathbf{I}_{(n-1)}$. This verifies the first condition. The third condition follows in a similar manner. The second condition holds as the matrix

$$\mathbf{H'AHH'A^{-1}H} = \mathbf{H'H},$$

is also the idempotent matrix $\mathbf{B} = [\ \mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n']$ and hence is symmetric. The fourth condition follows in a similar manner. Thus the matrix $\mathbf{H'A^{-1}H}$ is the minimum norm least squares inverse of **C**. □

In part (b) of this lemma the minimum norm least squares inverse of $\mathbf{C} \in \mathscr{C}(n)$ takes the form $\mathbf{H'A^{-1}H}$ whenever $\mathbf{C} = \mathbf{H'AH}$. It follows from (b) that $\mathbf{H'A^{-1}H}$ is also contained in $\mathscr{C}(n)$. This kind of an inverse for a matrix $\mathbf{C} \in \mathscr{C}(n)$ is thus denoted as $\mathbf{C}^{-1}_{\mathscr{C}(n)}$ in succeeding results. The function $\mathbf{H'A^{-1}H}$ is the second maximal invariant $\mathbf{U}(.,.)$ identified in part (c) of Lemma 3.3.1. Thus Lemma 3.3.2 (b) identifies an invertible relationship which relates the maximal invariants in that result. Part (a) of Lemma 3.3.2 allows the specification of the classes $\Xi(\mathbf{A}, \mathbf{H})$ through matrices $\mathbf{C} \in \mathscr{C}(n)$. Theorem 3.3.2 that follows denotes these classes as $\aleph_{\mathbf{C}}$ and establishes a partition of S_n^+ into the $\aleph_{\mathbf{C}}$ subsets. The notation $\aleph_{\mathbf{C}}$ and $\Xi(\mathbf{A}, \mathbf{H})$ will be used interchangeably for these subsets in the succeeding developments. A partition of the ordered set $\{S^+_{(n-1)}, \mathscr{H}\}$ into equivalence classes in part (a) of the theorem leads to the partition of the set of positive definite matrices S_n^+.

Theorem 3.3.2: (a) For $(\mathbf{X}, \mathbf{Y}) \in \{S^+_{(n-1)}, \mathscr{H}\}$, let $\{\mathscr{A}, \mathscr{G}\}_{(X,Y)}$ be the class of all ordered pairs in $\{S^+_{(n-1)}, \mathscr{H}\}$ such that for every element $(\mathbf{A}, \mathbf{H})$ in $\{\mathscr{A}, \mathscr{G}\}_{(X,Y)}$, $\mathbf{H'AH} = \mathbf{Y'XY}$. The collective family of all such classes obtained on varying $(\mathbf{X}, \mathbf{Y}) \in \{S^+_{(n-1)}, \mathscr{H}\}$ partitions $\{S^+_{(n-1)}, \mathscr{H}\}$.

(b) Let $\mathbf{M}(\gamma) = \gamma \mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n'$ as in Theorem 3.3.1. Then the set of matrices S^+_n can be partitioned as the family of subsets $\mathscr{D} = (\aleph_C | \mathbf{C} \in \mathscr{C}(n))$. For a chosen $\mathbf{C} \in \mathscr{C}(n)$ the subset $\aleph_C$ is given as

$$\aleph_C = [\,\Sigma\!: \Sigma = \mathbf{C} + \mathbf{M}(\gamma), \text{ for } \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\gamma < \bar{\gamma}\,].$$

where $\mathbf{C}^{-1}_{\mathscr{C}(n)}$ denotes the minimum norm least squares inverse of the matrix $\mathbf{C}$.

Proof: To prove (a) let $(\mathbf{A}_a, \mathbf{H}_a)$ and $(\mathbf{A}_b, \mathbf{H}_b)$ be related to each other if $\mathbf{H}_a'\mathbf{A}_a\mathbf{H}_a = \mathbf{H}_b'\mathbf{A}_b\mathbf{H}_b$. This relation is reflexive as every element relates to itself. Further if $(\mathbf{A}_a, \mathbf{H}_a)$ is related to $(\mathbf{A}_b, \mathbf{H}_b)$, and $(\mathbf{A}_b, \mathbf{H}_b)$ is related to $(\mathbf{A}_c, \mathbf{H}_c)$ then $(\mathbf{A}_a, \mathbf{H}_a)$ is related to $(\mathbf{A}_c, \mathbf{H}_c)$ as

$$\mathbf{H}_a'\mathbf{A}_a\mathbf{H}_a = \mathbf{H}_b'\mathbf{A}_b\mathbf{H}_b \text{ and } \mathbf{H}_b'\mathbf{A}_b\mathbf{H}_b = \mathbf{H}_c'\mathbf{A}_c\mathbf{H}_c \text{ implies that } \mathbf{H}_a'\mathbf{A}_a\mathbf{H}_a = \mathbf{H}_c'\mathbf{A}_c\mathbf{H}_c.$$

This shows the transitivity of the relation. Symmetry follows on noting that

$$\mathbf{H}_a'\mathbf{A}_a\mathbf{H}_a = \mathbf{H}_b'\mathbf{A}_b\mathbf{H}_b \text{ implies that } \mathbf{H}_b'\mathbf{A}_b\mathbf{H}_b = \mathbf{H}_a'\mathbf{A}_a\mathbf{H}_a.$$

Thus the relation is an equivalence relation and the class $\{\mathscr{A}, \mathscr{G}\}_{(X,Y)}$ defined through this equivalence relation is an equivalence class. The collective family of all such equivalence classes obtained on varying $\{\mathbf{X}, \mathbf{Y}\} \in \{S^+_{(n-1)}, \mathscr{H}\}$ partitions $\{S^+_{(n-1)}, \mathscr{H}\}$, as the collective family of all equivalence classes partitions an index set (Berge, 1963).

To prove (b) we note that for $\mathbf{T}[\mathbf{A}, \mathbf{H}] = \mathbf{H'AH}$, each equivalence class $\{\mathscr{A}, \mathscr{G}\}_{(X,Y)}$ as in (a) is obtained as the set $\mathbf{T}^{-1}[\mathbf{Y'XY}]$ in $\{S^+_{(n-1)}, \mathscr{H}\}$. Since $\mathbf{T}(.,.)$ is a maximal invariant under the group $g(\mathbf{A}, \mathbf{H}) = (\mathbf{OA}_1\mathbf{O'}, \mathbf{OH})$, each class constituting the partition in (a) is an orbit under this group of transformations. By the result in Lemma 3.3.1 (a), every such orbit is associated with a unique subset of S_n^+ which is specified by any choice of $(\mathbf{A}, \mathbf{H})$ in that orbit. These establish a correspondence between the orbits and the maximal invariant $\mathbf{T}(.,.)$, as well as one between the orbits and the subsets $\aleph_C$. The range $\mathbf{R}(\mathbf{T})$ of this maximal invariant is the class $\mathscr{C}(n)$ through Lemma 3.3.2 (a). This implies a partition of S_n^+ indexed by elements of $\mathscr{C}(n)$. The form of the elements in each subset follows from results in parts (a) and (e) of Theorem 3.3.1. Part (a) of this theorem implies that

$$\Sigma = \mathbf{H'AH} + \mathbf{M}(\gamma) = \mathbf{C} + \mathbf{M}(\gamma).$$

Further, when the matrix $\mathbf{C}$ has spectral decomposition $\mathbf{C} = \sum_{i=1}^{n-1} \xi_i \mathbf{g}_i \mathbf{g}_i'$, then $\mathbf{C}_{\mathscr{C}(n)}^{-1} = \sum_{i=1}^{n-1} \frac{\mathbf{g}_i \mathbf{g}_i'}{\xi_i}$. This simplifies the constraint in part (e) of Theorem 3.3.1 to

$$\gamma' \mathbf{C}_{\mathscr{C}(n)}^{-1} \gamma < \bar{\gamma}.$$

This establishes the form of elements in each partition set and completes the proof. □

The foregoing partition of S_n^+ supports the decomposition of any mixing probability measure $\mathbf{G}[\cdot;\mathbf{N}]$ over measurable subsets of S_n^+ into conditional probability measures $\mathbf{G}[\cdot|\mathbf{C}]$ for each $\mathbf{C}$ in $\mathscr{C}(n)$. The decomposition takes the form

$$\mathbf{G}[\cdot;\mathbf{N}] = \int_{\mathscr{C}(n)} \mathrm{G}[\cdot|\mathbf{C}] d\mathrm{N}(\mathbf{C}), \tag{3.3.4}$$

where $\mathrm{N}(\mathbf{C})$ is a probability measure over $\mathscr{C}(n)$. In introductory comments we had discussed Pitman estimation for mixture distributions obtained when the domain of $\mathbf{G}$ was restricted to a particular subset $\aleph_C$ for some $\mathbf{C}$ in $\mathscr{C}(n)$. The representation in (3.3.4) allows us to examine less restrictive mixing distributions than those discussed so far. Mixing distributions of interest here would be those for which restrictions would be applied to each probability measure $\mathbf{G}[\cdot|\mathbf{C}]$ in (3.3.4), while the probability measure $\mathrm{N}(\mathbf{C})$

over $\mathscr{C}(n)$ would be unrestricted. In the example that follows we illustrate the results in parts (a) and (b) of Theorem 3.3.2 through positive definite matrices of order 2.

Example 3.3: Matrices in $\aleph_{C_1}$ and $\aleph_{C_2}$ for $\mathbf{C}_1 = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}$ and $\mathbf{C}_2 = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$.

We note that $\mathbf{C}_1$ and $\mathbf{C}_2$ are in $\mathscr{C}(n)$ as they are of rank n-1 = 1, and as $\mathbf{1}_n'\mathbf{C}_1 = \mathbf{0}$ and $\mathbf{1}_n'\mathbf{C}_2 = \mathbf{0}$. This clarifies Definition 3.3. To verify the partition of $\{S^+_{(n-1)}, \mathscr{H}\}$ as in part (c), we note that for n=2 the ordered set $\{S^+_{(n-1)}, \mathscr{H}\}$ reduces to $\{\Re^+_1, \mathbf{h}\}$ where $\Re^+_1$ is the set of positive reals and $\mathbf{h}$ is the set of two dimensional vectors orthonormal to the unit vector. The two elements in the orbit corresponding to $\mathbf{C}_1$ are the pairs $\{1,(1/\sqrt{2},-1/\sqrt{2})'\}$ and $\{1,(-1/\sqrt{2},1/\sqrt{2})'\}$ from $\{\Re^+_1, \mathbf{h}\}$. This follows on noting that

$$\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} 1 \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} 1 \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix} = \mathbf{C}_1.$$

Similarly the two elements in the orbit in $\{\Re^+_1, \mathbf{h}\}$ corresponding to $\mathbf{C}_2$ are the ordered pairs $\{2,(1/\sqrt{2},-1/\sqrt{2})'\}$ and $\{2,(-1/\sqrt{2},1/\sqrt{2})'\}$. We have noted that $\mathbf{C}_1$ = $1\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}' \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$ and equals $\sum_{i=1}^{n-1} \xi_i \mathbf{g}_i \mathbf{g}_i'$ for $\xi_1 = 1$, and $\mathbf{g}_1' = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$. Hence using the representation in part (b) of Theorem 3.3.2 we obtain matrices in $\aleph_{C_1}$ as $\Sigma = \mathbf{C}_1 + \mathbf{M}(\gamma)$, where γ is a two dimensional vector of constants such that

$$\gamma' \mathbf{C}^{-1}_{\mathscr{C}(n)} \gamma = \sum_{i=1}^{n-1} \frac{\gamma' \mathbf{g}_i \mathbf{g}_i' \gamma}{\xi_i} = \frac{\gamma' \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}' \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \gamma}{1} < \bar{\gamma}.$$

A vector γ meeting this condition is $\gamma_{11}' = \begin{bmatrix} 4 & 6 \end{bmatrix}$ as

$$\bar{\gamma} = 5 > \frac{\begin{bmatrix} 4 & 6 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}' \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 4 & 6 \end{bmatrix}'}{1} = 2.$$

Similarly the vector $\gamma_{12}' = \begin{bmatrix} 8 & 5 \end{bmatrix}$ meets this condition. Using the vectors γ_{11} and γ_{12} we obtain Σ_{11} and Σ_{12} as elements of $\aleph_{C_1}$ as in the following.

$$\Sigma_{11} = \mathbf{C}_1 + \mathbf{M}(\gamma_{11}) = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix} + \begin{bmatrix} 4 & 4 \\ 6 & 6 \end{bmatrix} + \begin{bmatrix} 4 & 6 \\ 4 & 6 \end{bmatrix} - \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix} = \begin{bmatrix} 3.5 & 4.5 \\ 4.5 & 7.5 \end{bmatrix}$$

$$\Sigma_{12} = \mathbf{C}_1 + \mathbf{M}(\gamma_{12}) = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix} + \begin{bmatrix} 8 & 8 \\ 5 & 5 \end{bmatrix} + \begin{bmatrix} 8 & 5 \\ 8 & 5 \end{bmatrix} - \begin{bmatrix} 6.5 & 6.5 \\ 6.5 & 6.5 \end{bmatrix} = \begin{bmatrix} 10 & 6 \\ 6 & 4 \end{bmatrix}.$$

To obtain matrices in $\aleph_{C2}$ we note that $\mathbf{C}_2 = 2\left[1/\sqrt{2} \quad -1/\sqrt{2}\right]'\left[1/\sqrt{2} \quad -1/\sqrt{2}\right] = \sum_{i=1}^{n-1} \xi_i \mathbf{g}_i \mathbf{g}_i'$ for $\xi_1 = 2$, and $\mathbf{g}_1' = \left[1/\sqrt{2} \quad -1/\sqrt{2}\right]$. Admissible vectors γ now need to meet the condition

$$\gamma' \mathbf{C}^{-1}_{\mathscr{C}(n)} \gamma = \sum_{i=1}^{n-1} \frac{\gamma' \mathbf{g}_i \mathbf{g}_i' \gamma}{\xi_i} = \frac{\gamma'\left[1/\sqrt{2} \quad -1/\sqrt{2}\right]'\left[1/\sqrt{2} \quad -1/\sqrt{2}\right]\gamma}{2} < \bar{\gamma}.$$

A vector γ meeting this condition is $\gamma_{21}' = \left[3 \quad 7\right]$ as

$$\bar{\gamma} = 5 > \frac{\left[3 \quad 7\right]\left[1/\sqrt{2} \quad -1/\sqrt{2}\right]'\left[1/\sqrt{2} \quad -1/\sqrt{2}\right]\left[3 \quad 7\right]'}{2} = 4.$$

Similarly the vector $\gamma_{22}' = \left[9 \quad 4\right]$ meets the condition. Using these vectors we obtain two elements Σ_{21} and Σ_{22} as elements of $\aleph_{C2}$ as in the following.

$$\Sigma_{21} = \mathbf{C}_2 + \mathbf{M}(\gamma_{21}) = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + \begin{bmatrix} 3 & 3 \\ 7 & 7 \end{bmatrix} + \begin{bmatrix} 3 & 7 \\ 3 & 7 \end{bmatrix} - \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 4 & 10 \end{bmatrix}$$

$$\Sigma_{22} = \mathbf{C}_2 + \mathbf{M}(\gamma_{22}) = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + \begin{bmatrix} 9 & 9 \\ 4 & 4 \end{bmatrix} + \begin{bmatrix} 9 & 4 \\ 9 & 4 \end{bmatrix} - \begin{bmatrix} 6.5 & 6.5 \\ 6.5 & 6.5 \end{bmatrix} = \begin{bmatrix} 12.5 & 5.5 \\ 5.5 & 2.5 \end{bmatrix}.$$

To illustrate the decomposition of a mixing probability measure $\mathbf{G}$ as in (3.3.4), consider a measure which assigns 1/8 probability to each of Σ_{11} and Σ_{12}, 1/4 probability to Σ_{21} and 1/2 probability to Σ_{22}. This measure can be decomposed into the conditional measure $\mathbf{G}(.|\mathbf{C}_1)$ which assigns 1/2 probability to both Σ_{11} and Σ_{12}, and $\mathbf{G}(.|\mathbf{C}_2)$ which assigns 1/3 probability to Σ_{21} and 2/3 probability to Σ_{22}. Using the probability measure N($\mathbf{C}$) which assigns 1/4 probability to $\mathbf{C}_1$ and 3/4 to $\mathbf{C}_2$ we obtain $\mathbf{G}$.

A special subset amongst subsets $\aleph_C$ is the focus of the next section.

3.4 MATRICES IN THE CLASS $\Gamma(n)$

Amongst subsets $\Xi(\mathbf{A};\mathbf{H})$ of S_n^+ a subset denoted by $\Gamma(n)$ figures prominently in statistics. The subset satisfies $\Gamma(n) = \Xi(\mathbf{I}_{(n-1)};\mathbf{H})$. Matrices in this particular subset have stronger invariance properties than those for any other subset $\Xi(\mathbf{A};\mathbf{H})$. Note that while the definition that follows is similar to Definition 3.2, the property in part (b) of this definition holds for all $\mathbf{H} \in \mathscr{H}$. For elliptically symmetric distributions with dispersion matrices in $\Gamma(n)$, and for mixture distributions obtained through a mixing c.d.f. $\mathbf{G}$ defined over this subset, the Pitman estimate will be seen to emerge as a simpler correction of the sample average. Further it has been shown in the context of repeated measurements that for independent random vectors with multivariate normal distributions, dispersion structures within the class $\Gamma(n)$ preserve the distributions of F-ratios. These were studied independently by Huynh and Feldt (1970), and by Rouanet and Le′pine (1970). The definition of this subset draws on these results.

Definition 3.4: A $(n\times n)$ matrix Σ belongs to the class $\Gamma(n)$, if and only if

(a) $\Sigma \in S_n^+$ and

(b) For every $\Sigma \in \Gamma(n)$ the identity $\mathbf{H}\Sigma\mathbf{H}' = \mathbf{I}_{(n-1)}$ holds for all $\mathbf{H} \in \mathscr{H}$.

This definition implies that applying any set of (n-1) orthonormal contrasts to a normally distributed random vector having dispersion matrix $\Sigma \in \Gamma(n)$, we obtain an $(n-1)$-dimensional spherical normal distribution. In this dissertation it will be seen that mixtures of elliptically symmetric distributions using a c.d.f over $\Gamma(n)$ yield (n-1) dimensional spherical distributions for the contrasts. Properties of matrices meeting the conditions of the definition are in Theorem 3.4.1. The results in this theorem are identical to those in Jensen (1992) except for differences in parametrization of the solutions meeting the conditions of Definition 3.3. The results in (a) and (b) can be seen to follow from those in Theorem 3.3.1 and proofs are omitted.

Theorem 3.4.1: (a) Let $\mathbf{B} = [\ \mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n']$. Then a matrix Σ meets the second condition in Definition 3.3 if and only if it can be expressed as

$$\Sigma = \mathbf{B} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n'.$$

(b) The necessary and sufficient condition for matrices Σ in (a) to be positive definite is that

$$\sum_{i=1}^{n}(\gamma_i - \bar{\gamma})^2 < \bar{\gamma}.$$

(c) Let $\bar{\Gamma}(n)$ denote the subset of $\Gamma(n)$ obtained on restricting admissible vectors γ to those satisfying $\gamma = \bar{\gamma}\,\mathbf{1}_n$. Let $\Im(n)$ denote the equicorrelated class of matrices with elements given as

$$\Sigma(\rho) = [\ (1-\rho)\mathbf{I}_n + \rho\mathbf{1}_n\mathbf{1}_n'], \qquad -(n-1)^{-1} < \rho < 1.$$

Then every element of $\Im(n)$ can be obtained by scalar multiplication from a corresponding element in $\bar{\Gamma}(n)$, and vice-versa.

Proof to (c): By the result in (a) elements of $\Gamma(n)$ take the form

$$\Sigma = \mathbf{B} + \mathbf{M}(\gamma) = \mathbf{I}_n - n^{-1}\mathbf{1}_n\mathbf{1}_n' + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n', \text{ when } \sum_{i=1}^{n}(\gamma_i - \bar{\gamma})^2 < \bar{\gamma}.$$

Using $\gamma = \bar{\gamma}\,\mathbf{1}_n$ yields the elements for $\bar{\Gamma}(n)$ as

$$\Sigma(\bar{\gamma}) = \mathbf{I}_n + (\bar{\gamma} - n^{-1})\mathbf{1}_n\mathbf{1}_n', \text{ when } \bar{\gamma} > 0. \qquad (3.4.1)$$

We can rewrite $\Sigma(\bar{\gamma})$ as

$$\Sigma(\bar{\gamma}) = \frac{(n - n\bar{\gamma} - 1)}{n}[(1-\delta)\mathbf{I}_n + \delta\mathbf{1}_n\mathbf{1}_n'] \text{ with } \delta = \left[\frac{n\bar{\gamma} - 1}{n + n\bar{\gamma} - 1}\right].$$

The expression is a scalar multiple of a matrix in the equicorrelated class. Further the condition $\bar{\gamma} > 0$ implies the restriction $-(n-1)^{-1} < \delta < 1$ as in the equicorrelated class. Similarly, starting with an element in $\Im(n)$ we can rewrite it as

$$\Sigma(\rho) = [\ (1-\rho)\mathbf{I}_n + \rho\mathbf{1}_n\mathbf{1}_n'] = \ (1-\rho)\ [\ \mathbf{I}_n + (f - n^{-1})\mathbf{1}_n\mathbf{1}_n'\ ]$$

where $f = \dfrac{(n-1)\rho+1}{n-n\rho}$ and $-(n-1)^{-1} < \rho < 1$. The expression for the elements of $\Im(n)$ are scale multiples of matrices taking the form (3.4.1). Further the condition $-(n-1)^{-1} < \rho < 1$, implies the condition $f > 0$ as is the case for the parameter $\bar{\gamma}$ in (3.4.1). □

Figure 3.2 illustrates the constraint in part (b) for n = 2. The admissible vectors γ are inside the convex set symmetrical about the equiangular line, unlike $\Xi(\mathbf{A};\mathbf{H})$ where the admissible vectors were in sets elliptical about the equiangular line. Example 3.4 that follows constructs a matrix in the class $\Gamma(n)$ using the results in (a) and (b). The third result in the theorem shows that the dispersion structures of matrices in the equicorrelated class are identical to those of a restricted subset of $\Gamma(n)$ except for a scalar multiple. To the extent that an estimation problem does not depend on the scalar multiple (the value of the scale parameter), the set $\Im(n)$ can be considered a subset of $\Gamma(n)$ to which results applicable to $\Gamma(n)$ specialize.

Example 3.4: $\Sigma \in \Gamma(n)$, with $\gamma = (1/3)[\ 10, 13, 16]$.

We note that $\sum_{i=1}^{3}(\gamma_i - \bar{\gamma})^2 = 2 < \bar{\gamma} = 4.33$. Therefore a positive definite matrix in $\Gamma(n)$ is

$$\Sigma = \Sigma(\gamma) \ = \ \mathbf{B} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n'$$

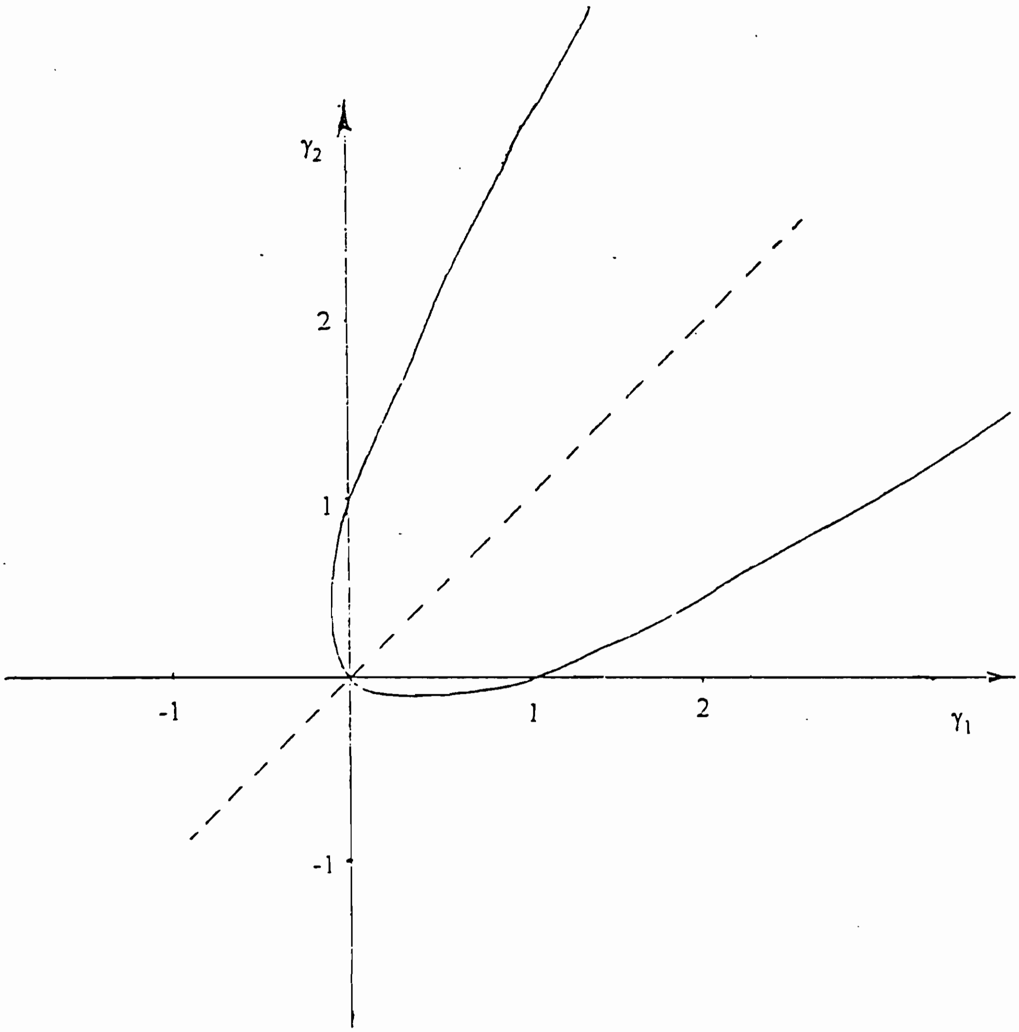

Figure 3.2: Admissible Vectors for positive definiteness of $\Sigma = \Sigma(\gamma)$ for n=2.

$$= (1/3)\left[\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} + \begin{bmatrix} 10 & 10 & 10 \\ 13 & 13 & 13 \\ 16 & 16 & 16 \end{bmatrix} + \begin{bmatrix} 10 & 13 & 16 \\ 10 & 13 & 16 \\ 10 & 13 & 16 \end{bmatrix} - \begin{bmatrix} 13 & 13 & 13 \\ 13 & 13 & 13 \\ 13 & 13 & 13 \end{bmatrix}\right] = \begin{bmatrix} 3 & 3 & 4 \\ 3 & 5 & 5 \\ 4 & 5 & 7 \end{bmatrix}.$$

3.5 MATRICES WITH EQUAL LINEAR FUNCTIONS ON COLUMNS

We will show that for an elliptically symmetric distribution with dispersion matrix Σ the Pitman estimate $\delta_{E,\Sigma}(\mathbf{x})$ for μ is given by the Aitkin (1934) estimator

$$\delta_{E,\Sigma}(\mathbf{x}) = (\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n)^{-1}\mathbf{1}_n'\Sigma^{-1}\mathbf{x}.$$

We note that this is of the form $\mathbf{w'x}$ such that $\sum_{i=1}^{n} w_i = 1$. Let $\mathcal{W}(n)$ denote the set of all n-dimensional vectors $\mathbf{w}$ with $\sum_{i=1}^{n} w_i = 1$. We are now interested in subsets of S_n^+ indexed by $\mathbf{w} \in \mathcal{W}(n)$ such that for every elliptically symmetric distribution with dispersion matrix in such a subset we obtain the same linear function as the Pitman estimate. Such subsets are denoted as $\Lambda(\mathbf{w})$. Matrices in each subset $\Lambda(\mathbf{w})$ are those for which the inner product of $\mathbf{w}'$ with each column vector yields the same scalar. These subsets are defined in the following.

Definition 3.5: For $\mathbf{w} \in \mathcal{W}(n)$, an (n×n) matrix Σ belongs to the class $\Lambda(\mathbf{w})$, if and only if

(a) $\Sigma \in S_n^+$ and

(b) $\mathbf{w}'\Sigma = c\mathbf{1}_n'$.

An example in a subset for a specified member of $\mathcal{W}(n)$ follows.

Example 3.5: Example of a matrix in the equal linear function class.

Consider $\mathbf{a}' = \begin{bmatrix} -1/4 & 3/4 & 1/2 \end{bmatrix}$. Note that $\sum_{i=1}^{3} a_i = 1.0$ and so $\mathbf{a} \in \mathscr{W}(n)$. Consider the matrix $\Sigma = \begin{bmatrix} 12 & 4 & 4 \\ 4 & 4 & 0 \\ 4 & 0 & 6 \end{bmatrix}$. The determinant of this matrix is 128 and the matrix is positive definite as all principal minors are positive. Membership in $\Lambda(\mathbf{a})$ follows on noting that

$$\mathbf{a}'\Sigma = \begin{bmatrix} -1/4 & 3/4 & 1/2 \end{bmatrix} \begin{bmatrix} 12 & 4 & 4 \\ 4 & 4 & 0 \\ 4 & 0 & 6 \end{bmatrix} = 2\mathbf{1}_n'.$$

Theorem 3.5.1 that follows explores connections between the classes $\Lambda(\mathbf{w})$ and the classes $\Xi(\mathbf{A};\mathbf{H})$. In this theorem we denote each subset $\Xi(\mathbf{A};\mathbf{H})$ as $\aleph_C$. Parts (a) and (b) characterize members common to $\Lambda(\mathbf{w})$ and $\aleph_C$ for any pair $(\mathbf{w},\mathbf{C})$ from the ordered set $[\mathscr{W}(n), \mathscr{C}(n)]$. Part (c) characterizes members in each subset $\Lambda(\mathbf{w})$.

Theorem 3.5.1: (a) As in Theorem 3.3.2 part (d) let $\mathscr{C}(n)$ denote the range of the function $\mathbf{T}[\mathbf{A}, \mathbf{H}] = \mathbf{H}'\mathbf{A}\mathbf{H}$ for $(\mathbf{A}, \mathbf{H}) \in \{S^+_{(n-1)}, \mathscr{H}\}$. Further as in part (e) let $\aleph_C$ be a subset of S^+_n generated through a matrix $\mathbf{C} \in \mathscr{C}(n)$ as

$$\aleph_C = [\, \Sigma \colon \Sigma = \mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n', \text{ for } \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\gamma < \bar{\gamma}\,]. \tag{3.5.1}$$

where $\mathbf{C}^{-1}_{\mathscr{C}(n)}$ denotes the minimum norm least squares inverse of the matrix $\mathbf{C}$. Then the necessary and sufficient condition for a matrix in $\aleph_C$ to be contained in a subset $\Lambda(\mathbf{w})$ for some $\mathbf{w} \in \mathscr{W}(n)$ is that $-\mathbf{C}\mathbf{w} = \mathbf{B}\gamma$ for $\mathbf{B} = [\, \mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n']$.

(b) For any $\mathbf{w} \in \mathscr{W}(n)$ and $\mathbf{C} \in \mathscr{C}(n)$ let $\beta(\mathbf{w},\mathbf{C})$ denote the subset of matrices of S^+_n which belong to $\Lambda(\mathbf{w}) \cap \aleph_C$. A matrix Σ is an element of $\beta(\mathbf{w},\mathbf{C})$ if and only if it has a representation as

$$\Sigma = \Sigma(\xi) = \mathbf{C} - \mathbf{C}\mathbf{w}\mathbf{1}_n' - \mathbf{1}_n\mathbf{w}'\mathbf{C}' + \xi\mathbf{1}_n\mathbf{1}_n',$$

in terms of the parameters $\mathbf{w}$ and $\mathbf{C}$ defining the subset and a scalar ξ such that $\xi > \mathbf{w}'\mathbf{C}\mathbf{w}$.

(c) For a vector $\mathbf{w} \in \mathcal{W}(n)$ let $\Lambda(\mathbf{w})$ be the corresponding subset of S_n^+. Then a matrix Σ is an element of $\Lambda(\mathbf{w})$ if and only if it can be expressed as

$$\Sigma = \Sigma(\mathbf{C}, \xi) = \mathbf{C} - \mathbf{Cw1}_n' - \mathbf{1}_n\mathbf{w}'\mathbf{C}' + \xi\mathbf{1}_n\mathbf{1}_n'$$

in terms of the parameter $\mathbf{w}$, a matrix $\mathbf{C}$ in $\mathcal{C}(n)$ and a scalar ξ such that $\xi > \mathbf{w'Cw}$.

<u>Proof</u>: To prove (a) we start by assuming that Σ is in $\Lambda(\mathbf{w})\cap\aleph_C$. Imposing the condition $\mathbf{w}'\Sigma = \mathbf{c1}_n'$ to the representation of matrices in $\aleph_C$ given by (3.5.1) yields

$$\mathbf{w}'[\mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n'] = \mathbf{c1}_n',$$

for some scalar c. On using $\sum_{i=1}^{n} w_i = 1$ this simplifies as

$$\mathbf{w'C} + \mathbf{w}'\gamma\mathbf{1}_n' + \gamma' - \bar{\gamma}\,\mathbf{1}_n' = \mathbf{c1}_n'.$$

Taking a transpose and premultiplying both sides of this equation by the matrix $\mathbf{B}$ yields

$$\mathbf{Cw} + \mathbf{B}\gamma = \mathbf{0} \text{ or } -\mathbf{Cw} = \mathbf{B}\gamma,$$

as $\mathbf{BC} = \mathbf{C}$ and $\mathbf{B1}_n = \mathbf{0}$. To show the converse we assume $-\mathbf{Cw} = \mathbf{B}\gamma$ and that Σ is in the subset $\aleph_C$. Note that $\mathbf{B}\gamma = [\,\mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n']\gamma = \gamma - \bar{\gamma}\,\mathbf{1}_n$. Hence using $-\mathbf{Cw} = \mathbf{B}\gamma$ we have $\gamma = \bar{\gamma}\,\mathbf{1}_n + \mathbf{B}\gamma = \bar{\gamma}\,\mathbf{1}_n - \mathbf{Cw}$. Substituting this in the representation in (3.5.1) yields

$$\Sigma = \mathbf{C} + [\bar{\gamma}\,\mathbf{1}_n - \mathbf{Cw}]\mathbf{1}_n' + \mathbf{1}_n[\bar{\gamma}\,\mathbf{1}_n - \mathbf{Cw}]' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n'.$$

A matrix of this form belongs in $\Lambda(\mathbf{w})$ as $\mathbf{w}'\Sigma$ works out as

$$\begin{aligned}\mathbf{w}'\Sigma &= \mathbf{w}'[\mathbf{C} + [\bar{\gamma}\,\mathbf{1}_n - \mathbf{Cw}]\mathbf{1}_n' + \mathbf{1}_n[\bar{\gamma}\,\mathbf{1}_n - \mathbf{Cw}]' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n'] \\ &= \mathbf{w'C} + \bar{\gamma}\,\mathbf{1}_n' - \mathbf{w'Cw1}_n' + \bar{\gamma}\,\mathbf{1}_n' - \mathbf{w'C} - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n' \\ &= \bar{\gamma}\,\mathbf{1}_n' - \mathbf{w'Cw1}_n' = [\bar{\gamma} - \mathbf{w'Cw}]\mathbf{1}_n',\end{aligned}$$

which is a scale multiple of the unit vector as required for membership in $\Lambda(\mathbf{w})$. This shows that the necessary and sufficient condition for a matrix in $\aleph_C$ to be contained in a subset $\Lambda(\mathbf{w})$ for some $\mathbf{w} \in \mathcal{W}(n)$ is that $-\mathbf{Cw} = \mathbf{B}\gamma$.

To prove (b) we first choose any matrix Σ in the subset $\beta(\mathbf{w},\mathbf{C})$ for some $\mathbf{w} \in \mathcal{W}(n)$ and $\mathbf{C} \in \mathcal{C}(n)$. Since Σ is in $\Lambda(\mathbf{w})$ we know from part (a) that $-\mathbf{Cw} = \mathbf{B}\gamma$ and hence

$$\gamma = \bar{\gamma}\,\mathbf{1}_n + \mathbf{B}\gamma = \bar{\gamma}\,\mathbf{1}_n - \mathbf{Cw}.$$

Using this expression in (3.5.1) we note that Σ takes the form

$$\Sigma = \mathbf{C} + [\bar{\gamma}\mathbf{1}_n - \mathbf{Cw}]\mathbf{1}_n' + \mathbf{1}_n[\bar{\gamma}\mathbf{1}_n - \mathbf{Cw}]' - \bar{\gamma}\mathbf{1}_n\mathbf{1}_n' \tag{3.5.2}$$

$$= \mathbf{C} + \bar{\gamma}\mathbf{1}_n\mathbf{1}_n' - \mathbf{Cw}\mathbf{1}_n' + \bar{\gamma}\mathbf{1}_n\mathbf{1}_n' - \mathbf{1}_n\mathbf{w}'\mathbf{C} - \bar{\gamma}\mathbf{1}_n\mathbf{1}_n'$$

$$= \mathbf{C} - \mathbf{Cw}\mathbf{1}_n' - \mathbf{1}_n\mathbf{w}'\mathbf{C} - \bar{\gamma}\mathbf{1}_n\mathbf{1}_n'.$$

This is the required representation on setting $\bar{\gamma} = \xi$. Further to show that $\xi > \mathbf{w}'\mathbf{Cw}$ for this representation we note that (3.5.2) is of the form

$$\Sigma = \mathbf{C} + \delta\mathbf{1}_n' + \mathbf{1}_n\delta' - \bar{\delta}\mathbf{1}_n\mathbf{1}_n'$$

for $\delta = [\bar{\gamma}\mathbf{1}_n - \mathbf{Cw}]$ as $\bar{\delta} = n^{-1}\mathbf{1}_n'[\bar{\gamma}\mathbf{1}_n - \mathbf{Cw}] = \bar{\gamma} - n^{-1}\mathbf{1}_n'\mathbf{Cw} = \bar{\gamma}$. Applying the condition of positive definiteness in (3.5.1) to δ we obtain

$$\delta'\mathbf{C}^{-1}_{\mathscr{C}(n)}\delta < \bar{\gamma} = \bar{\delta} = \xi.$$

The preceding expression leads to

$$[\bar{\gamma}\mathbf{1}_n - \mathbf{Cw}]'\mathbf{C}^{-1}_{\mathscr{C}(n)}[\bar{\gamma}\mathbf{1}_n - \mathbf{Cw}] = \mathbf{w}'\mathbf{C}\mathbf{C}^{-1}_{\mathscr{C}(n)}\mathbf{Cw} = \mathbf{w}'\mathbf{Cw} < \xi \tag{3.5.3}$$

as required on noting that $\mathbf{1}_n'\mathbf{C}^{-1}_{\mathscr{C}(n)} = \mathbf{0}$. To complete the proof to (b) we now need to demonstrate that a matrix Σ is a member of the subset $\beta(\mathbf{w},\mathbf{C})$ if it is given by

$$\Sigma = \Sigma(\xi) = \mathbf{C} - \mathbf{Cw}\mathbf{1}_n' - \mathbf{1}_n\mathbf{w}'\mathbf{C}' + \xi\mathbf{1}_n\mathbf{1}_n' \text{ for an } \xi > \mathbf{w}'\mathbf{Cw}.$$

To show that Σ is a member of the subset $\beta(\mathbf{w},\mathbf{C})$ we check to see if (i) it takes the form $\Sigma = \Sigma(\gamma) = \mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\mathbf{1}_n\mathbf{1}_n'$, (ii) it's positive definite, and (iii) if $\mathbf{w}'\Sigma$ is a scalar multiple of the unit vector. To show (i) we note that we can rewrite Σ as

$$\Sigma = \mathbf{C} - \mathbf{Cw}\mathbf{1}_n' - \mathbf{1}_n'\mathbf{w}'\mathbf{C}' + \xi\mathbf{1}_n\mathbf{1}_n'$$

$$= \mathbf{C} + \xi\mathbf{1}_n\mathbf{1}_n' - \mathbf{Cw}\mathbf{1}_n' + \xi\mathbf{1}_n\mathbf{1}_n' - \mathbf{1}_n\mathbf{w}'\mathbf{C} - \xi\mathbf{1}_n\mathbf{1}_n'$$

$$= \mathbf{C} + [\xi\mathbf{1}_n - \mathbf{Cw}]\mathbf{1}_n' + \mathbf{1}_n[\xi\mathbf{1}_n - \mathbf{Cw}]' - \xi\mathbf{1}_n\mathbf{1}_n',$$

which is of the form $\Sigma(\gamma)$ for $\gamma = [\xi\mathbf{1}_n - \mathbf{Cw}]$. For positive definiteness we need

$$\gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\gamma = [\xi\mathbf{1}_n - \mathbf{Cw}]'\mathbf{C}^{-1}_{\mathscr{C}(n)}[\xi\mathbf{1}_n - \mathbf{Cw}] = \mathbf{w}'\mathbf{Cw} < \xi.$$

The condition $\mathbf{w}'\mathbf{Cw} < \xi$ is assumed true by hypothesis and hence Σ is positive definite. To check for the third condition we find on using $\sum_{i=1}^{n} w_i = 1$ that

$$\mathbf{w}'\Sigma = \mathbf{w}'[\mathbf{C} - \mathbf{C}\mathbf{w}\mathbf{1}_n' - \mathbf{1}_n\mathbf{w}'\mathbf{C} + \xi\mathbf{1}_n\mathbf{1}_n']$$
$$= \mathbf{w}'\mathbf{C} - \mathbf{w}'\mathbf{C}\mathbf{w}\mathbf{1}_n' - \mathbf{w}'\mathbf{C} + \xi\mathbf{1}_n' = [\xi - \mathbf{w}'\mathbf{C}\mathbf{w}]\mathbf{1}_n'$$

is a scalar multiple of the unit vector. Demonstrating (i) through (iii) completes the proof to part (b). A matrix Σ is contained in $\beta(\mathbf{w},\mathbf{C})$ if and only if it has the required representation.

To prove (c) let Σ be a matrix of the form $\Sigma(\mathbf{C}, \xi)$ for some $\mathbf{C}$ in $\mathscr{C}(n)$ and some scalar ξ for which $\xi > \mathbf{w}'\mathbf{C}\mathbf{w}$. Then for that $\mathbf{C} \in \mathscr{C}(n)$ we note that Σ takes the form $\Sigma(\xi)$ as in part (b). Hence through the result in part (b) Σ is an element of $\beta(\mathbf{w},\mathbf{C})$ and therefore contained in $\Lambda(\mathbf{w})$. To complete the proof we now assume that a matrix Σ is an element of $\Lambda(\mathbf{w})$. By Definition 3.5 it then follows that this matrix is positive definite. By the partition of the set of positive definite matrices into subsets of the form $\aleph_C$ through Theorem 2.3.2(e) we know that Σ is an element of $\aleph_C$ for some $\mathbf{C}$ in $\mathscr{C}(n)$. For this $\mathbf{C}$ the matrix Σ therefore belongs to $\beta(\mathbf{w},\mathbf{C})$ and thus takes the form $\Sigma(\xi)$ as in part (b). Allowing $\mathbf{C}$ to be variable in $\Sigma(\xi)$ we obtain $\Sigma(\mathbf{C}, \xi)$ with the constraint $\xi > \mathbf{w}'\mathbf{C}\mathbf{w}$. With this we conclude that the stated expression for Σ and the constraint are necessary and sufficient for membership in $\Lambda(\mathbf{w})$. $\square$

Parts (a) and (b) of this theorem prescribe conditions for membership in $\Lambda(\mathbf{w}) \cap \aleph_C$. Matrices in such an intersection are obtained on using

$$\Sigma = \Sigma(\gamma) = \mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n'$$

for vectors γ meeting the constraints $\mathbf{B}\gamma = -\mathbf{C}\mathbf{w}$ and $\bar{\gamma} > \mathbf{w}'\mathbf{C}\mathbf{w}$. The first constraint forces admissible choices for γ from a straight line parallel to the equiangular line. Positive definiteness for Σ implies $\bar{\gamma} > \mathbf{w}'\mathbf{C}\mathbf{w}$. These observations are illustrated through the example that follows and through Figure 3.3. The example also computes a matrix in $\Lambda(\mathbf{w})$.

Example 3.6: Constraints on γ to generate matrices in $\Lambda(\mathbf{w}) \cap \aleph_C$ for $\mathbf{C} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ and $\mathbf{w}' = \begin{bmatrix} -0.5 & 1.5 \end{bmatrix}$.

Matrices in $\aleph_C$ are generated as

$$\Sigma = \mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\mathbf{1}_n\mathbf{1}_n'.$$

The constraints on γ are obtained as vectors γ for which $\gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\gamma < \bar{\gamma}$. When the matrix $\mathbf{C}$ has spectral decomposition $\mathbf{C} = \sum_{i=1}^{n-1} \xi_i \mathbf{g}_i\mathbf{g}_i'$, then $\mathbf{C}^{-1}_{\mathscr{C}(n)} = \sum_{i=1}^{n-1} \xi_i^{-1}(\mathbf{g}_i\mathbf{g}_i')$. The matrix $\mathbf{C}$ has the spectral decomposition $\mathbf{C} = 2\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}'\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$. Hence for this $\mathbf{C} \in \mathscr{C}(n)$ we need vectors γ such that

$$\gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\gamma = \frac{\gamma'\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}'\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}\gamma}{2} < \bar{\gamma}.$$

This works out to $(\gamma_1 - \bar{\gamma})^2 + (\gamma_2 - \bar{\gamma})^2 < 2\bar{\gamma}$, and constrains the two-dimensional vectors γ needed to generate $\aleph_C$ to those inside the convex set in Figure 3.3. The elements of the subset $\Lambda(\mathbf{w}) \cap \aleph_C$ are obtained under the further constraints $-\mathbf{Cw} = \mathbf{B}\gamma$ and $\mathbf{w}'\mathbf{Cw} < \bar{\gamma}$. In this example we have

$$\mathbf{B}\gamma = \begin{bmatrix} \gamma_1 - \bar{\gamma} \\ \gamma_2 - \bar{\gamma} \end{bmatrix} = -\mathbf{Cw} = -\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} -0.5 \\ 1.5 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix} \text{ and } \mathbf{w}'\mathbf{Cw} < \bar{\gamma}.$$

The identity $\begin{bmatrix} \gamma_1 - \bar{\gamma} \\ \gamma_2 - \bar{\gamma} \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$ forces choices of γ to be on a straight line parallel to the equiangular line as shown in Figure 3.3. The requirement $\mathbf{w}'\mathbf{Cw} < \bar{\gamma}$ leads to

$$\bar{\gamma} > \begin{bmatrix} -0.5 & 1.5 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} -0.5 \\ 1.5 \end{bmatrix} = \begin{bmatrix} -0.5 & 1.5 \end{bmatrix}\begin{bmatrix} 2 \\ -2 \end{bmatrix} = 4.$$

In Figure 3.3 we see that the constraint $\bar{\gamma} > 4$ makes us chose γ vectors on the straight line parallel to the equiangular line and also within the symmetric convex set. On using $\gamma' = \begin{bmatrix} 8 & 4 \end{bmatrix}$ from this constrained set, a matrix in $\Lambda(\mathbf{w})$ for $\mathbf{w}' = \begin{bmatrix} -0.5 & 1.5 \end{bmatrix}$ is obtained as

$$\Sigma = \mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\mathbf{1}_n\mathbf{1}_n' = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + \begin{bmatrix} 8 \\ 4 \end{bmatrix}\begin{bmatrix} 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 8 & 4 \end{bmatrix} - 6\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 11 & 5 \\ 5 & 3 \end{bmatrix}.$$

This matrix is also in $\aleph_C$ for the matrix $\mathbf{C}$ in this example.

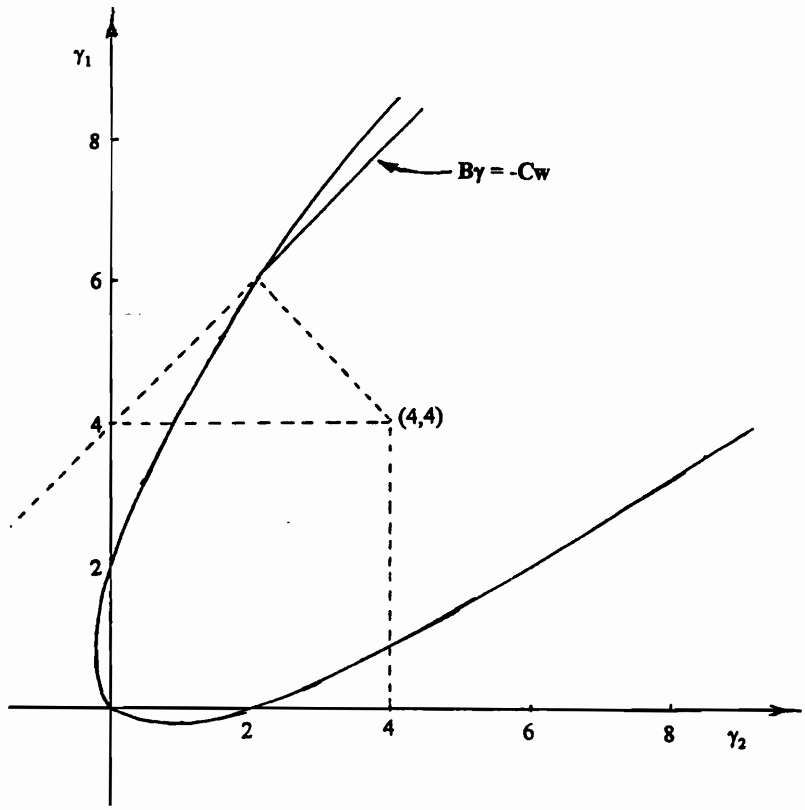

Figure 3.3: Constraints on γ in $\Re^2$ to generate matrices in $\beta(\mathbf{w},\mathbf{C})$ for $\mathbf{w}$ and $\mathbf{C}$ as in Example 3.6. The constraints $-\mathbf{Cw} = \mathbf{B}\gamma$ and $\mathbf{w}'\mathbf{Cw} < \bar{\gamma}$ together imply vectors γ chosen on the line inside the convex set as indicated.

Recall that a partition of S_n^+ into $\aleph_C$ sets was of interest in Section 3.3 as it allowed the decomposition of mixing distributions over S_n^+ into conditional distributions over each subset $\aleph_C$ in the partition. Using the representation in the preceding theorem, part (a) of the following result develops a partition of S_n^+ into $\Lambda(\mathbf{w})$ sets. These two partitions are then used in parts (b) and (c) to develop finer partitions for each $\Lambda(\mathbf{w})$ subset and each $\aleph_C$ subset of S_n^+.

Theorem 3.5.2: (a) The family $\ddot{\mathbf{A}} = (\ \Lambda(\mathbf{w})|\ \mathbf{w} \in \mathscr{W}(n)\ )$ of subsets of S_n^+ obtained on varying the vector $\mathbf{w}$ within $\mathscr{W}(n)$ partitions the set S_n^+ of positive definite matrices.
(b) For every $\mathbf{w} \in \mathscr{W}(n)$ the family $\ddot{\mathbf{E}}(\mathbf{w}) = \left(\beta(\mathbf{w},\mathbf{C})|\mathbf{C} \in \mathscr{C}(n)\right)$ of subsets of $\Lambda(\mathbf{w})$ obtained on varying $\mathbf{C}$ within $\mathscr{C}(n)$ partitions $\Lambda(\mathbf{w})$.
(c) Similarly for every $\mathbf{C} \in \mathscr{C}(n)$ the family $\ddot{\mathbf{O}}(\mathbf{C}) = \left(\beta(\mathbf{w},\mathbf{C})|\mathbf{w} \in \mathscr{W}(n)\right)$ of subsets of $\aleph_C$ obtained on varying the vector $\mathbf{w}$ within $\mathscr{W}(n)$ partitions the subset $\aleph_C$.
Proof: Following Berge (1963), we note that the required partition of S_n^+ in (a) is demonstrated if (i) for all $\mathbf{w} \in \mathscr{W}(n)$, $\Lambda(\mathbf{w}) \neq \varnothing$, and $\Lambda(\mathbf{w}) \subset S_n^+$, (ii) $\Lambda(\mathbf{w}_1) \cap \Lambda(\mathbf{w}_2) = \varnothing$ whenever $\mathbf{w}_1 \neq \mathbf{w}_2$, and (iii) $\bigcup_{\mathbf{w} \in \mathscr{W}} \Lambda(\mathbf{w}) = S_n^+$. From the representation in Theorem 3.5.1 part (c) note that for every $\mathbf{w} \in \mathscr{W}(n)$ a matrix Σ in $\Lambda(\mathbf{w})$ can be obtained as

$$\Sigma = \Sigma(\mathbf{C}, \xi) = \mathbf{C} - \mathbf{C}\mathbf{w}\mathbf{1}_n' - \mathbf{1}_n\mathbf{w}'\mathbf{C} + \xi\mathbf{1}_n\mathbf{1}_n' \tag{3.5.4}$$

for choices of $\mathbf{C} \in \mathscr{C}(n)$ and ξ such that $\xi > \mathbf{w}'\mathbf{C}\mathbf{w}$. Hence $\Lambda(\mathbf{w}) \neq \varnothing$. Further $\Lambda(\mathbf{w}) \subset S_n^+$ by definition. Preceding statements establish properties required in (i). To show (ii) choose any two distinct vectors $\mathbf{w}_1, \mathbf{w}_2 \in \mathscr{W}(n)$. The second condition follows if $\Sigma \in \Lambda(\mathbf{w}_1)$ implies that $\Sigma \notin \Lambda(\mathbf{w}_2)$. Choose any $\Sigma \in \Lambda(\mathbf{w}_1)$. Then by the results in part (c) of Theorem 3.5.1 this matrix is given by (3.5.4) for $\mathbf{w} = \mathbf{w}_1$, and for some $\mathbf{C} \in \mathscr{C}(n)$ and some $\xi \succ \mathbf{w}'\mathbf{C}\mathbf{w}$. For this matrix

$$\mathbf{w}_2'\Sigma = \mathbf{w}_2'[\mathbf{C} - \mathbf{C}\mathbf{w}_1\mathbf{1}_n' - \mathbf{1}_n\mathbf{w}_1'\mathbf{C}' + \xi\mathbf{1}_n\mathbf{1}_n']$$
$$= [\mathbf{w}_2\text{-}\mathbf{w}_1]'\mathbf{C} + [\xi - \mathbf{w}_2'\mathbf{C}\mathbf{w}_1]\mathbf{1}_n'.$$

In this expression $[\mathbf{w}_2\text{-}\mathbf{w}_1]'\mathbf{C}$ is not proportional to the unit vector as the null space of the matrix $\mathbf{C}$ in $\mathscr{C}(n)$ is spanned by the unit vector. Further $[\mathbf{w}_2\text{-}\mathbf{w}_1]'\mathbf{C}$ is non-zero as $[\mathbf{w}_2\text{-}\mathbf{w}_1]$ is not proportional to the unit vector. Thus $\mathbf{w}_2'\Sigma$ is not proportional to the unit vector and $\Sigma \notin \Lambda(\mathbf{w}_1)$. This establishes (ii). To verify (iii) we need to show that every matrix Σ in S_n^+ belongs to $\Lambda(\mathbf{w})$ for some $\mathbf{w} \in \mathscr{W}(n)$. Choose any Σ in S_n^+ and let $\mathbf{a}' = \left(\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n\right)^{-1}\mathbf{1}_n'\Sigma^{-1}$. Since $\mathbf{1}_n'\mathbf{a} = \mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n\left(\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n\right)^{-1} = 1.0$, it follows that $\mathbf{a} \in \mathscr{W}(n)$. Also

$$\mathbf{a}'\Sigma = \left(\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n\right)^{-1}\mathbf{1}_n'\Sigma^{-1}\Sigma = \left(\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n\right)^{-1}\mathbf{1}_n'$$

is a scalar multiple of the unit vector. Hence the chosen matrix belongs in the subset $\Lambda(\mathbf{a})$ for the specified $\mathbf{a} \in \mathscr{W}(n)$. This verifies (iii) and establishes the family $\ddot{\mathbf{A}} = \left(\Lambda(\mathbf{w}) \middle| \mathbf{w} \in \mathscr{W}(n)\right)$ of subsets of S_n^+ as a partition of S_n^+.

Statement (b) posits, for every $\mathbf{w} \in \mathscr{W}(n)$, a partition of $\Lambda(\mathbf{w})$ by the family $\ddot{\mathbf{E}}(\mathbf{w}) = \left(\beta(\mathbf{w},\mathbf{C}) \middle| \mathbf{C} \in \mathscr{C}(n)\right)$ of subsets of $\Lambda(\mathbf{w})$. For every $\mathbf{w} \in \mathscr{W}(n)$ this requires that (i) for all $\mathbf{C} \in \mathscr{C}(n)$, $\beta(\mathbf{w},\mathbf{C}) \neq \varnothing$ and $\beta(\mathbf{w},\mathbf{C}) \subset \Lambda(\mathbf{w})$, (ii) $\beta(\mathbf{w},\mathbf{C}_1) \cap \beta(\mathbf{w},\mathbf{C}_2) = \varnothing$ whenever $\mathbf{C}_1 \neq \mathbf{C}_2$, and (iii) $\bigcup_{\mathbf{C} \in \mathbf{R}(\mathbf{T})} \beta(\mathbf{w},\mathbf{C}) = \Lambda(\mathbf{w})$. From the representation in Theorem 3.5.1(b) we note that for every $\mathbf{w} \in \mathscr{W}(n)$ we can obtain a matrix Σ in $\beta(\mathbf{w},\mathbf{C})$ for all $\mathbf{C} \in \mathscr{C}(n)$ by choosing $\xi > \mathbf{w}'\mathbf{C}\mathbf{w}$. Further $\beta(\mathbf{w},\mathbf{C}) \subset \Lambda(\mathbf{w})$ by definition. This verifies (i). The second condition follows if $\Sigma \in \beta(\mathbf{w},\mathbf{C}_1)$ implies $\Sigma \notin \beta(\mathbf{w},\mathbf{C}_2)$ for each $\mathbf{w} \in \mathscr{W}(n)$ and all distinct matrices $\mathbf{C}_1$, $\mathbf{C}_2 \in \mathscr{C}(n)$. Consider any such pair of matrices and chose $\Sigma \in \beta(\mathbf{w},\mathbf{C}_1)$. This implies $\Sigma \in \aleph_{\mathbf{C}_1}$. The partition of S_n^+ into subsets of the form $\aleph_{\mathbf{C}}$ (Theorem 3.3.2 (b)), implies that $\Sigma \notin \aleph_{\mathbf{C}_2}$ as $\mathbf{C}_1 \neq \mathbf{C}_2$. Hence $\Sigma \notin \beta(\mathbf{w},\mathbf{C}_2)$. The third condition holds if for every $\mathbf{w} \in \mathscr{W}(n)$, each $\Sigma \in \Lambda(\mathbf{w})$ belongs in a subset $\beta(\mathbf{w},\mathbf{C})$ for some $\mathbf{C} \in \mathscr{C}(n)$. Choose any $\Sigma \in \Lambda(\mathbf{w})$. Then by definition this matrix is positive definite. By the partition S_n^+ into subsets of the form $\aleph_{\mathbf{C}}$, this $\Sigma \in \aleph_{\mathbf{C}}$ for some matrix $\mathbf{C}$

$\in \mathscr{C}(n)$. Then $\Sigma \in \beta(\mathbf{w},\mathbf{C})$ for this $\mathbf{C}$. With this we establish a partition of each $\Lambda(\mathbf{w})$ as the family $\ddot{\mathbf{E}}(\mathbf{C}) = \left(\beta(\mathbf{w},\mathbf{C}) \middle| \mathbf{C} \in \mathscr{C}(n)\right)$ of subsets of $\Lambda(\mathbf{w})$.

Statement (c) involves a similar partition of each $\aleph_C$ set. The conditions for this partition are verified using the partition of S_n^+ into $\Lambda(\mathbf{w})$ subsets in part (a) of this theorem. The proof follows on using similar arguments to those used for statement (b).

□

A special subset amongst the classes $\Lambda(\mathbf{w})$ is the focus of the next section.

3.6 THE EQUAL COLUMNS SUM CLASS

Since the sample average is a commonly used estimator for location, its admissibility under squared error loss has been examined. Kagan, Linnik, and Rao (1973) show that under conditions of independence and n>2, the admissibility of the sample average requires normality as noted earlier. We reexamine this in the absence of independence. It is in this context that a class, denoted by $\Omega(n)$, amongst classes $\Lambda(\mathbf{w})$, is of particular interest. This class satisfies $\Omega(n) = \Lambda(n^{-1}\mathbf{1}_n)$. For every elliptically symmetric random vector with a dispersion matrix in $\Omega(n)$ and a common scalar mean we find the sample average to be admissible. Further non-elliptically contoured distributions are identified for which the same conclusions apply. These derive from properties studied in this section. Matrices in this class have equal column sums. Example 3.7 gives a matrix belonging to this class. The definition of membership in $\Omega(n)$, obtained from Jensen (1989a), is formally stated in the following.

Definition 3.6: A (n×n) matrix Σ belongs to the class $\Omega(n)$, if and only if

(a) $\Sigma \in S_n^+$ and

(b) $\mathbf{1}_n'\Sigma = c\mathbf{1}_n$ for some constant c.

Example 3.7: A matrix Σ belonging to the class $\Omega(n)$.

$$\Sigma = \begin{bmatrix} 3 & 3 & 1 \\ 3 & 5 & -1 \\ 1 & -1 & 7 \end{bmatrix}$$

is in $\Omega(n)$, has column sums equal to 7, and is positive definite as it has a determinant of 28 > 0, and all principal minors are positive.

In part (a) of the theorem that follows the matrices Σ in $\Omega(n)$ are shown to be those in S_n^+ having the scaled unit vector $(1/\sqrt{n})\mathbf{1}_n$ as an eigenvector. Part (b) expresses each matrix in the subset obtained as $\Omega(n)\cap\aleph_C$ in terms of the parameter $\mathbf{C} \in \mathscr{C}(n)$, and a scalar ξ. The third part gives a representation of matrices in $\Omega(n)$. Part (d) gives a partition of this subset.

Theorem 3.6.1: (a) Let $\Sigma = \mathbf{EDE}'$ be a spectral decomposition for $\Sigma \in S_n^+$, where the column vectors of $\mathbf{E}$ are the n eigenvectors of Σ, and $\mathbf{D}$ is a diagonal matrix containing the corresponding eigenvalues. The necessary and sufficient condition for Σ to belong to $\Omega(n)$ is that $\mathbf{E}$ takes the form $\mathbf{K}$ in (3.2.1), consisting of the row vector $(1/\sqrt{n})\mathbf{1}_n'$ and some orthonormal completion $\mathbf{H}$ for it. The eigenvalue corresponding to $(1/\sqrt{n})\mathbf{1}_n$ is c, the common column sums for $\Sigma \in \Omega(n)$.

(b) For each $\mathbf{C} \in \mathscr{C}(n)$, a matrix Σ is in the intersection $\Omega(n)\cap\aleph_C$ if and only if it can be expressed as

$$\Sigma = \Sigma(\xi) = \mathbf{C} + \xi\mathbf{1}_n\mathbf{1}_n',$$

for some $\xi > 0$.

(c) For a matrix Σ to be contained in $\Omega(n)$ it is necessary and sufficient that the matrix has a representation as

$$\Sigma = \Sigma(\mathbf{C}, \xi) = \mathbf{C} + \xi\mathbf{1}_n\mathbf{1}_n',$$

for some $\mathbf{C} \in \mathscr{C}(n)$ and some $\xi > 0$.

(d) The family $\ddot{\mathbf{I}} = \left(\Omega(n)\cap\aleph_C \middle| \mathbf{C} \in \mathscr{C}(n)\right)$ of subsets of $\Omega(n)$ obtained on varying $\mathbf{C} \in \mathscr{C}(n)$ partitions $\Omega(n)$.

Proof: The result in Part (a) follows on noting that $(1/\sqrt{n})\mathbf{1}_n$ is an eigenvector of Σ if and only if

$(\Sigma-\lambda\mathbf{I}_n)(1/\sqrt{n})\mathbf{1}_n = \mathbf{0}$, where λ is the corresponding eigenvalue. This requires that

$$(1/\sqrt{n})\Sigma\mathbf{1}_n = (\lambda/\sqrt{n})\mathbf{1}_n$$

which is true if and only if the column sums of Σ are equal. Further it follows that the eigenvalue $\lambda = c$, the equal column sum.

In Theorem 3.5.1 (b) we noted that for any $\mathbf{w} \in \mathscr{W}(n)$ and any $\mathbf{C} \in \mathscr{C}(n)$ a matrix Σ belongs to the subset $\Lambda(\mathbf{w}) \cap \aleph_C$ if and only if it has a representation as

$$\Sigma = \Sigma(\xi) = \mathbf{C} - \mathbf{Cw1}_n' - \mathbf{1}_n\mathbf{w}'\mathbf{C}' + \xi\mathbf{1}_n\mathbf{1}_n',$$

in terms of the parameters $\mathbf{w}$ and $\mathbf{C}$ defining the subset and a scalar ξ such that $\xi > \mathbf{w'Cw}$. Note that $\mathbf{1}_n'\mathbf{C} = \mathbf{0}$, and $\Omega(n) = \Lambda(n^{-1}\mathbf{1}_n)$. These yield necessary and sufficient conditions for membership in $\Omega(n) \cap \aleph_C$ as the representation

$$\Sigma = \Sigma(\xi) = \mathbf{C} - \mathbf{Cw1}_n' - \mathbf{1}_n\mathbf{w}'\mathbf{C}' + \xi\mathbf{1}_n\mathbf{1}_n'$$

$$= \mathbf{C} - n^{-1}\mathbf{C1}_n\mathbf{1}_n' - n^{-1}\mathbf{1}_n\mathbf{1}_n'\mathbf{C}' + \xi\mathbf{1}_n\mathbf{1}_n' = \mathbf{C} + \xi\mathbf{1}_n\mathbf{1}_n',$$

for some $\xi > \mathbf{w'Cw} = n^{-2}\mathbf{1}_n'\mathbf{C1}_n = 0$. This proves (b). Part (c) follows in a similar manner from Theorem 3.5.1 (c).

The proposition in (d) follows from the partition established in Theorem 3.5.2 (b) as it applies to all $\mathbf{w} \in \mathscr{W}(n)$. □

This theorem illustrates how the $\Omega(n)$ and $\aleph_C$ subsets relate. With motivations similar to those expressed earlier we obtain the partition in (d). Part (b) provides a representation of matrices in each subset $\Omega(n) \cap \aleph_C$. Part (c) provides one for members of the equal column sums class of matrices. These representations can be used to compute matrices in these subsets. Example 3.8 does this. For the matrix $\mathbf{C} \in \mathscr{C}(n)$ in Example 2.2, two matrices in $\Omega(n) \cap \aleph_C$ are obtained.

Example 3.8: Matrices in $\aleph_C$ and $\Omega(n) \cap \aleph_C$ for $C = (1/9)\begin{bmatrix} 20 & 2 & -22 \\ 2 & 11 & -13 \\ -22 & -13 & 35 \end{bmatrix}$.

From Example 2.2(c), we note that the matrices $\Sigma_1 = \begin{bmatrix} 5 & 3 & 1 \\ 3 & 4 & 2 \\ 1 & 2 & 8 \end{bmatrix}$ and $\Sigma_2 = \begin{bmatrix} 7 & 6 & 5 \\ 6 & 8 & 7 \\ 5 & 7 & 14 \end{bmatrix}$ are in $\aleph_C$ for this $\mathbf{C} \in \mathscr{C}(n)$. However these matrices are not in $\Omega(n)$. In the expression in Theorem 3.6.1 (c), using ξ as 1 and 2 respectively yield

$$\Sigma_3 = \Sigma(\xi) = \mathbf{C} + \xi \mathbf{1}_n \mathbf{1}_n' = (1/9)\begin{bmatrix} 20 & 2 & -22 \\ 2 & 11 & -13 \\ -22 & -13 & 35 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = (1/9)\begin{bmatrix} 29 & 11 & -13 \\ 11 & 20 & -4 \\ -13 & -4 & 44 \end{bmatrix}$$

$$\text{and } \Sigma_4 = \mathbf{C} + 2\mathbf{1}_n \mathbf{1}_n' = (1/9)\begin{bmatrix} 20 & 2 & -22 \\ 2 & 11 & -13 \\ -22 & -13 & 35 \end{bmatrix} + 2\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = (1/9)\begin{bmatrix} 38 & 20 & -4 \\ 20 & 29 & 31 \\ -4 & 5 & 53 \end{bmatrix},$$

as elements of $\Omega(n)$ and also as members of $\aleph_C$.

The next section excerpts results in this chapter on the various subsets of the class of positive definite matrices and its dual partition.

3.7 OVERVIEW

This section summarizes notation and results developed about matrices of scale parameters. Section 3.3 developed results about the classes $\aleph_C$ (also referred to as $\Xi(\mathbf{A},\mathbf{H})$) of matrices in S_n^+. The index matrix $\mathbf{C}$ is an element of the class $\mathscr{C}(n)$ of n-dimensional positive semidefinite matrices of rank (n-1) with null space spanned by the unit vector. The notation $\mathbf{C}^{-1}_{\mathscr{C}(n)}$ denotes the minimum norm least squares inverse of the matrix $\mathbf{C}$. For a chosen $\mathbf{C} \in \mathscr{C}(n)$, the class $\aleph_C$ contains matrices which can be expressed as

$$\Sigma = \Sigma(\gamma) = \mathbf{C} + \gamma \mathbf{1}_n' + \mathbf{1}_n \gamma' - \bar{\gamma}\, \mathbf{1}_n \mathbf{1}_n',$$

for the parameter $\mathbf{C}$, and some vector γ such that $\gamma' \mathbf{C}^{-1}_{\mathscr{C}(n)} \gamma < \bar{\gamma}$. The family of subsets obtained on varying $\mathbf{C} \in \mathscr{C}(n)$ partitions the set of positive definite matrices. Figure 3.4 presents S_n^+ symbolically as a grid of subsets. The horizontal strips represent the $\aleph_C$ classes in S_n^+. A special subset amongst subsets $\aleph_C$ is examined in Section 3.4. This is the subset $\Gamma(n)$ containing matrices of the type studied by Huynh and Feldt (1970) and Rouanet and Le'pine (1970). These are represented as the central horizontal strip in the figure. On choosing $\mathbf{B} = [\mathbf{I}_n - (1/n)\mathbf{1}_n \mathbf{1}_n']$ from $\mathscr{C}(n)$ this subset is obtained as $\aleph_{\mathbf{B}}$.

The classes $\Lambda(\mathbf{w})$ of S_n^+ are studied in Section 3.5. For every elliptically symmetric random vector $\mathbf{X}$ having a dispersion matrix in each class $\Lambda(\mathbf{w})$, $\mathbf{w}'\mathbf{X}$ will be seen to be the Pitman estimator. This motivates $\Lambda(\mathbf{w})$. The vector $\mathbf{w}$ identifying each subset is selected from the class $\mathscr{W}(n)$ of all n-dimensional vectors $\mathbf{w}$ such that $\sum_{i=1}^{n} w_i = 1$. For a chosen $\mathbf{w} \in \mathscr{W}(n)$, matrices in the class $\Lambda(\mathbf{w})$ can be expressed as

$$\Sigma = \Sigma(\mathbf{C}, \xi) = \mathbf{C} - \mathbf{C}\mathbf{w}\mathbf{1}_n' - \mathbf{1}_n \mathbf{w}'\mathbf{C} + \xi \mathbf{1}_n \mathbf{1}_n'$$

in terms of the parameter $\mathbf{w}$, a matrix $\mathbf{C}$ in $\mathscr{C}(n)$ and a scalar ξ such that $\xi > \mathbf{w}'\mathbf{C}\mathbf{w}$. The family of subsets of S_n^+ obtained on varying $\mathbf{w} \in \mathscr{W}(n)$ partitions S_n^+. The vertical strips

	•	$\Lambda(\mathbf{w}_1)$	$\Lambda(\mathbf{w}_2)$	$\Lambda(n^{-1}\mathbf{1}_n{}') = \Omega(n)$	$\Lambda(\mathbf{w}_3)$	$\Lambda(\mathbf{w}_4)$	•
•							
$\aleph_{C_1}$		•	•	•	•	•	
$\aleph_{C_2}$		•	$\beta(\mathbf{w}_2,\mathbf{C}_2)$	$\aleph_{C_2} \cap \Omega(n)$	$\beta(\mathbf{w}_3,\mathbf{C}_2)$	•	
$\aleph_B = \Gamma(n)$		•	$\beta(\mathbf{w}_2,\mathbf{B})$	$\Gamma(n) \cap \Omega(n) = \overline{\Gamma}(n)$	$\beta(\mathbf{w}_3,\mathbf{B})$	•	
$\aleph_{C_3}$		•	$\beta(\mathbf{w}_2,\mathbf{C}_3)$	$\aleph_{C_3} \cap \Omega(n)$	$\beta(\mathbf{w}_3,\mathbf{C}_3)$	•	
$\aleph_{C_4}$		•	•	•	•	•	
•							

Figure 3.4: Partition of the set S_n^+ of positive definite matrices. The partition is represented as a grid. The horizontal strips are the $\aleph_C$ classes and the vertical strips are the $\Lambda(\mathbf{w})$ classes. The $\Gamma(n)$ and $\Omega(n)$ subsets are special subsets of these classes. They are the subsets containing matrices studied by Huynh and Feldt (1970) and the equal column sum matrices respectively. For any pair $(\mathbf{w},\mathbf{C})$ from the ordered set $[\mathcal{W},\mathbf{R}(\mathbf{T})]$, the subset $\beta(\mathbf{w},\mathbf{C}) = \Lambda(\mathbf{w}) \cap \aleph_C$. The intersection of $\Gamma(n)$ and $\Omega(n)$ is $\overline{\Gamma}(n)$. The matrices in this subset are equicorrelated.

in Figure 3.4 represent the classes $\Lambda(\mathbf{w})$. The middle strip represents the special class $\Omega(n)$ consisting of matrices with equal column sums. Each cell in the figure corresponds to subsets obtained as the intersections $\aleph_C \cap \Lambda(\mathbf{w})$. These subsets are denoted as $\beta(\mathbf{w},\mathbf{C})$. The intersections $\aleph_C \cap \Omega(n)$ obtained as $\mathbf{C}$ varies over $\mathscr{C}(n)$ partitions $\Omega(n)$. For each $\mathbf{w} \in \mathscr{W}(n)$ a similar conclusion holds for the intersections of $\Lambda(\mathbf{w})$ with $\aleph_C$ as $\mathbf{C}$ is varied over $\mathscr{C}(n)$. This and the partition of each $\aleph_C$ set as $\mathbf{w}$ varies over $\mathscr{W}(n)$ are illustrated by the figure. Amongst the intersection subsets $\beta(\mathbf{w},\mathbf{C})$ the subset $\Gamma(n) \cap \Omega(n)$ is denoted as $\overline{\Gamma}(n)$. The subset $\overline{\Gamma}(n)$ consists of elements in the class of equicorrelated matrices.

Mixtures of elliptically symmetric distributions are of interest in the next chapter. Results in this chapter relate as the mixtures arise through distributions over S_n^+. Mixtures for probability measures restricted in two ways are of particular interest. Firstly, the domain of the probability measure is restricted to each subset of S_n^+ in the partitions discussed in this chapter. Secondly, lesser restrictions pertain to probability measures with domain extending across these partitions. Restrictions then apply as properties required conditionally over each subset in the domain of the probability measure. This motivates the study of subsets and partitions of the set of positive definite matrices in this chapter as foundations to developments in the next chapter.

Chapter 4

PITMAN ESTIMATION FOR ENSEMBLES AND MIXTURES

4.1 INTRODUCTION:

This chapter provides an application for the various subsets of positive definite matrices studied in Chapter 3. This application has similar objectives to those of Kagan, Linnik, and Rao (1973). Their results are reviewed in greater detail in Chapter 2. Their study examines distributions in the location family assuming independence and moments to order 2. For these conditions they demonstrate that the unbiased admissibility of the best linear estimator is a property which characterizes normality.

In this chapter the assumptions of independence is relaxed through distributional assumptions from within two classes of distributions. These classes of distributions were introduced in Section 2.3. They are the elliptically symmetric distributions and the distributions obtained as mixtures of these. The MRE estimator under squared error loss (the Pitman estimator) is obtained for members of these classes. For both contexts we find that amongst equivariant estimators the admissible estimator is linear. Section 4.2 derives the general form of the Pitman estimate for elliptically symmetric distributions. Section 4.3 applies this result to contexts where the scale parameter Σ belongs to each class $\aleph_C$ and to each class $\Lambda(\mathbf{w})$. In Section 4.4 a general form for the Pitman estimate for mixtures is obtained. Recall that these mixtures are obtained through the use of a probability measure $\mathbf{G}$ over the set of positive definite matrices. Section 4.5 applies this to contexts where the domain of the probability measure $\mathbf{G}$ is restricted. The restrictions on the domain are to each class $\aleph_C$ and to each class $\Lambda(\mathbf{w})$. Section 4.6 provides conditions for admissibility of linear functions with particular emphasis on the sample average.

4.2 THE PITMAN ESTIMATOR FOR ELLIPTICALLY SYMMETRIC DISTRIBUTIONS

The following notation and ideas in Chapters 2 and 3 are pertinent. Random vectors are denoted as $\mathbf{X}$ and take values $\mathbf{x}$ in $\Re^n$. The class of matrices $\mathcal{H}$ consist of all (n-1)×n dimensional matrices $\mathbf{H}$ which are semiorthogonal completions of the row vector $(1/\sqrt{n})\mathbf{1}_n'$. The minimal risk equivariant (MRE) estimate for location under squared error loss is a Pitman estimate. This section deals with the general form taken by the Pitman estimate in the context of elliptically symmetric random vectors. The distribution of such a random vector is denoted as $L(\mathbf{X}) = E_n(\mu\mathbf{1}_n, \Sigma, \phi)$, where ϕ specifies the particular elliptically symmetric distribution. Pitman estimates are obtained by using the distribution of any equivariant estimator $\delta_0(\mathbf{X})$ conditional on the value observed for any maximal invariant function of $\mathbf{x}$ under translations of the form $\mathbf{x} \rightarrow \mathbf{x} + t\mathbf{1}_n$. In this application, $\delta_0(\mathbf{X})$ is chosen to be the sample average $\overline{X}$. For any $\mathbf{H} \in \mathcal{H}$, the function $\mathbf{d}(\mathbf{x}) = \mathbf{Hx}$ provides the required maximal invariant. This is demonstrated in the following.

Lemma 4.2.1: For $\mathbf{x} \in \Re^n$, consider the group of transformations

$$\mathbf{g}(\mathbf{x}) = \mathbf{x} + t\mathbf{1}_n,$$

for $t \in \Re^1$, the real line. Then for all $\mathbf{H} \in \mathcal{H}$, the function $\mathbf{d}(\mathbf{x}) = \mathbf{Hx}$ is maximal invariant.

Proof: To prove this choose any $\mathbf{H} \in \mathcal{H}$ and let $\mathbf{d}(\mathbf{x}) = \mathbf{Hx}$. This is maximal invariant if it is invariant, i.e.

$$\mathbf{d}(\mathbf{x}) = \mathbf{d}[\mathbf{g}(\mathbf{x})]$$

for all $\mathbf{x} \in \Re^n$, and if the condition $\mathbf{d}(\mathbf{x}_1) = \mathbf{d}(\mathbf{x}_2)$ implies that

$$\mathbf{x}_1 = \mathbf{g}(\mathbf{x}_2) = \mathbf{x}_2 + t\mathbf{1}_n,$$

for some $t \in \Re^1$. Using the property $\mathbf{H1}_n = \mathbf{0}$, the first requirement follows as

$$\mathbf{d}[\mathbf{g}(\mathbf{x})] = \mathbf{H}[\mathbf{x} + t\mathbf{1}_n] = \mathbf{Hx} + \mathbf{0} = \mathbf{d}(\mathbf{x}),$$

for any $\mathbf{x} \in \Re^n$ and any $t \in \Re^1$. To verify the second requirement impose the condition $\mathbf{d}(\mathbf{x}_1) = \mathbf{d}(\mathbf{x}_2)$ or $\mathbf{Hx}_1 = \mathbf{Hx}_2$. Premultiplying both sides by $\mathbf{H}'$ and using $\mathbf{H}'\mathbf{H} = [\mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n']$ yields

$$\mathbf{H}'\mathbf{Hx}_1 = \mathbf{H}'\mathbf{Hx}_2$$

$$\Rightarrow [\,\mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n']\mathbf{x}_1 = [\,\mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n']\mathbf{x}_2$$

$$\Rightarrow \mathbf{x}_1 - \bar{x}_1\mathbf{1}_n = \mathbf{x}_2 - \bar{x}_2\mathbf{1}_n.$$

Thus $\mathbf{x}_1 = \mathbf{x}_2 + (\bar{x}_1 - \bar{x}_2)\mathbf{1}_n$ is of the form $\mathbf{x}_1 = \mathbf{x}_2 + t\mathbf{1}_n$ for $t = \bar{x}_1 - \bar{x}_2$. Hence $\mathbf{x}_1 = \mathbf{g}(\mathbf{x}_2)$ whenever $\mathbf{d}(\mathbf{x}_1) = \mathbf{d}(\mathbf{x}_2)$. This completes the proof. □

In succeeding results we choose $\delta_0(\mathbf{X}) = \overline{X}$, and $\mathbf{d}(\mathbf{x})$ as the maximal invariant. The Pitman estimate is obtained on using the result in Lehmann (1983) reproduced as Theorem 2.2.1 in this dissertation. This estimate is a correction of any equivariant estimator by the conditional mean of this estimator when $\mu = 0$. This correction to $\overline{X}$ is denoted by $\mathscr{E}_0[\overline{X}|\mathbf{d}]$. The estimate is given by

$$\delta\,(\mathbf{x}\,;\mathbf{d}) = \overline{X} - \mathscr{E}_0[\overline{X}|\mathbf{d}], \tag{4.2.1}$$

for some $\mathbf{d}$ as in the lemma. Replacing the conditional mean in (4.2.1) by a conditional median $\mathbf{M}_0[\overline{X}|\mathbf{d}]$ yields the MRE estimate under Laplace loss. These conditional parameters are obtained for elliptically symmetric distributions in the following.

<u>Lemma 4.2.2:</u> For $L(\mathbf{X}) = E_n(\mu\mathbf{1}_n, \Sigma, \phi)$, let $\mathbf{D} = \mathbf{D}(\mathbf{X})$ be the random vector $\mathbf{HX}$ for some $\mathbf{H} \in \mathscr{H}$. Then for a sample vector $\mathbf{x} \in \Re^n$, the distribution of $\overline{X}$ given that $\mathbf{D}$ takes the value $\mathbf{d} = \mathbf{Hx}$ is

$$L(\overline{X} \mid \mathbf{D} = \mathbf{Hx}) = E_1\left[\mu + n^{-1}\mathbf{c'H'(H\Sigma H')^{-1}Hx},\ n^{-2}\left(\sum_{i,j}\sigma_{ij} - \mathbf{c'H'(H\Sigma H')^{-1}Hc}\right), \phi^*\right] \tag{4.2.2}$$

where $[\,\sigma_{ij}\,]$ are the elements of the matrix Σ, $\mathbf{c}$ is the vector of column sums of Σ, and ϕ^* depends on ϕ through the quadratic form $\mathbf{x'H'(H\Sigma H')^{-1}Hx}$.

<u>Proof:</u> For the chosen matrix $\mathbf{H} \in \mathscr{H}$, let $\mathbf{T}$ be the matrix

$$\mathbf{T} = \begin{bmatrix} n^{-1}\mathbf{1}_n \\ \mathbf{H} \end{bmatrix}$$

On using properties of elliptically symmetric distributions in Theorem 2.3.1 (b) and (c), we have

$$L(\mathbf{TX}) = L\begin{bmatrix} \overline{X} \\ \mathbf{HX} \end{bmatrix} = E_n\left[\begin{bmatrix} \mu \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} n^{-1}\mathbf{1}_n' \\ \mathbf{H} \end{bmatrix}\Sigma\begin{bmatrix} n^{-1}\mathbf{1}_n \\ \mathbf{H}' \end{bmatrix}\right]$$

$$= E_n\left[\begin{bmatrix} \mu \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{bmatrix}\right] \tag{4.2.3}$$

where $\Delta_{11} = n^{-1}\mathbf{1}_n'\Sigma\mathbf{1}_n n^{-1} = n^{-2}\sum_{i,j}\sigma_{ij}$, $\Delta_{12} = n^{-1}\,\mathbf{1}_n'\Sigma\mathbf{H}' = n^{-1}\mathbf{c'H'}$, and $\Delta_{22} = \mathbf{H\Sigma H'}$.

Using Theorem 2.3.1 (d), which gives the conditional distribution for elliptically symmetric distributions, we get the required result. □

A one-dimensional elliptically symmetric distribution is a one-dimensional symmetric distribution. The location parameter in (4.2.2) is the conditional mean of the distribution when first order moments for the distribution of $\mathbf{X}$ exist, and is also the conditional median. Note that obtaining the conditional distribution in the preceding lemma by conditioning on any one maximal invariant is equivalent to conditioning on any other. This follows as invertible transformations connect maximal invariant functions. Consider the set of maximal invariants obtained as $\mathbf{Hx}$ on varying $\mathbf{H} \in \mathscr{H}$. Then the parameters of the conditional distribution in Lemma 4.2.2 are invariant to the particular choice of the maximal invariant from this set. To see this consider $\mathbf{H}_1 \neq \mathbf{H}_2$ from $\mathscr{H}$. Then by Lemma 3.2.1 $\mathbf{H}_1 = \mathbf{OH}_2$ for some orthogonal matrix $\mathbf{O}$ of order (n-1). This implies

$$\mathbf{H}_1'[\mathbf{H}_1\Sigma\mathbf{H}_1']^{-1}\mathbf{H}_1 = \mathbf{H}_2'\mathbf{O}'[\mathbf{OH}_2\Sigma\mathbf{H}_2'\mathbf{O}']^{-1}\mathbf{H}_2 = \mathbf{H}_2'[\mathbf{H}_2\Sigma\mathbf{H}_2']^{-1}\mathbf{H}_2,$$

leading to the invariance of the parameters in Lemma 4.2.2. Setting $\mu = 0$ for the location parameter gives the correction which when applied to the sample average yields the MRE estimator. This estimator is in Theorem 4.2.1. For elliptically symmetric distributions the estimator continues to take the Aitkin (1934) form. An algebraic identity to this form in terms of matrices $\mathbf{H} \in \mathscr{H}$ is provided and applied in the next section. The estimate is also MRE under Laplace loss functions .

Theorem 4.2.1: For $L(\mathbf{X}) = E_n(\mu\mathbf{1}_n, \Sigma, \phi)$, let $\overline{X}$ have finite risk. Then the MRE estimate for the location parameter μ is given for each $\mathbf{H} \in \mathscr{H}$ as

$$\delta(\mathbf{x}\,;\mathbf{d}) = \overline{X} - n^{-1}\,\mathbf{c}'\mathbf{H}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{Hx} = (\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n)^{-1}\mathbf{1}_n'\Sigma^{-1}\mathbf{x}, \tag{4.2.4}$$

where $\mathbf{c}$ is the vector of column sums of Σ. The variance of this estimator is given by

$$\mathrm{Var}\,[\delta(\mathbf{x};\mathbf{d})] = \alpha_\phi n^{-2}\Big(\sum_{i,j}\sigma_{ij} - \mathbf{c}'\mathbf{H}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{Hc}\Big) = \alpha_\phi(\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n)^{-1}, \tag{4.2.5}$$

where $\alpha_\phi > 0$ is a constant depending on ϕ. This constant equals 1 when the distribution of $\mathbf{X}$ is normal.

Proof: When $\mu = 0$, the mean and median of the distribution in (4.2.2) is

$$n^{-1}\,\mathbf{c}'\mathbf{H}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{Hx}.$$

Hence the MRE estimate is obtained by subtracting this expression from $\overline{X}$. That yields the expression in the middle of (4.2.4) as the Pitman estimate. The estimate depends on Σ but does not depend on ϕ implying that the estimate is the same for elliptically symmetric distributions. The expression to the right in (4.2.4) is the generalized least squares estimate which is uniformly minimal variance unbiased (UMVU) when **X** is normally distributed. Equivariant UMVU estimators are identical to the MRE estimator under squared error loss (Lehmann 1983, p. 163). This yields the second equality in (4.2.4). From (4.2.2), the variance of the estimator is given by

$$\mathrm{Var}\,[\delta\,(\mathbf{x};\,\mathbf{d})] = \alpha_\phi n^{-2}\big(\sum_{i,j} \sigma_{ij} - \mathbf{c}'\mathbf{H}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{H}\mathbf{c}\big) = \alpha_\phi(\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n)^{-1}$$

as required in the middle of expression (4.2.5). The expression to the right in (4.2.5) follows using arguments similar to those used for the right hand side of expression (4.2.4). □

This theorem gives the general form taken by the MRE estimator for elliptically symmetric random vectors. This estimator has the following three properties. Firstly, the estimator is a weighted average of the elements of the random vector. This follows on noting that $\delta(\mathbf{x};\,\mathbf{d}) = \mathbf{k}'\mathbf{x}$ with $\mathbf{1}_n'\mathbf{k} = \mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n(\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n)^{-1} = 1$. Secondly, the first point implies that the estimator is unbiased as each component of **X** is unbiased. Thirdly, the Aitkin (1934) form reveals that the estimator is independent of ϕ and the value of the maximal invariant **d**. The estimator depends on the parameter only through Σ. This supports the alternate notation of $\delta_\Sigma(\mathbf{X})$ for the estimator in succeeding results. The next section applies (4.2.4) to positive definite matrices which belong to each class $\aleph_C$ and to each class $\Lambda(\mathbf{w})$.

4.3 ESTIMATION IN ELLIPTICALLY SYMMETRIC DISTRIBUTIONS: THE CLASSES $\aleph_C$ AND $\Lambda(w)$

The general form of the Pitman estimator obtained in the preceding section is now specialized to each class $\aleph_C$. From earlier developments the matrix $\mathbf{C}$ indexing each class is a member of the class $\mathscr{C}(n)$ comprising the positive semidefinite matrices of order n with a one-dimensional null space spanned by the unit vector. Each matrix $\mathbf{C}$ in $\mathscr{C}(n)$ can be expressed as $\mathbf{H'AH}$ for some $(\mathbf{A}, \mathbf{H})$ from the ordered set $\{S^+_{(n-1)}, \mathscr{H}\}$. The minimum norm least squares inverse of $\mathbf{C}$ is also contained in $\mathscr{C}(n)$ and is denoted as $\mathbf{C}^{-1}_{\mathscr{C}(n)}$. This inverse equals $\mathbf{H'A^{-1}H}$ whenever $\mathbf{C} = \mathbf{H'AH}$. For each $\mathbf{C}$, the necessary and sufficient condition for a matrix Σ to be a member of $\aleph_C$ is that it have an expansion as

$$\Sigma = \Sigma(\gamma) = \mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n' \qquad (4.3.1)$$

for a vector γ of constants such that $\gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\gamma < \bar{\gamma}$. The theorem that follows obtains the Pitman estimator for members of each class $\aleph_C$.

Theorem 4.3.1: Let $L(\mathbf{X}) = E_n(\mu\mathbf{1}_n, \Sigma, \phi)$ be the distribution for an n-dimensional random vector $\mathbf{X}$ with dispersion matrix $\Sigma(\gamma)$ in $\aleph_C$ as in (4.3.1). Then the MRE estimator for μ is

$$\delta_\Sigma(\mathbf{X}) = \overline{X} - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\mathbf{X}.$$

The variance of this estimator is given by

$$\mathrm{Var}\,[\,\delta_\Sigma(\mathbf{X};\phi)\,] = \alpha_\phi(\,\bar{\gamma} - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\gamma)$$

with α_ϕ depending on the particular elliptically contoured distribution and being equal to 1 when the distribution is normal.

Proof: Using Lemma 3.3.2 express $\mathbf{C}$ in $\mathscr{C}(n)$ as $\mathbf{C} = \mathbf{H'AH}$ for some pair of matrices $(\mathbf{A}, \mathbf{H})$ from the ordered set $\{S^{+}_{(n-1)}, \mathscr{H}\}$. For this $\mathbf{H}$ it follows from Theorem 4.2.1 that the Pitman estimator can be expressed as

$$\delta_{\Sigma}(\mathbf{X}) = \overline{X} - n^{-1}\, \mathbf{c'H'(H\Sigma H')^{-1}HX},$$

with variance

$$\mathrm{Var}\,[\, \delta_{\Sigma}(\mathbf{X})\,] = \alpha_{\phi} n^{-2}\Big(\sum_{i,j} \sigma_{ij} - \mathbf{c'H'(H\Sigma H')^{-1}Hc}\Big),$$

where α_{ϕ} depends on the particular elliptically contoured distribution. Since $\Sigma \in \Xi(\mathbf{A}, \mathbf{H})$ it follows by Definition 3.2 that $\mathbf{H\Sigma H'} = \mathbf{A}$, so that

$$\mathbf{H'(H\Sigma H')^{-1}H} = \mathbf{H'A^{-1}H} = \mathbf{C}^{-1}_{\mathscr{C}(n)}.$$

Further note that

$$n^{-1}\,\mathbf{c'} = n^{-1}\mathbf{1}_n'\Sigma = n^{-1}\mathbf{1}_n'[\mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n'] = \gamma'.$$

Similarly $\sum_{i,j} \sigma_{ij} = \mathbf{1}_n'\Sigma\mathbf{1}_n = n^2\bar{\gamma}$. Substituting these in the general expressions the estimator and the variance of the estimator are given as

$$\delta_{\Sigma}(\mathbf{X}) = \overline{X} - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\mathbf{X}, \text{ and } \mathrm{Var}\,[\, \delta_{\Sigma}(\mathbf{X})\,] = \alpha_{\phi}(\, \bar{\gamma} - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\gamma),$$

respectively, as required. □

Example 4.1 illustrates the theorem. The corollary that follows applies the theorem to the context of matrices studied by Huynh and Feldt (1970) and Rouanet and Le'pine (1970).

Example 4.1: MRE estimation for μ for a sample vector from an elliptically contoured distribution with $\Sigma \in \aleph_C$ with $\mathbf{C} = (1/9)\begin{bmatrix} 20 & 2 & -22 \\ 2 & 11 & -13 \\ -22 & -13 & 35 \end{bmatrix}$.

Let $\mathbf{x}' = (10\ 12\ 14)$ be a sample vector from a distribution with $\Sigma = \begin{bmatrix} 5 & 3 & 1 \\ 3 & 4 & 2 \\ 1 & 2 & 8 \end{bmatrix}$.

Through Example 3.2, this matrix was seen to belong to the class $\aleph_C$ with $\mathbf{C}$ as specified. Further, $\Sigma = \mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' + \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n'$ for $\gamma' = [3, 3, 11/3\,]$ and $\mathbf{C} = \mathbf{H'AH}$ for the matrices $\mathbf{H} = \begin{bmatrix} -1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix}$ and $\mathbf{A} = \begin{bmatrix} 11/6 & 5/\sqrt{12} \\ 5/\sqrt{12} & 11/2 \end{bmatrix}$. Hence

$$\mathbf{C}^{-1}_{\mathscr{C}(n)} = \mathbf{H'A^{-1}H} = \begin{bmatrix} -1/\sqrt{6} & 1/\sqrt{2} \\ 2/\sqrt{6} & 0 \\ -1/\sqrt{6} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 11/16 & -5/8\sqrt{12} \\ -5/8\sqrt{12} & 11/48 \end{bmatrix} \begin{bmatrix} -1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix}$$

$$= (1/24)\begin{bmatrix} 8 & -8 & 0 \\ -8 & 11 & -3 \\ 0 & -3 & 3 \end{bmatrix}.$$

Using this matrix the Pitman estimate is obtained as

$$\delta_\Sigma(\mathbf{X}) = \overline{\mathrm{X}} - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\mathbf{x}$$

$$= 12 - [3\;\;3\;\;11/3\,](1/24)\begin{bmatrix} 8 & -8 & 0 \\ -8 & 11 & -3 \\ 0 & -3 & 3 \end{bmatrix}\begin{bmatrix} 10 \\ 12 \\ 14 \end{bmatrix} = 11\tfrac{5}{6}.$$

The variance for the estimator is

$$\mathrm{Var}\,[\,\delta_\Sigma(\mathbf{X};\phi)\,] = \alpha_\phi(\,\bar{\gamma} - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\gamma)$$

$$= 29\alpha_\phi/9 - \alpha_\phi[3\;\;3\;\;11/3\,](1/24)\begin{bmatrix} 8 & -8 & 0 \\ -8 & 11 & -3 \\ 0 & -3 & 3 \end{bmatrix}\begin{bmatrix} 3 \\ 3 \\ 11/3 \end{bmatrix} = 29\alpha_\phi/9 - \alpha_\phi/18 = 57\alpha_\phi/18.$$

with α_ϕ being equal to 1 when the distribution is normal.

Corollary 4.3.1: For $L(\mathbf{X}) = \mathrm{E}_n(\,\mu\mathbf{1}_n, \Sigma, \phi\,)$, with $\Sigma \in \Gamma(n)$ obtained from a vector of constants γ as $\Sigma = \Sigma(\gamma) = \mathbf{B} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n'$, the MRE estimate for μ is

$$\delta_\Sigma(\mathbf{x}) = \overline{\mathrm{X}} - \gamma'\mathbf{e}$$

where $\mathbf{e}$ is the vector of OLS residuals given by $e_i = x_i - \bar{x}$. Its variance is

$$\mathrm{Var}\,[\,\delta_\Sigma(\mathbf{X})\,] = \alpha_\phi[\,\bar{\gamma} - \sum_{i=1}^{n}(\gamma_i - \bar{\gamma})^2\,],$$

where γ_i are the elements of γ, with average $\bar{\gamma}$, and α_ϕ is a constant depending on the particular elliptically contoured distribution. This constant equals 1 when the distribution is normal.

Proof: Observe that $\Gamma(n) = \aleph_B$ for $\mathbf{B} = [\, \mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n']$. Using $\mathbf{Bx} = \mathbf{e}$, $\gamma'\mathbf{B}\gamma = \sum_{i=1}^{n}(\gamma_i - \bar{\gamma})^2$, and noting that $\mathbf{B}^{-1}_{\mathscr{C}(n)} = \mathbf{B}$, we obtain the required results from expressions in Theorem 4.3.1. □

The expression for the variance in this corollary has an interesting feature. The variance decreases with increasing $\sum_{i=1}^{n}(\gamma_i - \bar{\gamma})^2$. It follows that for any two matrices $\Sigma(\gamma_1)$ and $\Sigma(\gamma_2)$ with the vectors γ_1 and γ_2 having the same $\bar{\gamma}$, the matrix associated with the vector more dispersed in the sense of majorization, has lower variance for the Pitman estimator. This conclusion is counter-intuitive as Jensen (1992) shows that on requiring any two admissible vectors γ_1 and γ_2 to have equal sum for their components, the more ill-conditioned matrix is the one associated with the more dispersed vector. An application of the result in Corollary 4.3.1 is in the following.

Example 4.2: MRE estimate for μ for elliptically contoured distribution with $\Sigma \in \Gamma(3)$.

Let $\mathbf{x}' = (10\ 12\ 14)$ be a sample vector from a distribution with

$$\Sigma = \begin{bmatrix} 3 & 3 & 4 \\ 3 & 5 & 5 \\ 4 & 5 & 7 \end{bmatrix}.$$

In Example 3.4, this matrix was seen to belong to $\Gamma(n)$ for $\gamma' = (1/3)(10\ 13\ 16)$. Hence the estimate is given by

$$\delta_\Sigma(\mathbf{X}) = \overline{\mathrm{X}} - \gamma'\mathbf{e} = 36/3 - (1/3)(10\ 13\ 16)\begin{bmatrix} -2 \\ 0 \\ 2 \end{bmatrix} = 12 - (1/3)(-20+32) = 8.$$

Further

$$\mathrm{Var}\,[\,\delta_\Sigma(\mathbf{X})\,] = \alpha_\phi[\,\bar{\gamma} - \sum_{i=1}^{n}(\gamma_i - \bar{\gamma})^2\,] = 39\alpha_\phi/9 - 18\,\alpha_\phi/9 = 2.33\alpha_\phi.$$

with α_ϕ being equal to 1 when the distribution is normal.

Succeeding results in this section concern the subset $\Lambda(\mathbf{w})$ of S_n^+. Such subsets are defined through n-dimensional vectors $\mathbf{w}$ such that $\mathbf{1}_n'\mathbf{w} = 1$. The collection of all such vectors is denoted as $\mathscr{W}(n)$ as before. For each $\mathbf{w} \in \mathscr{W}(n)$, $\Lambda(\mathbf{w})$ is the subset of S_n^+ within which every element Σ has $\mathbf{w}'\Sigma = c\mathbf{1}_n'$ for some scalar c. A useful identity concerning matrices in $\Lambda(\mathbf{w})$ is in the following.

Lemma 4.3.1: For $\Sigma \in S_n^+$ and $\mathbf{w} \in \mathscr{W}(n)$ the identity

$$n^{-1}\mathbf{1}_n' - n^{-1}\mathbf{c}'\mathbf{H}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{H} = \mathbf{w}$$

holds for each $\mathbf{H} \in \mathscr{H}$, if and only if $\Sigma \in \Lambda(\mathbf{w})$.

Proof: Choose any $\mathbf{w} \in \mathscr{W}(n)$. To demonstrate sufficiency consider $\Sigma \in \Lambda(\mathbf{w})$ for this $\mathbf{w} \in \mathscr{W}(n)$. By Theorem 3.5.1 (c) this matrix can be written as

$$\Sigma = \Sigma(\mathbf{C}, \xi) = \mathbf{C} - \mathbf{C}\mathbf{w}\mathbf{1}_n' - \mathbf{1}_n\mathbf{w}'\mathbf{C}' + \xi\mathbf{1}_n\mathbf{1}_n'$$

in terms of the parameter $\mathbf{w}$, a matrix $\mathbf{C}$ in $\mathscr{C}(n)$ and a scalar ξ such that $\xi > \mathbf{w}'\mathbf{C}\mathbf{w}$. Sufficiency follows if

$$n^{-1}\mathbf{1}_n' - n^{-1}\mathbf{c}'\mathbf{H}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{H} = \mathbf{w}'.$$

Choose any $\mathbf{H} \in \mathscr{H}$. Then $\mathbf{C} = \mathbf{H}'\mathbf{A}\mathbf{H}$ for some $\mathbf{A} \in S_{(n-1)}^+$. Using this and the properties $\mathbf{H}'\mathbf{H} = \mathbf{I}_{(n-1)}$ and $\mathbf{H}\mathbf{1}_n = \mathbf{0}$ for matrices $\mathbf{H} \in \mathscr{H}$, it follows that

$$(\mathbf{H}\Sigma\mathbf{H}')^{-1} = [\mathbf{H}[\mathbf{C} - \mathbf{C}\mathbf{w}\mathbf{1}_n' - \mathbf{1}_n\mathbf{w}'\mathbf{C}' + \xi\mathbf{1}_n\mathbf{1}_n']\mathbf{H}']^{-1} = (\mathbf{A})^{-1}.$$

Using this, expanding $\mathbf{c}$ as $\mathbf{c}' = \mathbf{1}_n'\Sigma$, and using $\mathbf{H}\mathbf{H}' = [\,\mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n']$ yields

$$n^{-1}\mathbf{1}_n' - n^{-1}\mathbf{c}'\mathbf{H}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{H} = n^{-1}\mathbf{1}_n' - n^{-1}\mathbf{1}_n'\Sigma\mathbf{H}'(\mathbf{A})^{-1}\mathbf{H}$$

$$= n^{-1}\mathbf{1}_n' - n^{-1}\mathbf{1}_n'[\mathbf{C} - \mathbf{C}\mathbf{w}\mathbf{1}_n' - \mathbf{1}_n\mathbf{w}'\mathbf{C}' + \xi\mathbf{1}_n\mathbf{1}_n']\mathbf{H}'(\mathbf{A})^{-1}\mathbf{H}$$

$$= n^{-1}\mathbf{1}_n' - \mathbf{w}'\mathbf{H}'\mathbf{A}\mathbf{H}\mathbf{H}'(\mathbf{A})^{-1}\mathbf{H} = n^{-1}\mathbf{1}_n' - \mathbf{w}'[\,\mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n'] = \mathbf{w}',$$

as required. To verify necessity let

$$\mathbf{w}' = n^{-1}\mathbf{1}_n' - n^{-1}\mathbf{c}'\mathbf{H}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{H}$$

The right hand side of this identity equals $(\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n)^{-1}\mathbf{1}_n'\Sigma^{-1}$ through Theorem 4.2.1. This implies that

$$\mathbf{w}' = (\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n)^{-1}\mathbf{1}_n'\Sigma^{-1}.$$

We need to show that $\Sigma \in \Lambda(\mathbf{w})$ for this vector $\mathbf{w}$. This follows as

$$\mathbf{w}'\Sigma = (\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n)^{-1}\mathbf{1}_n'\Sigma^{-1}\Sigma = (\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n)^{-1}\mathbf{1}_n'$$

is a scale multiple of the unit vector as required for membership in the class $\Lambda(\mathbf{w})$. This ascertains necessity and completes the proof. □

The theorem that follows derives the Pitman estimate whenever $\Sigma \in \Lambda(\mathbf{w})$ for some $\mathbf{w} \in \mathcal{W}(n)$. Example 4.3 illustrates the theorem.

Theorem 4.3.2: For $L(\mathbf{X}) = E_n(\mu\mathbf{1}_n, \Sigma, \phi)$ and each $\mathbf{w} \in \mathcal{W}(n)$, the estimator $\mathbf{w}'\mathbf{X}$ is the MRE estimator for μ if and only if $\Sigma \in \Lambda(\mathbf{w})$.

Proof: The Pitman estimate for $L(\mathbf{X}) = E_n(\mu\mathbf{1}_n, \Sigma, \phi)$ is given as

$$\delta_\Sigma(\mathbf{X}) = \overline{X} - n^{-1}\,\mathbf{c}'\mathbf{H}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{H}\mathbf{X}$$

in Theorem 4.2.1. Through Lemma 4.3.1 this equals $\mathbf{w}'\mathbf{x}$ if and only if $\Sigma \in \Lambda(\mathbf{w})$. □

Example 4.3: MRE estimate of μ for elliptically symmetric distributions with $\Sigma \in \Lambda(\mathbf{w})$ for $\mathbf{w}' = [-1/4 \;\; 3/4 \;\; 1/2]$.

Let $\mathbf{x}' = (10\ 12\ 14)$ be a sample vector from a distribution with $\Sigma = \begin{bmatrix} 12 & 4 & 4 \\ 4 & 4 & 0 \\ 4 & 0 & 6 \end{bmatrix}$.

In Example 3.5 this matrix was seen to belong to $\Lambda(\mathbf{w})$ for the specified $\mathbf{w}$. Hence the estimate is given by

$$\delta_\Sigma(\mathbf{x}) = \mathbf{w}'\mathbf{x} = [-1/4 \;\; 3/4 \;\; 1/2]\begin{bmatrix} 10 \\ 12 \\ 14 \end{bmatrix} = 13.5.$$

We now consider the special subset $\Omega(n)$ amongst the subsets $\Lambda(\mathbf{w})$. This subset consists of positive definite matrices with equal column sums, where $\Omega(n) = \Lambda(n^{-1}\mathbf{1}_n)$.

<u>Corollary 4.3.2:</u> For $L(\mathbf{X}) = E_n(\mu\mathbf{1}_n, \Sigma, \phi)$, the sample average $\overline{X}$ is the MRE estimator for μ if and only if $\Sigma \in \Omega(n)$.

<u>Proof</u>: The result follows from Theorem 4.3.2 as $\Omega(n) = \Lambda(n^{-1}\mathbf{1}_n)$ and $n^{-1}\mathbf{1}_n\mathbf{X} = \overline{X}$. ❑

This corollary demonstrates that membership in $\Omega(n)$ preserves admissibility under squared error loss of $\overline{X}$ amongst equivariant estimators. In Chapter 3 the ill-conditioned matrix in the following example was seen to belong to $\Omega(n)$. The appropriate estimate corresponding to it is the sample average.

<u>Example 4.4:</u> MRE estimate of μ for elliptically contoured distribution with $\Sigma \in \Omega(3)$.

Let $\mathbf{x}' = (10\;12\;14)$ be a sample vector from a distribution with

$$\Sigma = \begin{bmatrix} 3 & 3 & 1 \\ 3 & 5 & -1 \\ 1 & -1 & 7 \end{bmatrix}.$$

In Example 3.7 it was noted that the above matrix belongs to $\Omega(3)$ with column sums equal to 7. Hence $\delta_\Sigma(\mathbf{x}) = \overline{X} = 12$.

The next section considers Pitman estimation for non-elliptical distributions obtained as mixtures of elliptically symmetric distributions.

4.4 PITMAN ESTIMATION: GENERAL RESULTS FOR MIXTURES

This section obtains Pitman estimates for μ for the case of random vectors having mixture distributions. These distributions are translation families of distributions. Densities of such random vectors take the form

$$h(\mathbf{x};\,\mu,\,\phi,\,\mathbf{G}) = \int_{S_n^+} f_\phi\left[(\mathbf{x}-\mu\mathbf{1}_n)'\Sigma^{-1}(\mathbf{x}-\mu\mathbf{1}_n)\right]d\mathbf{G}(\Sigma)$$

where $f_\phi(.)$ is the density of the elliptically symmetric components in the mixture, and $\mathbf{G}$ is a probability measure over S_n^+. The distribution of a random vector $\mathbf{X}$ with a mixture distribution is succinctly specified as $L(\mathbf{X})$ = $EM_n(\mu,\phi,\mathbf{G})$. The Pitman estimates for such distributions are obtained as before using

$$\delta(\mathbf{x};\,\mathbf{D}=\mathbf{d}) = \overline{X} - \mathscr{E}_0\left[\overline{X}|\mathbf{D}=\mathbf{d}\right]$$

where $\mathbf{D}$ = $\mathbf{HX}$ for some matrix $\mathbf{H}$ chosen from the set $\mathscr{H}$ of orthonormal completions of the unit vector; $\overline{X}$ denotes the sample average; and $\mathscr{E}_0\left[\overline{X}|\mathbf{D}=\mathbf{d}\right]$ is the conditional expectation when μ = 0. The lemma that follows provides the general expression for $\mathscr{E}_0\left[\overline{X}|\mathbf{D}=\mathbf{d}\right]$ for densities in the location family. This result applies to succeeding developments.

<u>Lemma 4.4.1</u>: Consider a random vector $\mathbf{X}$ with pdf taking the form $f(\mathbf{x}-\mu\mathbf{1}_n)$. Fix $\mathbf{H} \in \mathscr{H}$ and consider $\mathbf{D}$ = $\mathbf{HX}$. Then

$$\mathscr{E}_0\left[\overline{X}|\mathbf{D}=\mathbf{d}\right] = \frac{\int_{-\infty}^{\infty} t\sqrt{n}f\left[\mathbf{K}'\begin{bmatrix} t\sqrt{n} \\ \mathbf{d}\end{bmatrix}\right]dt}{h(\mathbf{d})}$$

where $\mathbf{K} = \begin{bmatrix}(1/\sqrt{n})\mathbf{1}_n' \\ \mathbf{H}\end{bmatrix}$ and $h(\mathbf{d})$ is the density of $\mathbf{D}$.

Proof: When $\mu = 0$, the density $f(\mathbf{x}-\mu\mathbf{1}_n)$ takes the form $f(\mathbf{x})$. To compute $\mathscr{E}_0\left[\overline{X}\middle|\mathbf{D}=\mathbf{d}\right]$ consider the following sequence of change of variables. Firstly, let

$$\mathbf{y}^* = \begin{bmatrix} y_1^* \\ \mathbf{y}_2^* \end{bmatrix} = \begin{bmatrix} (1/\sqrt{n})\mathbf{1}_n' \\ \mathbf{H} \end{bmatrix}\mathbf{x} = \mathbf{K}\mathbf{x}$$

for the matrix $\mathbf{H}$ chosen from $\mathscr{H}$. The matrix $\mathbf{K}$ is orthogonal and this gives the joint density of the transformed variables as

$$g(\mathbf{y}^*) = |K|^{-1}\, f[\mathbf{K}'\mathbf{y}^*] = f[\mathbf{K}'\mathbf{y}^*]$$

Next consider $\mathbf{d} = \mathbf{y}_2^* = \mathbf{H}\mathbf{x}$ and $y_1 = (1/\sqrt{n})y_1^* = \overline{x}$. This yields

$$h(y_1,\mathbf{d}) = \sqrt{n}g\begin{bmatrix} y_1\sqrt{n} \\ \mathbf{d} \end{bmatrix} = \sqrt{n}f\left[\mathbf{K}'\begin{bmatrix} y_1\sqrt{n} \\ \mathbf{d} \end{bmatrix}\right].$$

Thus the conditional density of Y_1 given $\mathbf{D} = \mathbf{d}$ is

$$h(y_1|\mathbf{d}) = \frac{\sqrt{n}f\left[K'\begin{bmatrix} y_1\sqrt{n} \\ \mathbf{d} \end{bmatrix}\right]}{h(\mathbf{d})}$$

Since $y_1 = \overline{x}$, it follows that

$$\mathscr{E}_0\left[\overline{X}\middle|\mathbf{D}=\mathbf{d}\right] = \mathscr{E}_0\left[Y_1\middle|\mathbf{D}=\mathbf{d}\right] = \frac{\int_{-\infty}^{\infty} t\sqrt{n}f\left[\mathbf{K}'\begin{bmatrix} t\sqrt{n} \\ \mathbf{d} \end{bmatrix}\right]dt}{h(\mathbf{d})}$$

as required. ❑

We now apply the preceding lemma to elliptically symmetric distributions with densities. For these distributions the conditional expectation $\mathscr{E}_0\left[\overline{X}\middle|\mathbf{D}=\mathbf{d}\right]$ is denoted as $\mathscr{E}_\Sigma\left[\overline{X}\middle|\mathbf{d}\right]$. For $\mathbf{H} \in \mathscr{H}$ the distribution of the random vector $\mathbf{D} = \mathbf{HX}$ for $L(\mathbf{X}) = E_n(\mu\mathbf{1}_n, \Sigma, \phi)$ is given by $L(\mathbf{D}) = E_{n-1}\left(\mu\mathbf{H}\mathbf{1}_n, \mathbf{H}\Sigma\mathbf{H}', \phi\right) = E_{n-1}\left(\mathbf{0}, \mathbf{H}\Sigma\mathbf{H}', \phi\right)$. The density of $\mathbf{D}$ is given by

$$m(\mathbf{d}; \Sigma, \phi) = g_\phi[\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d}].$$

The identity in the following lemma will be applied to demonstrate Theorem 4.4.1.

Lemma 4.4.2: For $L(\mathbf{X}) = \mathrm{E}_n(\mu\mathbf{1}_n, \Sigma, \phi)$ having density

$$f_\phi\left[(\mathbf{x}-\mu\mathbf{1}_n)'\Sigma^{-1}(\mathbf{x}-\mu\mathbf{1}_n)\right],$$

fix $\mathbf{H} \in \mathscr{H}$, and let $\mathbf{K}$ and $\mathbf{D}$ be as before. Then

$$\mathscr{E}_\Sigma\left[\overline{\mathrm{X}}\middle|\mathbf{d}\right] g_\phi[\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d}] = \int_{-\infty}^{\infty} t\sqrt{n} f_\phi\left[\begin{bmatrix} t\sqrt{n} & \mathbf{d}' \end{bmatrix}\mathbf{K}\Sigma^{-1}\mathbf{K}'\begin{bmatrix} t\sqrt{n} \\ \mathbf{d} \end{bmatrix}\right] dt.$$

Proof: Applying Lemma 4.4.1 note that

$$\mathscr{E}_\Sigma\left[\overline{\mathrm{X}}\middle|\mathbf{d}\right] = \frac{\int_{-\infty}^{\infty} t\sqrt{n} f_\phi\left[\begin{bmatrix} t\sqrt{n} & \mathbf{d}' \end{bmatrix}\mathbf{K}\Sigma^{-1}\mathbf{K}'\begin{bmatrix} t\sqrt{n} \\ \mathbf{d} \end{bmatrix}\right] dt}{g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d})}.$$

This leads to the required identity. ❑

In Theorem 4.4.1 that follows the preceding lemmas are applied for mixtures of the type $L(\mathbf{X}) = \mathrm{EM}_n(\mu, \phi, \mathbf{G})$. Here the conditional expectation $\mathscr{E}_0\left[\overline{\mathrm{X}}\middle|\mathbf{D}=\mathbf{d}\right]$ in Lemma 4.4.1 is denoted as $\mathscr{E}_{\mathbf{G},\phi}\left[\overline{\mathrm{X}}\middle|\mathbf{d}\right]$. The density of $\mathbf{D} = \mathbf{HX}$ for $\mathbf{H} \in \mathscr{H}$ is obtained as

$$b(\mathbf{d}; \mathbf{G}, \phi) = \int_{S_n^+} g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d})d\mathbf{G}(\Sigma),$$

where $g_\phi(.)$ is the density of $\mathbf{D}$ when $L(\mathbf{X}) = \mathrm{E}_n(\mu\mathbf{1}_n, \Sigma, \phi)$. The Pitman estimate for the random vector with the mixture distribution is denoted as $\delta_{\mathbf{G},\phi}(\mathbf{x};\mathbf{d})$. The subscripts indicate the parameters through which the estimate depends on $\mathbf{x}$ and $\mathbf{d}$. For each elliptically symmetric distribution in the mixture $\delta_\Sigma(\mathbf{x})$ denotes the Pitman estimate. This was seen to depend on $\mathbf{x}$ through Σ in Section 4.2. The following theorem expresses $\delta_{\mathbf{G},\phi}(\mathbf{x};\mathbf{d})$ in terms of $b(.)$, $g_\phi(.)$ and $\delta_\Sigma(\mathbf{x})$.

Theorem 4.4.1: Fix $\mathbf{H} \in \mathscr{H}$ and let $\mathbf{D} = \mathbf{HX}$ for this $\mathbf{H}$. Then for $L(\mathbf{X}) = \mathrm{EM}_n(\mu, \phi, \mathbf{G})$ the Pitman estimate is given by

$$\delta_{\mathbf{G},\phi}(\mathbf{x};\mathbf{d}) = \frac{\int_{S_n^+} \delta_\Sigma(\mathbf{x}) g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d}) d\mathbf{G}(\Sigma)}{b(\mathbf{d};\mathbf{G},\phi)}.$$

Proof: By using Lemma 4.4.1, interchanging integrals through Fubini's Theorem, and then applying Lemma 4.4.2, the conditional expectation works out to

$$\mathscr{E}_{\mathbf{G},\phi}\left[\overline{X}|\mathbf{d}\right] = \frac{\int_{-\infty}^{\infty} t\sqrt{n} \int_{S_n^+} f_\phi\left[\left[t\sqrt{n} \quad \mathbf{d}'\right]\mathbf{K}\Sigma^{-1}\mathbf{K}'\begin{bmatrix} t\sqrt{n} \\ \mathbf{d} \end{bmatrix}\right] d\mathbf{G}(\Sigma) dt}{b(\mathbf{d};\mathbf{G},\phi)}$$

$$= \frac{\int_{S_n^+} \mathscr{E}_\Sigma\left[\overline{X}|\mathbf{d}\right] g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d}) d\mathbf{G}(\Sigma)}{b(\mathbf{d};\mathbf{G},\phi)}.$$

Now using $\delta_{\mathbf{G},\phi}(\mathbf{x};\mathbf{d}) = \overline{X} - \mathscr{E}_{\mathbf{G},\phi}\left[\overline{X}|\mathbf{d}\right]$ we get

$$\delta_{\mathbf{G},\phi}(\mathbf{x};\mathbf{d}) = \overline{X} - \frac{\int_{S_n^+} \mathscr{E}_\Sigma\left[\overline{X}|\mathbf{d}\right] g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d}) d\mathbf{G}(\Sigma)}{b(\mathbf{d};\mathbf{G},\phi)}$$

$$= \frac{\int_{S_n^+} (\overline{X} - \mathscr{E}_\Sigma\left[\overline{X}|\mathbf{d}\right]) g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d}) d\mathbf{G}(\Sigma)}{b(\mathbf{d};\mathbf{G},\phi)}$$

$$= \frac{\int_{S_n^+} \delta_\Sigma(\mathbf{x}) g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d}) d\mathbf{G}(\Sigma)}{b(\mathbf{d};\mathbf{G},\phi)}$$

as required. ❑

It should be noted that the Pitman estimate $\delta_{\mathbf{G},\phi}(\mathbf{x};\mathbf{d})$ can depend on the observed value of the maximal invariant. Corollary 4.4.1 clarifies this. The estimate takes the form $\mathbf{k}'\mathbf{x}$, where the vector of coefficients $\mathbf{k}$ is a function of $\mathbf{d}$ through the parameters ϕ and $\mathbf{G}$. Further, as for the Pitman estimate for elliptically symmetric components (denoted by $\mathbf{k}_\Sigma$), it is seen that $\mathbf{1}_n'\mathbf{k} = 1$, implying that the estimate is an unbiased weighted average.

Corollary 4.4.1: Let $L(\mathbf{X}) = \mathrm{EM}_n(\mu,\phi,\mathbf{G})$. Fix $\mathbf{H} \in \mathscr{H}$ and let $\mathbf{D} = \mathbf{HX}$. Then the Pitman estimate is an unbiased weighted average of the form $\mathbf{k'x}$ with

$$\mathbf{k} = \mathbf{k}(\mathbf{d},\mathbf{G},\phi) = \frac{\int_{S_n^+} \mathbf{k}_\Sigma g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d})d\mathbf{G}(\Sigma)}{b(\mathbf{d};\mathbf{G},\phi)}.$$

Proof: Note that

$$\begin{aligned}
\delta_{\mathbf{G},\phi}(\mathbf{x};\mathbf{d}) &= \frac{\int_{S_n^+} \delta_\Sigma(\mathbf{x}) g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d})d\mathbf{G}(\Sigma)}{b(\mathbf{d};\mathbf{G},\phi)} \\
&= \frac{\int_{S_n^+} \mathbf{x}'\mathbf{k}_\Sigma g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d})d\mathbf{G}(\Sigma)}{b(\mathbf{d};\mathbf{G},\phi)} \\
&= \frac{\mathbf{x}'\int_{S_n^+} \mathbf{k}_\Sigma g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d})d\mathbf{G}(\Sigma)}{b(\mathbf{d};\mathbf{G},\phi)} = \mathbf{x}'\mathbf{k},
\end{aligned}$$

for $\mathbf{k}$ depending on $\mathbf{d}$, $\mathbf{G}$, and ϕ as required. To show that the Pitman estimate for mixture distributions is a weighted average note that $\mathbf{1}_n'\mathbf{k}_\Sigma = \mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n(\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n)^{-1} = 1$ for all Σ. Hence

$$\begin{aligned}
\mathbf{1}_n'\mathbf{k}(\mathbf{d},\mathbf{G},\phi) &= \frac{\mathbf{1}_n'\int_{S_n^+} \mathbf{k}_\Sigma g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d})d\mathbf{G}(\Sigma)}{b(\mathbf{d};\mathbf{G},\phi)} \\
&= \frac{\int_{S_n^+} (\mathbf{1}_n'\mathbf{k}_\Sigma) g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d})d\mathbf{G}(\Sigma)}{b(\mathbf{d};\mathbf{G},\phi)} = 1.
\end{aligned}$$

Unbiasedness follows as the expected value of each component of $\mathbf{X}$ is μ, and hence the expected value of the weighted average is μ. ❑

Theorem 4.4.1 and the preceding corollary illustrate that the Pitman estimate, in general, can depend on the maximal invariant through ϕ and $\mathbf{G}$. This is illustrated in Example 4.5 for a 50% mixture of two bivariate normal distributions as in Figure 4.1. The sample vectors $\mathbf{x}_1' = \begin{bmatrix}5 & 1\end{bmatrix}$, and $\mathbf{x}_2' = \begin{bmatrix}5 & 4\end{bmatrix}$ are considered. For $\mathbf{H} = (1/\sqrt{2})\begin{bmatrix}1 & -1\end{bmatrix}$

the function $\mathbf{d} = \mathbf{Hx}$ is chosen as the maximal invariant under translations. Since $\mathbf{d}_1 = \mathbf{Hx}_1 = 4/\sqrt{2}$ differs from $\mathbf{d}_2 = \mathbf{Hx}_2 = 1/\sqrt{2}$, the estimating linear form differs for the two sample vectors.

Example 4.5 : Illustration of non-unique linear forms for the Pitman estimator.

Consider the sample vectors $\mathbf{x}_1' = [\,5,1\,]$, and $\mathbf{x}_2' = [\,5,4\,]$ from a distribution which is a 50% mixture of bivariate normals with $\Sigma_1 = \begin{bmatrix} 5 & -4.5 \\ -4.5 & 5 \end{bmatrix}$ and $\Sigma_2 = \begin{bmatrix} 11 & 7 \\ 7 & 5 \end{bmatrix}$. The vectors of coefficients $\mathbf{k}(\mathbf{d}_1,\mathbf{G},\phi)$ and $\mathbf{k}(\mathbf{d}_2,\mathbf{G},\phi)$ which yield the Pitman estimators will be computed by (i) computing the vector of coefficients $\mathbf{k}_{\Sigma_1}$ and $\mathbf{k}_{\Sigma_2}$ for each of the Gaussian components, (ii) calculating $g_\phi(\mathbf{d}_1'(\mathbf{H}\Sigma_1\mathbf{H}')^{-1}\mathbf{d}_1)$ and $g_\phi(\mathbf{d}_1'(\mathbf{H}\Sigma_2\mathbf{H}')^{-1}\mathbf{d}_1)$ to compute $\mathbf{k}(\mathbf{d}_1,\mathbf{G},\phi)$, and (iii) calculating $g_\phi(\mathbf{d}_2'(\mathbf{H}\Sigma_1\mathbf{H}')^{-1}\mathbf{d}_2)$ and $g_\phi(\mathbf{d}_2'(\mathbf{H}\Sigma_2\mathbf{H}')^{-1}\mathbf{d}_2)$ to compute $\mathbf{k}(\mathbf{d}_2,\mathbf{G},\phi)$. First we calculate $\mathbf{k}_{\Sigma_1}$ and $\mathbf{k}_{\Sigma_2}$. The vector of coefficients for Σ_1 is $\mathbf{k}_{\Sigma_1}' = \begin{bmatrix} 1/2 & 1/2 \end{bmatrix}$ as $\Sigma_1 \in \Omega(n)$. Further

$$\mathbf{k}_{\Sigma_2}' = \left[\left[\mathbf{1}_n'\Sigma_2^{-1}\mathbf{1}_n\right]^{-1}\mathbf{1}_n'\Sigma_2^{-1}\right]' = \begin{bmatrix} -1 & 2 \end{bmatrix}.$$

The density $g_\phi(\mathbf{d}'(\mathbf{H}\Sigma_1\mathbf{H}')^{-1}\mathbf{d})$ is univariate normal with variance

$$\sigma^2 = \mathbf{H}\Sigma_1\mathbf{H}' = 1/2\begin{bmatrix} 1 & -1 \end{bmatrix}\begin{bmatrix} 5 & -4.5 \\ -4.5 & 5 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = 9.5.$$

Evaluating this univariate normal at $\mathbf{d}_1 = \mathbf{Hx}_1 = 4/\sqrt{2}$ we obtain

$$g_\phi(\mathbf{d}_1'(\mathbf{H}\Sigma_1\mathbf{H}')^{-1}\mathbf{d}_1) = \frac{1}{\sqrt{2\Pi}.\sqrt{9.5}} e^{-1/2\frac{(4/\sqrt{2})^2}{9.5}} = 0.084955.$$

Similarly $g_\phi(\mathbf{d}_1'(\mathbf{H}\Sigma_2\mathbf{H}')^{-1}\mathbf{d}_1)$ equals the univariate normal density with variance

$$\sigma^2 = \mathbf{H}\Sigma_2\mathbf{H}' = 1/2\begin{bmatrix} 1 & -1 \end{bmatrix}\begin{bmatrix} 11 & 7 \\ 7 & 5 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = 1$$

Hence

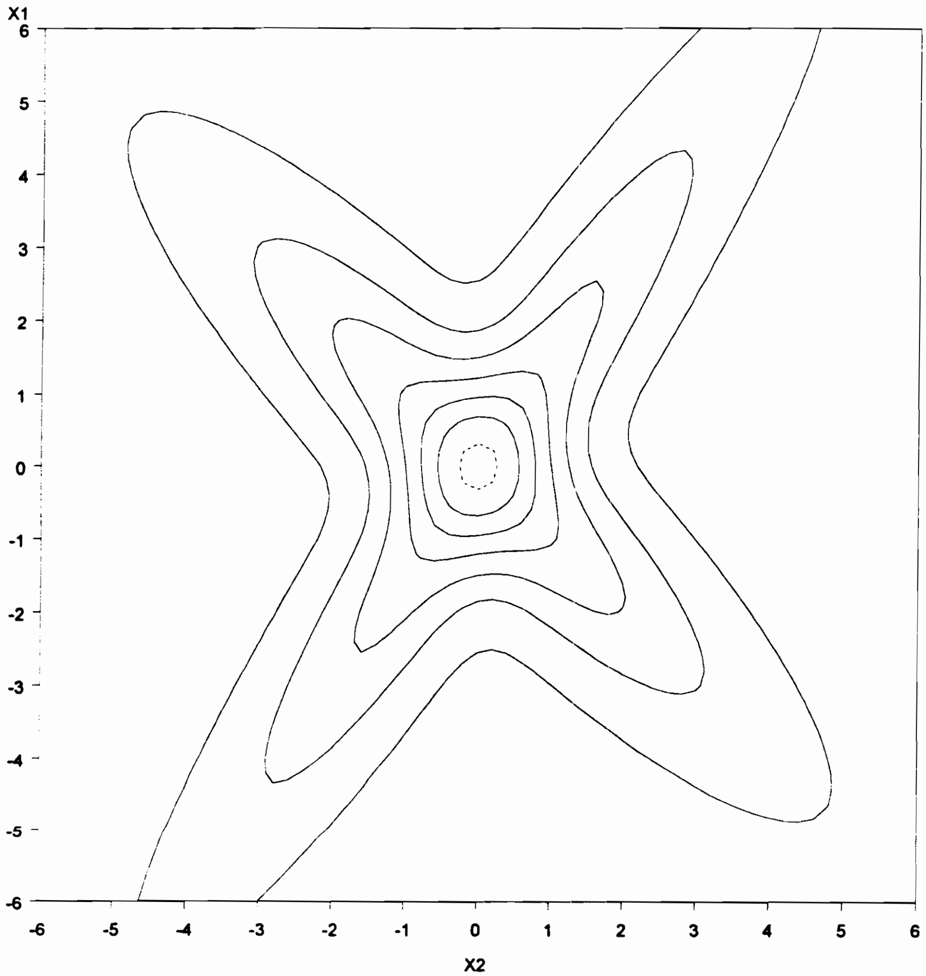

Figure 4.1: Mixture of bivariate normals with **G** assigning equal probability to $\Sigma_1 = \begin{bmatrix} 5 & -45 \\ -4.5 & 5 \end{bmatrix}$ and to $\Sigma_2 = \begin{bmatrix} 11 & 7 \\ 7 & 5 \end{bmatrix}$.

$$g_\phi(\mathbf{d}_1'(\mathbf{H}\Sigma_2\mathbf{H}')^{-1}\mathbf{d}_1) = \frac{1}{\sqrt{2\Pi.1}}e^{-1/2\frac{(4/\sqrt{2})^2}{1}} = 0.007307.$$

Using these computations

$$\mathbf{k}'(\mathbf{d}_1,\mathbf{G},\phi) = \frac{0.5g_\phi(\mathbf{d}_1'(\mathbf{H}\Sigma_1\mathbf{H}')^{-1}\mathbf{d}_1)\mathbf{k}'_{\Sigma_1} + 0.5g_\phi(\mathbf{d}_1'(\mathbf{H}\Sigma_2\mathbf{H}')^{-1}\mathbf{d}_1)\mathbf{k}'_{\Sigma_2}}{0.5g_\phi(\mathbf{d}_1'(\mathbf{H}\Sigma_1\mathbf{H}')^{-1}\mathbf{d}_1) + 0.5g_\phi(\mathbf{d}_1'(\mathbf{H}\Sigma_2\mathbf{H}')^{-1}\mathbf{d}_1))}$$

$$= \frac{0.5\times 0.084955\mathbf{k}'_{\Sigma_1} + 0.5\times 0.007307\mathbf{k}'_{\Sigma_2}}{0.5\times 0.084955 + 0.5\times 0.007307} = \begin{bmatrix}0.3813 & 0.6187\end{bmatrix}$$

This vector of coefficients is closer to $\mathbf{k}_{\Sigma_1}$ than it is to $\mathbf{k}_{\Sigma_2}$. Using this vector

$$\delta_{\mathbf{G},\phi}(\mathbf{x};\mathbf{d}_1) = [\mathbf{k}(\mathbf{d}_1,\mathbf{G},\phi)]'\mathbf{x}_1 = \begin{bmatrix}0.3813 & 0.6187\end{bmatrix}\begin{bmatrix}5\\1\end{bmatrix} = 2.524.$$

Similar calculations for $\mathbf{x}_2 = [\ 5,4\ ]$ using $\mathbf{d}_2 = \mathbf{h}\mathbf{x}_2 = 1/\sqrt{2}$ yield

$$g_\phi(\mathbf{d}_2'(\mathbf{H}\Sigma_1\mathbf{H}')^{-1}\mathbf{d}_2) = 0.1261, \text{ and } g_\phi(\mathbf{d}_2'(\mathbf{H}\Sigma_2\mathbf{H}')^{-1}\mathbf{d}_2) = 0.3107.$$

Hence

$$\mathbf{k}'(\mathbf{d}_2,\mathbf{G},\phi) = \frac{0.5\times 0.1261\mathbf{k}'_{\Sigma_1} + 0.5\times 0.3107\mathbf{k}'_{\Sigma_2}}{0.5\times 0.1261 + 0.5\times 0.3107} = \begin{bmatrix}-0.567 & 1.567\end{bmatrix}.$$

Note that this vector of coefficients is closer to $\mathbf{k}_{\Sigma_2}$ than it is to $\mathbf{k}_{\Sigma_1}$. Using this vector

$$\delta_{\mathbf{G},\phi}(\mathbf{x};\mathbf{d}_2) = \mathbf{k}'(\mathbf{d}_2,\mathbf{G},\phi)\mathbf{x}_2 = \begin{bmatrix}-0.567 & 1.567\end{bmatrix}\begin{bmatrix}5\\4\end{bmatrix} = 3.433.$$

The preceding example illustrates that when $\mathbf{G}$ is not restricted then the estimate does depend on a maximal invariant function of the data through ϕ and $\mathbf{G}$. In the following section mixture distributions with restrictions on the probability measure $\mathbf{G}$ over S_n^+ are studied. For random vectors with these distributions the Pitman estimate depends solely on a parameter Ψ which is proportional to the variance-covariance matrix for the random vector having the mixture distribution.

4.5 PITMAN ESTIMATION FOR MIXTURES: RESTRICTIONS ON THE DOMAIN OF G TO THE CLASSES $\aleph_C$ AND $\Lambda(\mathbf{w})$

For $L(\mathbf{X}) = EM_n(\mu,\phi,\mathbf{G})$ we now examine Pitman estimation for μ when the domain of the probability measure **G** is restricted to the subsets $\aleph_C$ and $\Lambda(\mathbf{w})$. The domain of **G** is denoted as Dom(**G**). For each vector **w** such that $\mathbf{1}_n'\mathbf{w} = 1$, the subset $\Lambda(\mathbf{w})$ consists of positive definite matrices such that $\mathbf{w}'\Sigma$ is proportional to the unit vector. The index set for the subsets $\aleph_C$ are real symmetric matrices of order n with a one-dimensional null space spanned by the unit vector. It was shown in Section 3.3 that matrices in $\aleph_C$ can be expressed as

$$\Sigma(\mathbf{C},\gamma) = \mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\mathbf{1}_n\mathbf{1}_n'$$

for a vector γ of constants such that $\gamma'\mathbf{C}^{-1}_{\gamma(n)}\gamma < \bar{\gamma}$. In those contexts where Dom(**G**) is restricted, the Pitman estimate depends in a simpler manner on a parameter Ψ obtained as

$$\Psi = \int_{S_n^+} \Sigma d\mathbf{G}(\Sigma).$$

The matrix Ψ is positive definite and is proportional to Var(**X**). Lemma 4.5.1 illustrates this.

Lemma 4.5.1: (a) For a probability measure **G** over S_n^+, the matrix $\Psi = \int_{S_n^+} \Sigma d\mathbf{G}(\Sigma)$ is positive definite.

(b) For $L(\mathbf{X}) = EM_n(\mu,\phi,\mathbf{G})$, $\mathrm{Var}(\mathbf{X}) = \alpha_\phi\Psi$, where α_ϕ is a constant depending on ϕ.

Proof: To prove (a) choose any $\mathbf{a} \in \Re^n$. Since each $\Sigma \in$ Dom(**G**) is a member of S_n^+ it follows that $\mathbf{a}'\Sigma\mathbf{a} > 0$ for all Σ. Hence

$$\mathbf{a}'\Psi\mathbf{a} = \mathbf{a}'\left[\int_{S_n^+} \Sigma d\mathbf{G}(\Sigma)\right]\mathbf{a} = \int_{S_n^+} (\mathbf{a}'\Sigma\mathbf{a})d\mathbf{G}(\Sigma) > 0.$$

This implies $\Psi \in S_n^+$.

To prove (b) write Var(**X**) as

$$\mathrm{Var}(\mathbf{X}) = \mathrm{Var}_G\{\mathscr{E}[\,\mathbf{X} \mid \Sigma\,]\} + \mathscr{E}_G\{\mathrm{Var}\,[\,\mathbf{X} \mid \Sigma\,]\}.$$

Since $\mathscr{E}[\,\mathbf{X} \mid \Sigma\,] = \mu \mathbf{1}_n$ for each Σ, $\mathrm{Var}_G\{\mathscr{E}[\,\mathbf{X} \mid \Sigma\,]\} = 0$. This leads to

$$\mathrm{Var}(\mathbf{X}) = \mathscr{E}_G\{\mathrm{Var}\,[\,\mathbf{X} \mid \Sigma\,]\} = \alpha_\phi \int_{S_n^+} \Sigma dG(\Sigma) = \alpha_\phi \Psi,$$

as required. ❑

Recall that for elliptically symmetric distributions where $\Sigma \in \Lambda(\mathbf{w})$ for some $\mathbf{w} \in \mathscr{W}(n)$, then the Pitman estimate is **w'x** for this **w**. A similar result follows for mixtures. When Dom(**G**) = $\Lambda(\mathbf{w})$ for some **w**, then $\Psi \in \Lambda(\mathbf{w})$ and **w'x** is the Pitman estimate for location.

<u>Theorem 4.5.1</u>: (a) When Dom(G) = $\Lambda(\mathbf{w})$ for some $\mathbf{w} \in \mathscr{W}(n)$, then $\Psi \in \Lambda(\mathbf{w})$ for the chosen vector.

(b) When $L(\mathbf{X}) = \mathrm{EM}_n(\mu, \phi, \mathbf{G})$ with Dom(**G**) = $\Lambda(\mathbf{w})$, then the Pitman estimate for μ is **w'x** for this **w**.

<u>Proof</u>: Since Dom(**G**) = $\Lambda(\mathbf{w})$, $\mathbf{w}'\Sigma = c_\Sigma \mathbf{1}_n'$ for all $\Sigma \in$ Dom(**G**). Using this note that

$$\mathbf{w}'\Psi = \mathbf{w}' \int_{\Lambda(\mathbf{w})} \Sigma d\mathbf{G}(\Sigma) = \int_{\Lambda(\mathbf{w})} (\mathbf{w}'\Sigma) d\mathbf{G}(\Sigma) = \mathbf{1}_n' \int_{\Lambda(\mathbf{w})} c_\Sigma d\mathbf{G}(\Sigma)$$

which is a scale multiple of the unit vector. This together with positive definiteness from Lemma 4.5.1(a), implies that $\Psi \in \Lambda(\mathbf{w})$.

The proof to (b) follows on applying Theorem 4.4.1 and noting that the Pitman estimate for μ for each elliptically symmetric component is **w'x** independent of $\Sigma \in \Lambda(\mathbf{w})$. This follows from

$$\delta_{G,\phi}(\mathbf{x};\mathbf{d}) = \frac{\int_{S_n^+} \delta_\Sigma(\mathbf{x}) g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d}) d\mathbf{G}(\Sigma)}{b(\mathbf{d};\mathbf{G},\phi)}$$

$$= \frac{\mathbf{w}'\mathbf{x}\int_{S_n^+} g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d}) d\mathbf{G}(\Sigma)}{b(\mathbf{d};\mathbf{G},\phi)} = \mathbf{w}'\mathbf{x}.$$

❑

In Figure 3.4, the $\Lambda(\mathbf{w})$ subsets are depicted as the vertical strips. From Theorem 4.5.1 we conclude that arbitrary probability measures over a strip $\Lambda(\mathbf{w})$ generate mixture distributions for which **w'x** is Pitman. The following corollary applies Theorem 4.5.1 to obtain a class of mixture distributions for which the sample average is the Pitman estimate. These distributions are those for which Dom(**G**) is the subset $\Omega(n)$ of S_n^+, containing positive definite matrices which have equal column sums.

<u>Corollary 4.5.1</u>: (a) When Dom(**G**) = $\Omega(n)$, $\Psi \in \Omega(n)$.

(b) When $L(\mathbf{X}) = EM_n(\mu,\phi,\mathbf{G})$ with Dom(**G**) = $\Omega(n)$, then the sample average $\overline{X}$ is the Pitman estimate for μ.

<u>Proof</u>: Results in this Corollary follow from Theorem 4.5.1 as $\Omega(n) = \Lambda(n^{-1}\mathbf{1}_n)$. ❑

In Figure 3.4, $\Omega(n)$ is the central vertical strip. The preceding corollary concludes that arbitrary probability measures **G** over this strip preserve $\overline{X}$ as the MRE estimate under squared error. Figure 4.2 depicts a mixture of two bivariate normals each having a dispersion matrix in $\Omega(2)$. Note the symmetry of the density. In Section 4.3 elements of $\Omega(n)$ were seen to have an eigenvector equal to the scaled unit vector, which accounts for the symmetry in Figure 4.2. The Pitman estimate is shown to be the sample average in the following example.

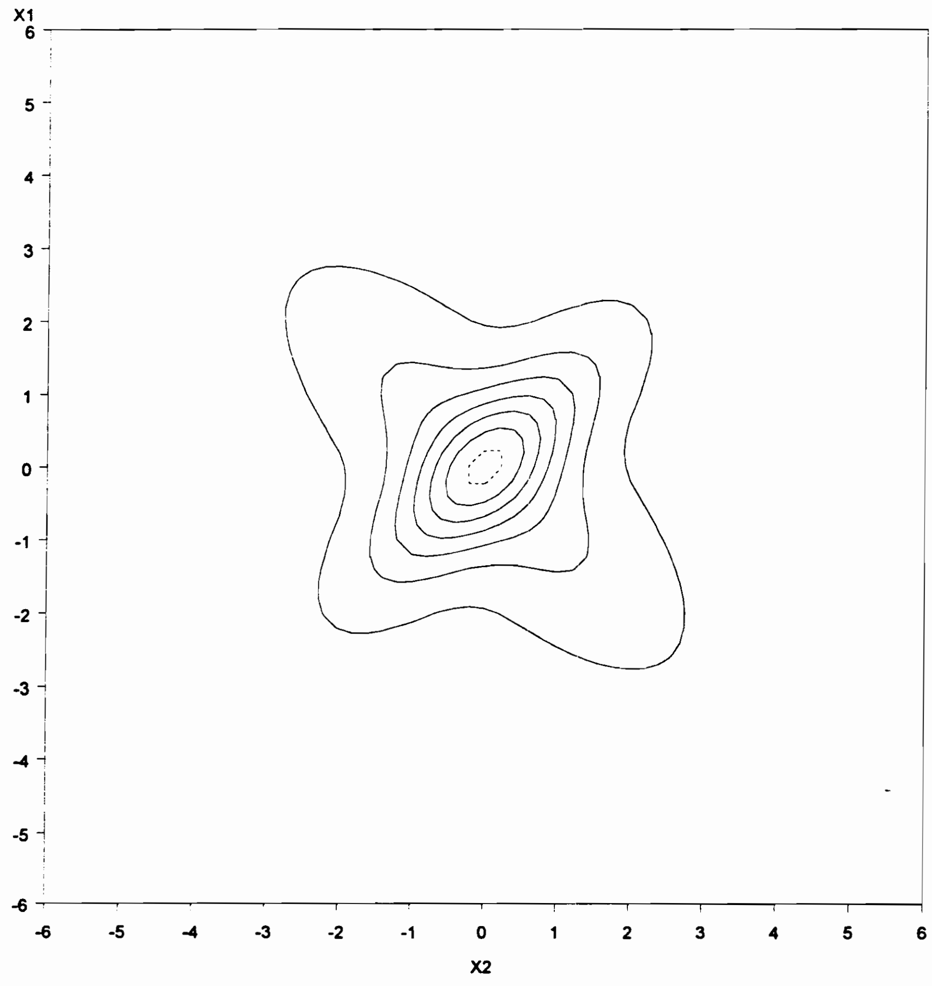

Figure 4.2: Mixture of bivariate normals with Dom(**G**) = $\Omega(2)$ with **G** assigning 0.5 probability to $\Sigma_1 = \begin{bmatrix} 2 & -1.5 \\ -1.5 & 2 \end{bmatrix}$ and to $\Sigma_2 = \begin{bmatrix} 1 & 0.75 \\ 0.75 & 1 \end{bmatrix}$.

Example 4.6: MRE estimation for mixtures with Dom($\mathbf{G}$) = $\Omega(2)$.

Let $\Sigma_1 = \begin{bmatrix} 2 & -1.5 \\ -1.5 & 2 \end{bmatrix}$ and $\Sigma_2 = \begin{bmatrix} 1 & 0.75 \\ 0.75 & 1 \end{bmatrix}$, and let $\mathbf{G}$ be a c.d.f. which assigns equal probability of 0.5 to each of these. Figure 4.2 is the mixture density when two bivariate normal distributions with these dispersion matrices are used. If $\mathbf{x}' = (5, 4)$ were the sample vector, then $\delta_\Psi(\mathbf{x}) = \overline{X} = 4.5$ is the MRE estimator for μ.

Recall that for $L(\mathbf{X}) = E_n(\mu\mathbf{1}_n, \Sigma, \phi)$, with $\Sigma \in \aleph_\mathbf{C}$ for some $\mathbf{C} \in \mathscr{C}(n)$, the Pitman estimate for μ is $\delta_\Sigma(\mathbf{x}) = \overline{X} - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\mathbf{x}$ for γ as in the expansion

$$\Sigma = \mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\mathbf{1}_n\mathbf{1}_n'.$$

Theorem 4.5.2 provides a result similar to this for mixtures in terms of the parameter Ψ for the mixture distributions. Further necessary and sufficient conditions are provided for each estimate $\mathbf{w}'\mathbf{x}$ to be Pitman.

Theorem 4.5.2: (a) When Dom($\mathbf{G}$) = $\aleph_\mathbf{C}$ for some $\mathbf{C} \in \mathscr{C}(n)$, then $\Psi \in \aleph_\mathbf{C}$.

(b) For $\mathbf{G}$ as in (a) let $L(\mathbf{X}) = EM_n(\mu, \phi, \mathbf{G})$. Then

$$\Psi = \mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\mathbf{1}_n\mathbf{1}_n'$$

for some vector γ of constants such that $\gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\gamma < \bar{\gamma}$. The Pitman estimate is given through this vector as $\delta_\Psi(x) = \overline{X} - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\mathbf{x}$.

(c) For $L(\mathbf{X}) = EM_n(\mu, \phi, \mathbf{G})$ with Dom($\mathbf{G}$) = $\aleph_\mathbf{C}$ for some $\mathbf{C} \in \mathscr{C}(n)$, $\mathbf{w}'\mathbf{x}$ is the Pitman estimate if and only if $\Psi \in \Lambda(\mathbf{w})$.

(d) For a random vector with a distribution as in (c), the sample average $\overline{X}$ is Pitman if and only if $\Psi \in \Omega(n)$.

Proof: (a) For $\mathbf{C} \in \mathscr{C}(n)$ let $(\mathbf{A}, \mathbf{H})$ be chosen from $\{S^+_{(n-1)}, \mathscr{H}\}$ such that $\mathbf{H}'\mathbf{A}\mathbf{H} = \mathbf{C}$ in this and succeeding parts of this proof. Then for all $\Sigma \in \aleph_\mathbf{C}$, $\mathbf{H}\Sigma\mathbf{H}' = \mathbf{A}$. Using this note that

$$\mathbf{H\Psi H'} = \mathbf{H}\left[\int_{\aleph_C} \Sigma d\mathbf{G}(\Sigma)\right]\mathbf{H'} = \int_{\aleph_C}(\mathbf{H}\Sigma\mathbf{H'})d\mathbf{G}(\Sigma) = \mathbf{A}.$$

Since $\mathbf{H\Psi H'} = \mathbf{A}$ for the chosen $(\mathbf{A}, \mathbf{H}) \in \{S^{+}_{(n-1)}, \mathscr{H}\}$ and since Ψ is positive definite, $\Psi \in \aleph_C$.

The first part of (b) follows from the characterization of matrices in $\aleph_C$. To show the second part note (i) that the vector γ in the expansion for Ψ can be expressed in terms of the vectors γ_Σ in the expansion for each Σ as $\gamma = \int_{\aleph_C}\gamma_\Sigma d\mathbf{G}(\Sigma)$, since

$$\int_{\aleph_C}\Sigma d\mathbf{G}(\Sigma) = \int_{\aleph_C}(\mathbf{C}+\gamma_\Sigma\mathbf{1_n}'+\mathbf{1_n}\gamma_\Sigma'-\bar{\gamma}_\Sigma\mathbf{1_n}\mathbf{1_n}')d\mathbf{G}(\Sigma),$$

and (ii) that the density $g_\phi[\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d}] = g_\phi[\mathbf{d}'(\mathbf{A})^{-1}\mathbf{d}]$ invariantly for all $\Sigma \in \aleph_C$. Using this we get

$$\begin{aligned}
\delta_{\mathbf{G},\phi}(\mathbf{x};\mathbf{d}) &= \frac{\int_{\aleph_C}\delta_\Sigma(\mathbf{x})g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d})d\mathbf{G}(\Sigma)}{b(\mathbf{d};\mathbf{G},\phi)} \\
&= \frac{\int_{\aleph_C}\delta_\Sigma(\mathbf{x})g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d})d\mathbf{G}(\Sigma)}{\int_{\aleph_C}g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d})d\mathbf{G}(\Sigma)} \\
&= \frac{\int_{\aleph_C}\delta_\Sigma(\mathbf{x})g_\phi(\mathbf{d}'(\mathbf{A})^{-1}\mathbf{d})d\mathbf{G}(\Sigma)}{\int_{\aleph_C}g_\phi(\mathbf{d}'(\mathbf{A}')^{-1}\mathbf{d})d\mathbf{G}(\Sigma)} \\
&= \int_{\aleph_C}\delta_\Sigma(\mathbf{x})d\mathbf{G}(\Sigma) = \int_{\aleph_C}(\overline{\mathrm{X}} - \gamma_\Sigma\mathbf{C}^{-1}_{\mathscr{C}(n)}\mathbf{x})d\mathbf{G}(\Sigma) \\
&= \overline{\mathrm{X}} - \mathbf{x}'\mathbf{C}^{-1}_{\mathscr{C}(n)}\int_{\aleph_C}\gamma_\Sigma d\mathbf{G}(\Sigma) = \overline{\mathrm{X}} - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\mathbf{x},
\end{aligned}$$

for the vector γ such that $\Psi = \mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\mathbf{1}_n\mathbf{1}_n'$.

To prove (c) we need to show that the estimate $\overline{\mathrm{X}} - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\mathbf{x}$ in part (b) equals $\mathbf{w}'\mathbf{x}$ if and only if $\Psi \in \Lambda(\mathbf{w})$. Equivalently we need $n^{-1}\mathbf{1}_n - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)} = \mathbf{w}$. To show this note the identities (i) $\gamma' = n^{-1}\mathbf{1}_n'\Psi$ and (ii) $\mathbf{C}^{-1}_{\mathscr{C}(n)} = \mathbf{H}'(\mathbf{A})^{-1}\mathbf{H}' = \mathbf{H}'(\mathbf{H\Psi H}')^{-1}\mathbf{H}'$. These lead to

$$n^{-1}\mathbf{1}_n - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)} = n^{-1}\mathbf{1}_n - n^{-1}\mathbf{1}_n'\Psi\mathbf{H}'(\mathbf{H\Psi H}')^{-1}\mathbf{H}',$$

which by Lemma 4.3.1 equals $\mathbf{w}$ if and only if $\Psi \in \Lambda(\mathbf{w})$.

Part (d) follows from (c) as $\Omega(n) = \Lambda(n^{-1}\mathbf{1}_n)$. ❑

Note that $\Psi \in \Lambda(\mathbf{w})$ in part (c) implies that $\Psi \in \Lambda(\mathbf{w}) \cap \aleph_C$. This intersection was denoted as the subset $\beta(\mathbf{w}, \mathbf{C})$ in Chapter 3. Part (d) has $\Psi \in \Omega(n) \cap \aleph_C = \beta(n^{-1}\mathbf{1}_n, \mathbf{C})$. In Figure 3.4 the $\beta(\mathbf{w}, \mathbf{C})$ subsets are the cells of the grid and the $\aleph_C$ subsets are the horizontal strips. In Theorem 4.5.2 Dom(G) is restricted to a horizontal strip $\aleph_C$. For this $\mathbf{C} \in \mathscr{C}(n)$, if Ψ lies in a cell $\beta(\mathbf{w}, \mathbf{C})$ for some $\mathbf{w} \in \mathscr{W}(n)$, than $\mathbf{w}'\mathbf{x}$ is Pitman for this $\mathbf{w}$. The estimate is $\bar{x}$ when Ψ is in the cell $\beta(n^{-1}\mathbf{1}_n, \mathbf{C})$ where the horizontal strip $\aleph_C$ intersects the vertical strip $\Omega(n)$. The corollary that follows applies the results of Theorem 4.5.2 to the special subset consisting of the matrices in the class $\Gamma(n)$. The corollary follows on using $\Gamma(n) = \aleph_B$ for $\mathbf{B} = [\mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n']$.

<u>Corollary 4.5.2</u>: (a) When Dom(**G**) = $\Gamma(n)$, then $\Psi \in \Gamma(n)$.

(b) For **G** as in (a) let $L(\mathbf{X}) = \mathrm{EM}_n(\mu, \phi, \mathbf{G})$. Then

$$\Psi = \mathbf{B} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\mathbf{1}_n\mathbf{1}_n'$$

for some vector γ of constants such that $\sum_{i=1}^{n}(\gamma_i - \bar{\gamma})^2 < \bar{\gamma}$. The Pitman estimate is given through this vector as $\delta_\Psi(\mathbf{x}) = \bar{x} - \gamma'\mathbf{e}$.

(c) For $L(\mathbf{X}) = \mathrm{EM}_n(\mu, \phi, \mathbf{G})$ with Dom(**G**) = $\Gamma(n)$, $\mathbf{w}'\mathbf{x}$ is the Pitman estimate if and only if $\Psi \in \Lambda(\mathbf{w})$.

(d) For a random vector with a distribution as in (c), the sample average is Pitman if and only if $\Psi \in \Omega(n)$.

<u>Proof:</u> Observe that $\Gamma(n) = \aleph_B$ for $\mathbf{B} = [\mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n']$. Using $\mathbf{Bx} = \mathbf{e}$, $\gamma'\mathbf{B}\gamma = \sum_{i=1}^{n}(\gamma_i - \bar{\gamma})^2$ and noting that $\mathbf{B}^{-1}_{\mathscr{C}(n)} = \mathbf{B}$, the required results follow from the corresponding expressions in Theorem 4.5.1.

□

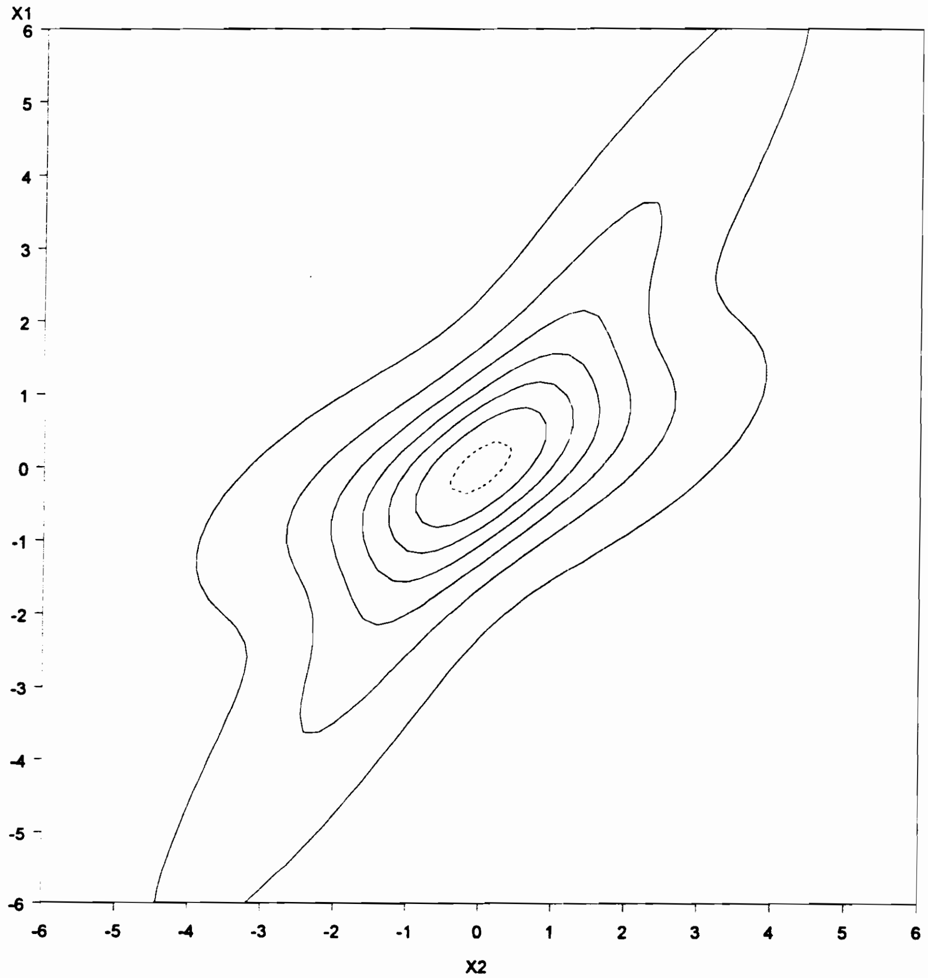

Figure 4.3: Mixture of bivariate normals with Dom(**G**) = $\Gamma(2)$ with **G** assigning 0.5 probability to $\Sigma_1 = \begin{bmatrix} 11 & 7 \\ 7 & 5 \end{bmatrix}$ and to $\Sigma_2 = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}$.

Parts (c) and (d) of the corollary imply that Ψ belongs in $\beta(\mathbf{w}, \mathbf{B})$ and $\overline{\Gamma}(n)$, respectively, where $\overline{\Gamma}(n)$ consists of equicorrelated matrices. Corollary 4.5.2 restricts **G** to the central strip in Figure 3.4. If the parameter Ψ lies in the cell $\beta(\mathbf{w}, \mathbf{B})$ for some $\mathbf{w} \in \mathcal{W}(n)$, then $\mathbf{w}'\mathbf{x}$ is Pitman. Further the Pitman estimate is $\overline{x}$ whenever Ψ lies in the cell $\overline{\Gamma}(n)$ where $\Gamma(n)$ intersects $\Omega(n)$. Figure 4.3 depicts a mixture of two bivariate normal distributions, each having a dispersion matrix in $\Gamma(2)$. Using part (b) of the corollary, the MRE estimate under squared error loss is obtained for this distribution in the following example.

Example 4.7: MRE estimation for a mixture with Dom(**G**) = $\Gamma(2)$

Let $\Sigma_1 = \begin{bmatrix} 11 & 7 \\ 7 & 5 \end{bmatrix}$ and $\Sigma_2 = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}$, and let **G** be a c.d.f. which assigns equal probability of 0.5 to each of these. These matrices are members of $\Gamma(2)$ and $\Psi = \int_{\Gamma(2)} \Sigma d\mathbf{G}(\Sigma) = \begin{bmatrix} 6 & 4 \\ 4 & 4 \end{bmatrix}$.

When the bivariate distributions with scale parameters Σ_1 and Σ_2 are mixed the resulting mixture density is shown in Figures 2.2 and 4.3. If $\mathbf{x}' = (5, 4)$ were the sample vector, then

$$\delta_{\Psi}(\mathbf{x}) = \overline{x} - \gamma'\mathbf{e} = \overline{x} - \begin{bmatrix} 5 & 4 \end{bmatrix} \begin{bmatrix} 1/2 \\ -1/2 \end{bmatrix} = 4.5 - 0.5 = 4.$$

As in the preceding figure, Figure 4.4 depicts a mixture with **G** restricted to $\Gamma(2)$. The mixture distribution now has the additional property in Corollary 4.5.2(d), namely, that Ψ now is a member of $\Omega(2)$ as well as the class $\overline{\Gamma}(2)$containing two-dimensional matrices proportional to an equicorrelated matrix. The Pitman estimate is the sample average as shown in the following example.

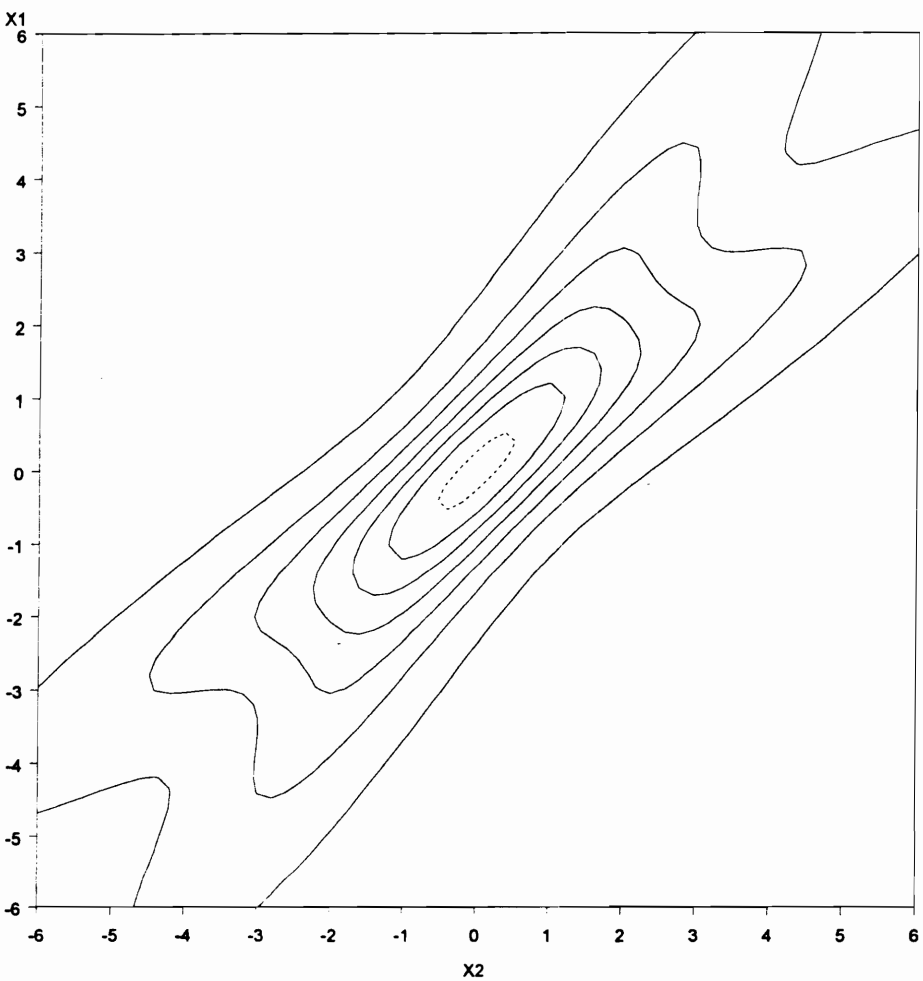

Figure 4.4: Mixture of bivariate normals with Dom($\mathbf{G}$) = $\Gamma(2)$ and $\Psi = \int_{\Gamma(2)} \Sigma d\mathbf{G}(\Sigma)$ belonging in $\Omega(n)$ and $\overline{\Gamma}(2)$.The measure $\mathbf{G}$ assigns 0.5 probability to $\Sigma_1 = \begin{bmatrix} 11 & 7 \\ 7 & 5 \end{bmatrix}$ and to $\Sigma_2 = \begin{bmatrix} 5 & 7 \\ 7 & 11 \end{bmatrix}$.

Example 4.8: MRE estimation for mixtures with Dom(**G**) = $\Gamma(2)$ with $\Psi = \int_{\Gamma(2)} \Sigma d\mathbf{G}(\Sigma)$ belonging to $\Omega(2)$ and $\overline{\Gamma}(2)$.

Let $\Sigma_1 = \begin{bmatrix} 11 & 7 \\ 7 & 5 \end{bmatrix}$ and $\Sigma_2 = \begin{bmatrix} 5 & 7 \\ 7 & 11 \end{bmatrix}$, and let **G** be a c.d.f. which assigns equal probability of 0.5 to each of these. These matrices are members of $\Gamma(n)$ and $\int_{\Gamma(n)} \Sigma d\mathbf{G}(\Sigma) = \Psi = \begin{bmatrix} 8 & 7 \\ 7 & 8 \end{bmatrix}$.

When the bivariate distributions with scale parameters Σ_1 and Σ_2 are mixed the resulting mixture density is shown in Figure 4.4. If $\mathbf{x}' = (5, 4)$ were the sample vector, then

$$\delta_\Psi(\mathbf{x}) = \overline{X} = 4.5.$$

The next section presents exhaustive conditions under which the sample average is MRE under squared error loss. As in this section the subset $\Omega(n)$ figures prominently.

4.6: CONDITIONS FOR THE ADMISSIBILITY OF THE SAMPLE AVERAGE

In this section we consider conditions for the admissibility of the sample average within the class of equivariant estimators. Kagan, Linnik and Rao (1973) considered distributions in the location family and assumed independence. The admissibility of the sample average $\overline{X}$ for this assumption was a unique characteristic of normality. The components of a random vector with a mixture distribution of the type studied here are not independent. In this context restrictions on the mixing measure **G** often preserve the admissibility of $\overline{X}$. This section explores looser restrictions than those explored in Section 4.5. These conditions are obtained in Corollary 4.6.1. Theorem 4.6.1 that follows leads to this corollary. Recall from Corollary 4.4.1 that the Pitman estimate for the mixture distribution is a weighted average, i.e., it takes the form $\mathbf{a}'\mathbf{x}$ for some vector $\mathbf{a}$ such that $\mathbf{1}_n'\mathbf{a} = 1$. The class containing all such vectors is denoted by $\mathscr{W}(n)$ as before. In Example 4.5 the appropriate vector of coefficients which yielded the weighted average depends on the sample vector $\mathbf{x}$ through the maximal invariant $\mathbf{d}$. For every chosen $\mathbf{w} \in \mathscr{W}(n)$, Theorem 4.6.1 now identifies a set of conditions on the distribution of a random vector $\mathbf{X}$ such that the Pitman estimate invariantly equals $\mathbf{w}'\mathbf{x}$ for all $\mathbf{x} \in \Re^n$. The partition of S_n^+ into a family $\mathscr{D} = \{\aleph_C | \mathbf{C} \in \mathscr{C}(n)\}$ in Theorem 3.3.2 (b) allows the decomposition of a probability measure $\mathbf{G}(.)$ into probability measures $\mathbf{G}(.|\ \mathbf{C})$. This takes the form

$$\mathbf{G}[\cdot;\mathbf{N}] = \int_{\mathscr{C}(n)} \mathbf{G}[\cdot|\mathbf{C}]d\mathbf{N}(\mathbf{C})$$

where $\mathbf{N}$ is a probability measure over $\mathscr{C}(n)$. In Theorem 4.6.1 the condition for $\mathbf{w}'\mathbf{x}$ to be Pitman for all $\mathbf{x} \in \Re^n$ is that

$$\int_{\aleph_C} \Sigma d\mathbf{G}(\Sigma|\mathbf{C}) \in \Lambda(\mathbf{w})$$

for each $\mathbf{C}$ in Dom($\mathbf{N}$). The probability measure $\mathbf{N}$ over $\mathscr{C}(n)$ can be arbitrary.

Theorem 4.6.1: (a) Let $L(\mathbf{X}) = \mathrm{EM}_n(\mu,\phi,\mathbf{G})$, and suppose that $\int_{\aleph_C} \Sigma d\mathbf{G}(\Sigma|\mathbf{C}) \in \Lambda(\mathbf{w})$ for all $\mathbf{C} \in \mathscr{C}(n)$ except for subsets of measure zero. Then $\mathbf{w}'\mathbf{x}$ is the Pitman estimate for μ for every $\mathbf{x} \in \Re^n$.

(b) For a mixing probability measure $\mathbf{G}(.)$ such that $\int_{\aleph_C} \Sigma d\mathbf{G}(\Sigma|\mathbf{C}) \in \Lambda(\mathbf{w})$ for all $\mathbf{C} \in \mathscr{C}(n)$, the matrix $\Psi = \int_{S_n^+} \Sigma d\mathbf{G}(\Sigma) \in \Lambda(\mathbf{w})$.

Proof: Write $\mathbf{G}(.)$ as

$$\mathbf{G}(.) = \mathbf{G}[\cdot;\mathrm{N}] = \int_{\mathscr{C}(n)} \mathbf{G}[\cdot|\mathbf{C}]d\mathrm{N}(\mathbf{C})$$

for some probability measure N over $\mathscr{C}(n)$. Fix $\mathbf{H} \in \mathscr{H}(n)$. Then by Theorem 4.4.1 the Pitman estimate is

$$\delta_{G,\phi}(\mathbf{x};\mathbf{d}) = \frac{\int_{S_n^+} \delta_\Sigma(\mathbf{x})g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d})d\mathbf{G}(\Sigma)}{\int_{S_n^+} g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d})d\mathbf{G}(\Sigma)}$$

$$= \frac{\int_{\mathscr{C}(n)}\int_{\aleph_C} \delta_\Sigma(\mathbf{x})g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d})d\mathbf{G}(\Sigma|\mathbf{C})d\mathrm{N}(\mathbf{C})}{\int_{\mathscr{C}(n)}\int_{\aleph_C} g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d})d\mathbf{G}(\Sigma|\mathbf{C})d\mathrm{N}(\mathbf{C})},$$

where $g_\phi(.)$ is the density of the random vector $\mathbf{D} = \mathbf{H}\mathbf{X}$ for $L(\mathbf{X}) = \mathrm{E}_n(\mu\mathbf{1}_n,\Sigma,\phi)$. To simplify the expression further consider for each $\mathbf{C} \in \mathscr{C}(n)$ the matrix $\mathbf{A}_C \in S_{n-1}^+$ such that $\mathbf{C} = \mathbf{H}'\mathbf{A}_C\mathbf{H}$. Then $L(\mathbf{D}) = \mathrm{E}_{n-1}(\mu\mathbf{H}\mathbf{1}_n,\mathbf{H}\Sigma\mathbf{H}',\phi) = \mathrm{E}_{n-1}(\mathbf{0},\mathbf{A}_C,\phi)$ is invariant as Σ varies over each subset $\aleph_C$. The density $g_\phi(.)$ can be reexpressed as

$$g_\phi[\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d}] = g_\phi[\mathbf{x}'\mathbf{H}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{H}\mathbf{x}] = g_\phi[\mathbf{x}'\mathbf{H}'(\mathbf{A}_C)^{-1}\mathbf{H}\mathbf{x}] = g_\phi[\mathbf{x}'\mathbf{C}_{\mathscr{C}(n)}^{-1}\mathbf{x}].$$

Factoring $g_\phi(.)$ out of the middle integrals yields

$$\delta_{G,\phi}(\mathbf{x};\mathbf{d}) = \frac{\int_{\mathscr{C}(n)} g_\phi(\mathbf{x}'\mathbf{C}_{\mathscr{C}(n)}^{-1}\mathbf{x}) \int_{\aleph_C} \delta_\Sigma(\mathbf{x})d\mathbf{G}(\Sigma|\mathbf{C})d\mathrm{N}(\mathbf{C})}{\int_{\mathscr{C}(n)} g_\phi(\mathbf{x}'\mathbf{C}_{\mathscr{C}(n)}^{-1}\mathbf{x}) \int_{\aleph_C} d\mathbf{G}(\Sigma|\mathbf{C})d\mathrm{N}(\mathbf{C})}. \qquad (4.6.1)$$

For each $\mathbf{C}$ the integral $\int_{\aleph_C} \delta_\Sigma(\mathbf{x})d\mathbf{G}(\Sigma|\mathbf{C}) = \mathbf{w}'\mathbf{x}$ for some $\mathbf{w} \in \mathscr{W}(n)$ if and only if $\Psi_C = \int_{\aleph_C} \Sigma d\mathbf{G}(\Sigma|\mathbf{C})$ belongs in $\Lambda(\mathbf{w})$. This follows on noting that (i) $\Psi_C = \mathbf{C}+\gamma\mathbf{1}_n'+\mathbf{1}_n\gamma'-\bar{\gamma}\mathbf{1}_n\mathbf{1}_n'$

for some vector γ, (ii) $\int_{\aleph_C} \delta_\Sigma(\mathbf{x})d\mathbf{G}(\Sigma) = \overline{X} - \gamma' \mathbf{C}_{\mathscr{C}(n)}^{-1}\mathbf{x}$ for this vector, and (iii) that $\overline{X} - \gamma' \mathbf{C}_{\mathscr{C}(n)}^{-1}\mathbf{x} = \mathbf{w'x}$ if and only if $\Psi_C \in \Lambda(\mathbf{w})$. These follow using arguments as in parts (b) and (c) of Theorem 4.5.2. Thus on assuming the condition

$$\int_{\aleph_C} \Sigma d\mathbf{G}(\Sigma|\mathbf{C}) \in \Lambda(\mathbf{w})$$

for all $\mathbf{C} \in \mathscr{C}(n)$, the integral $\int_{\aleph_C} \delta_\Sigma(\mathbf{x})d\mathbf{G}(\Sigma|\mathbf{C})$ equals $\mathbf{w'x}$ for all $\mathbf{C}$. Using this in (4.6.1) we obtain

$$\delta_{G,\phi}(\mathbf{x};\mathbf{d}) = \frac{\int_{\mathscr{C}(n)} g_\phi(\mathbf{x}'\mathbf{C}_{\mathscr{C}(n)}^{-1}\mathbf{x})(\mathbf{w}'\mathbf{x})d\mathbf{N}(\mathbf{C})}{\int_{\mathscr{C}(n)} g_\phi(\mathbf{x}'\mathbf{C}_{\mathscr{C}(n)}^{-1}\mathbf{x})d\mathbf{N}(\mathbf{C})} = \mathbf{w'x},$$

as required.

To prove (b) we need to show that $\mathbf{w}'\Psi = a\mathbf{1}_n'$ for some scalar a. Compute

$$\mathbf{w}'\Psi = \mathbf{w}'\left[\int_{\mathscr{C}(n)}\int_{\aleph_C} \Sigma d\mathbf{G}(\Sigma|\mathbf{C})d\mathbf{N}(\mathbf{C})\right] = \mathbf{w}'\left[\int_{\mathscr{C}(n)} \Psi_C d\mathbf{N}(\mathbf{C})\right] = \int_{\mathscr{C}(n)} (\mathbf{w}'\Psi_C)d\mathbf{N}(\mathbf{C}).$$

Since $\Psi_C = \int_{\aleph_C} \Sigma d\mathbf{G}(\Sigma|\mathbf{C})$ belongs in $\Lambda(\mathbf{w})$, $\mathbf{w}'\Psi_C = a_C\mathbf{1}_n'$ for some scalar a_C. Using this gives

$$\mathbf{w}'\Psi = \int_{\mathscr{C}(n)} (\mathbf{w}'\Psi_C)d\mathbf{N}(\mathbf{C}) = \int_{\mathscr{C}(n)} (a_C\mathbf{1}_n')d\mathbf{N}(\mathbf{C}) = \left[\int_{\mathscr{C}(n)} a_C d\mathbf{N}(\mathbf{C})\right]\mathbf{1}_n',$$

which is a scalar multiple of the unit vector as required. ❑

In terms of the grid of subsets in Figure 3.4, Theorem 4.6.1 considers probability measures **G** which are not restricted to each vertical or horizontal strip as was the case in Section 4.5. The measure **G** now has domain across the horizontal strips with the property that the parameters

$$\Psi_C = \int_{\aleph_C} \Sigma d\mathbf{G}(\Sigma|\mathbf{C})$$

belong in a subset $\Lambda(\mathbf{w})$ for all $\mathbf{C}$ (except with measure zero). The mixing measure $\mathbf{N}$ over $\mathscr{C}(n)$ can be arbitrary. The corollary that follows applies Theorem 4.6.1 to obtain conditions for the sample average to be the Pitman estimate.

Corollary 4.6.1: (a) Let $L(\mathbf{X}) = EM_n(\mu, \phi, \mathbf{G})$, and suppose that $\int_{\aleph_C} \Sigma d\mathbf{G}(\Sigma|\mathbf{C}) \in \Omega(n)$ for all $\mathbf{C} \in \mathscr{C}(n)$ except for subsets of measure zero. Then the sample average $\overline{x}$ is the Pitman estimate for μ for every $\mathbf{x} \in \Re^n$.

(b) For a mixing probability measure $\mathbf{G}(.)$ such that $\int_{\aleph_C} \Sigma d\mathbf{G}(\Sigma|\mathbf{C}) \in \Omega(n)$ for all $\mathbf{C} \in \mathscr{C}(n)$, the matrix $\Psi = \int_{S_n^+} \Sigma d\mathbf{G}(\Sigma) \in \Omega(n)$.

Proof: These follow from Theorem 4.6.1 as $\Omega(n) = \Lambda(n^{-1}\mathbf{1}_n')$. ❑

In this Corollary we have exhaustive conditions under which the sample average is equivariantly admissible. This is illustrated through Figure 4.5. Each univariate plot pictured represents a conditional mixing distribution $\mathbf{G}(.|\mathbf{C})$ over each subset $\aleph_C$. The symmetries of these distributions yield expectations

$$\Psi_C = \int_{\aleph_C} \Sigma d\mathbf{G}(\Sigma|\mathbf{C})$$

within the class $\Omega(n)$. Such conditional distributions give rise to a wide range of mixing measures $\mathbf{G}(.)$ over S_n^+. These mixing measures, in turn, generate broad classes of irregularly contoured mixture densities of the form

$$h(\mathbf{x}; \mu, \phi, \mathbf{G}) = \int_{S_n^+} f_\phi\left[(\mathbf{x} - \mu\mathbf{1}_n)'\Sigma^{-1}(\mathbf{x} - \mu\mathbf{1}_n)\right] d\mathbf{G}(\Sigma)$$

where $f_\phi(.)$ is the density of the elliptically symmetric components in the mixture, and $\mathbf{G}$ is the probability measure over S_n^+. For random vectors with such densities the sample average continues to be the Pitman estimator.

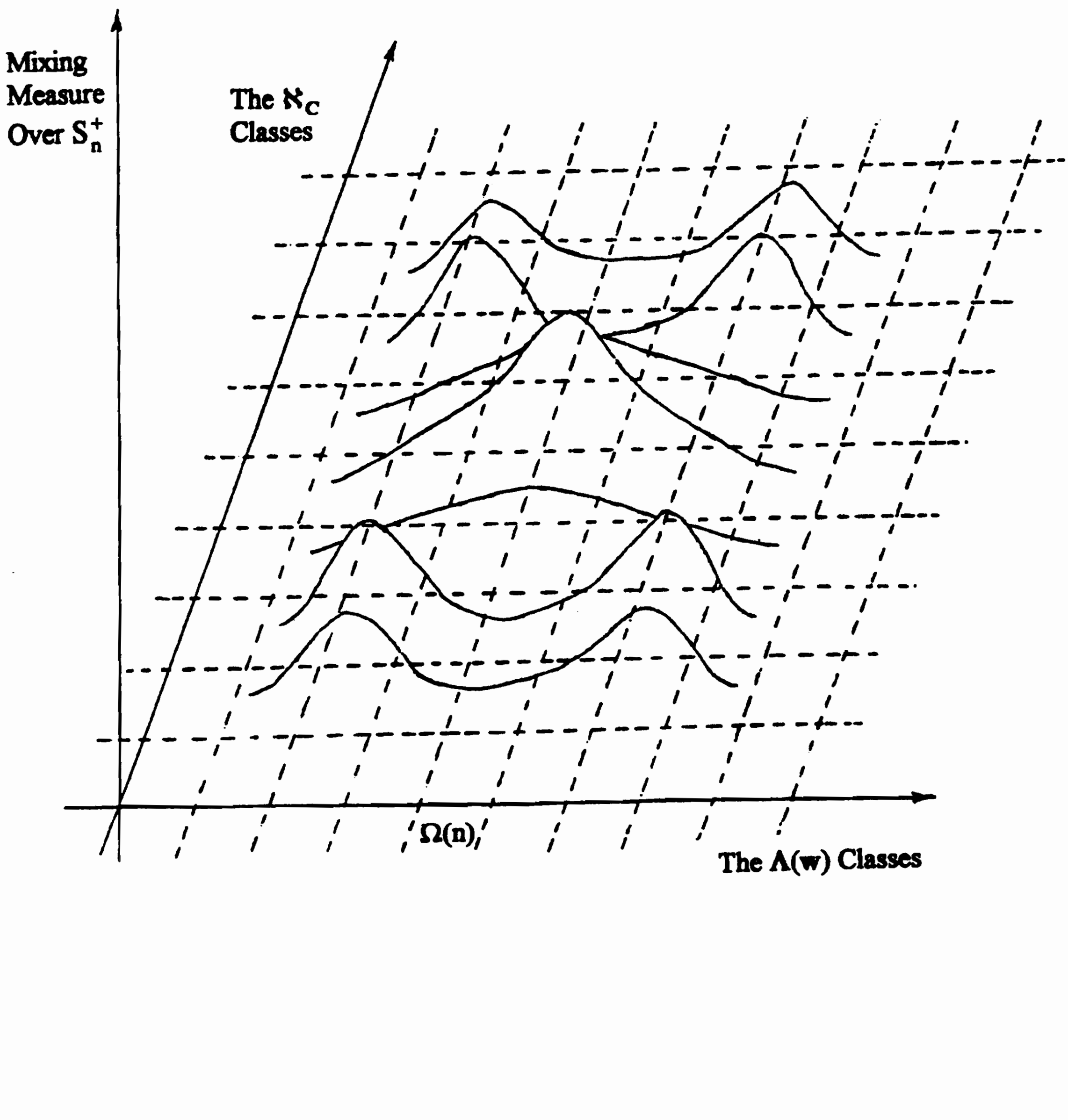

Figure 4.5: Illustration of the conditions on the mixing distribution $\mathbf{G}$ for the equivariant admissibility of the sample average when $L(\mathbf{X}) = \mathrm{EM}_n(\mu,\phi,\mathbf{G})$. The expectation $\Psi_C = \int_{\aleph_C} \Sigma d\mathbf{G}(\Sigma|C)$ for each $\aleph_C$ strip needs to be in $\Omega(n)$.

The next section reviews results in this chapter dealing with Pitman estimation for ensembles of symmetric distributions and for mixtures over these.

4.7: RETROSPECTIVE

This section summarizes results dealing with Pitman estimation developed in this chapter. Section 4.2 develops the general form of the Pitman estimate for elliptically symmetric distributions. The estimate is a linear unbiased weighted average. It is given through the scale parameter Σ of the elliptically symmetric distributions as the Aitkin (1934) estimate in the expression below. Further the following identity to the Aitkin form is obtained in terms of matrices $\mathbf{H} \in \mathscr{H}$ as

$$\delta_\Sigma(\mathbf{x}) = \overline{X} - n^{-1}\,\mathbf{c}'\mathbf{H}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{H}\mathbf{x} = (\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n)^{-1}\mathbf{1}_n'\Sigma^{-1}\mathbf{x}.$$

Section 4.3 specializes this to the subsets $\aleph_C$ and the subsets $\Lambda(\mathbf{w})$. First each elliptically symmetric random vector with distribution $L(\mathbf{X}) = E_n(\,\mu\mathbf{1}_n, \Sigma, \phi\,)$ with $\Sigma \in \aleph_C$, is considered. The index matrix $\mathbf{C}$ specifying each class $\aleph_C$ is a member of the class $\mathscr{C}(n)$, and $\mathbf{C}^{-1}_{\mathscr{C}(n)}$ denotes the minimum norm least squares inverse of $\mathbf{C}$. When Σ has the expansion

$$\Sigma = \Sigma(\gamma) = \mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n'$$

for a vector γ of constants such that $\gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\gamma < \bar{\gamma}$, then the Pitman estimate for μ is shown to be

$$\delta_\Sigma(\mathbf{X}) = \overline{X} - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\mathbf{X}.$$

This result is applied to the subset $\Gamma(n)$. For elliptically symmetric distributions with dispersion matrices in this subset the estimate is given by

$$\delta_\Sigma(\mathbf{X}) = \overline{X} - \gamma'\mathbf{e},$$

where the vector γ is as in the expansion of Σ, and $\mathbf{e}$ is the vector of residuals $e_i = x_i - \bar{x}$ for i = 1, 2, ..., n. The general form for the Pitman estimate is applied to the subsets $\Lambda(\mathbf{w})$. The vector identifying each subset is selected from the class $\mathcal{W}(n)$ of all n-dimensional vectors $\mathbf{w}$ such that $\mathbf{1}_n'\mathbf{w} = 1$. For every elliptically symmetric random vector $\mathbf{X}$ having a dispersion matrix in the class $\Lambda(\mathbf{w})$, $\mathbf{w'x}$ is the Pitman estimator.

Section 4.4 obtains the general form of the Pitman estimate for mixture distributions. Random vectors with these distributions have densities taking the form

$$h(\mathbf{x};\, \mu,\, \phi,\, \mathbf{G}) = \int_{S_n^+} f_\phi\left[(\mathbf{x}-\mu\mathbf{1}_n)'\Sigma^{-1}(\mathbf{x}-\mu\mathbf{1}_n)\right]d\mathbf{G}(\Sigma)$$

where $f_\phi(.)$ is the density of the elliptically symmetric components in the mixture with parameter ϕ, and $\mathbf{G}$ is a probability measure over S_n^+. For any (n-1)×n-dimensional matrix $\mathbf{H}$ obtained as an orthonormal completion of the unit vector, the Pitman estimate is

$$\delta_{\mathbf{G},\phi}(\mathbf{x};\mathbf{d}) = \frac{\int_{S_n^+} \delta_\Sigma(\mathbf{x})g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d})d\mathbf{G}(\Sigma)}{b(\mathbf{d};\mathbf{G},\phi)}$$

where $g_\phi(.)$ is the density of the (n-1)-dimensional random vector $\mathbf{D} = \mathbf{HX}$ for each elliptically symmetric component $\mathbf{X}$ in the mixture. Further results in Section 4.4 show that this estimator is an unbiased weighted average of the elements of the random vector $\mathbf{X}$. It takes the form $\mathbf{k'x}$ where $\mathbf{k}$ can depend on the maximal invariant $\mathbf{d}$ through the parameters ϕ and $\mathbf{G}$. An example where it does depend on the value taken by a maximal invariant is in Example 4.5.

Section 4.5 places restrictions on $\mathbf{G}$. The domain of $\mathbf{G}$ is restricted to the classes $\aleph_C$ and $\Lambda(\mathbf{w})$. Under these restrictions the Pitman estimate is not a function of the maximal invariant $\mathbf{d}$, and depends on $\mathbf{G}$ through a parameter

$$\Psi = \int_{S_n^+} \Sigma d\mathbf{G}(\Sigma).$$

This parameter is the dispersion matrix for the mixture random variable. When Dom($\mathbf{G}$) = $\Lambda(\mathbf{w})$ for some $\mathbf{w}$, then $\Psi \in \Lambda(\mathbf{w})$, and $\mathbf{w}'\mathbf{x}$ is the Pitman estimate. In particular, the sample average is Pitman when Dom($\mathbf{G}$) = $\Omega(n)$. In terms of the representation of S_n^+ in Figure 3.4 these results imply arbitrary measures over each vertical strip $\Lambda(\mathbf{w})$. Succeeding results in Section 4.5 have $\mathbf{G}$ restricted to each subset $\aleph_C$. The parameters Ψ is in $\aleph_C$ and is expressible as

$$\Psi = \Sigma(\gamma) = \mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n'$$

for some vector γ. For this γ, the Pitman estimate for μ is

$$\delta_\Sigma(\mathbf{X}) = \overline{\mathbf{X}} - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\mathbf{X}.$$

This result is applied to the case Dom($\mathbf{G}$) = $\Gamma(n)$. Then Ψ is in $\Gamma(n)$ and has an expansion as above with $\mathbf{C}$ = $[\,\mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n']$ and the Pitman estimate simplifies as in the elliptically symmetric case. In terms of Figure 3.4 the preceding results concern mixtures obtained by mixing over each horizontal strip. The two estimators equal $\mathbf{w}'\mathbf{x}$ whenever the parameter Ψ is an element of the corresponding vertical strip $\Lambda(\mathbf{w})$.

In Section 4.6 conditions for equivariant admissibility of a linear estimator are explored when $L(\mathbf{X})$ = $EM_n(\mu,\phi,\mathbf{G})$. To find these the probability measure $\mathbf{G}(.)$ is decomposed as

$$\mathbf{G}[\cdot;\mathbf{N}] = \int_{\mathscr{C}(n)} \mathbf{G}[\cdot|\mathbf{C}]d\mathbf{N}(\mathbf{C})$$

where $\mathbf{N}$ is a probability measure over $\mathscr{C}(n)$. Theorem 4.6.1 considers $\Psi_C = \int_{\aleph_C} \Sigma d\mathbf{G}(\Sigma|\mathbf{C})$.

When $\Psi_C \in \Lambda(\mathbf{w})$ for all $\mathbf{C}$ in Dom($\mathbf{N}$), then $\mathbf{w}'\mathbf{x}$ is Pitman. The mixing measure over $\mathbf{N}$ can be arbitrary. In particular, when $\Psi_C \in \Omega(n)$ for all $\mathbf{C}$ then the sample average is equivariantly admissible under squared error loss. A mixing distribution over S_n^+ which meets this condition is depicted in Figure 4.5. This summarizes essential results. The next chapter presents conclusions and further developments.

Chapter 5

CONCLUSIONS AND FURTHER STUDY

5.1 CONCLUSIONS

This dissertation uncovers symmetries in two contexts. Firstly, we derive a two-way partition of the class S_n^+ of positive definite matrices. Secondly, we find minimal risk equivariant (MRE) estimators for μ for the case of vector observations with structured dispersion in subsets of the partition. Each positive definite matrix is obtained as a simple linear function of some vector **w** from the class $\mathcal{W}(n)$ and some matrix **C** from the class $\mathcal{C}(n)$. The class $\mathcal{W}(n)$ contains all n-dimensional vectors **w** such that $\sum_{i=1}^{n} w_i = 1$, and the class $\mathcal{C}(n)$ contains n-dimensional real symmetric matrices with one-dimensional null space spanned by the unit vector. Through elements of these classes we can write each matrix Σ of S_n^+ as

$$\Sigma = \Sigma(\mathbf{C}, \mathbf{w}, \xi) = \mathbf{C} - \mathbf{Cw1}_n' - \mathbf{1}_n\mathbf{w}'\mathbf{C}' + \xi\mathbf{1}_n\mathbf{1}_n',$$

for some scalar ξ such that $\xi > \mathbf{w}'\mathbf{Cw}$. For fixed **w**, varying **C** and ξ within constraints generates the class $\Lambda(\mathbf{w})$. Each class $\Lambda(\mathbf{w})$ contains matrices for which $\mathbf{w}'\Sigma$ is a scalar multiple of the unit vector. For fixed **C**, varying **w** and ξ generates the class $\aleph_C$. Characterizations of the classes $\aleph_C$ and others studied are summarized in Table 5.1. Some subsets of interest to statisticians are obtained in our classification. The class $\Gamma(n)$ contains matrices studied by Huynh and Feldt (1970) and Rouanet and Le'pine (1970). Dispersion structures in $\Gamma(n)$ validate the use of F-ratios in the context of experiments using repeated measures. A second class of matrices, denoted as $\Omega(n)$, is studied in Jensen (1989a). Dispersion structures within this class are seen to lead to the independence of S^2 and $\overline{X}$. A third class of matrices, the equicorrelated class, arises as

Table 5.1: Characterization of Classes of Positive Definite Matrices.

Classes of Matrices	*Characterization*
The Classes $\aleph_C$.	A matrix Σ is in each class $\aleph_C$ if and only if $\Sigma = \Sigma(\gamma) = \mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n'$, for some vector γ such that $\gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\gamma < \bar{\gamma}$.
The Class $\Gamma(n)$.	The subset $\aleph_\mathbf{B}$ for $\mathbf{B} = [\,\mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n']$.
The Classes $\Lambda(\mathbf{w})$.	A matrix is in each class $\Lambda(\mathbf{w})$ if and only if $\Sigma = \Sigma(\mathbf{C},\xi) = \mathbf{C} - \mathbf{Cw1}_n' - \mathbf{1}_n\mathbf{w'C'} + \xi\mathbf{1}_n\mathbf{1}_n'$ for $\mathbf{C}$ and ξ such that $\xi > \mathbf{w'Cw}$.
The Class $\Omega(n)$.	The subset $\Lambda(n^{-1}\mathbf{1}_n)$.
The Classes $\beta(\mathbf{w},\mathbf{C})$.	The subsets obtained as intersections of the classes $\Lambda(\mathbf{w})$ with $\aleph_C$.
The Class $\bar{\Gamma}(n)$.	The subset containing the equicorrelated matrices obtained as $\Gamma(n)\cap\Omega(n)$.

Table 5.2: Pitman Estimators for Elliptical Ensembles.

Elliptical Ensemble	*Pitman Estimator*
Distributions $L(\mathbf{X}) = E_n(\mu\mathbf{1}_n, \Sigma, \phi)$ with Σ from the subsets $\aleph_C$.	$\delta_\Sigma(\mathbf{X}) = \bar{X} - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\mathbf{X}$. for the vector γ in the expansion of Σ as $\Sigma(\gamma) = \mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n'$.
Distributions $L(\mathbf{X}) = E_n(\mu\mathbf{1}_n, \Sigma, \phi)$ with Σ from the subsets $\Gamma(n)$.	$\delta_\Sigma(\mathbf{X}) = \bar{X} - \gamma'\mathbf{e}$, where the vector γ is as in the expansion of Σ, and $\mathbf{e}$ is the vector of residuals $e_i = x_i - \bar{x}$ for $i = 1, 2, \ldots, n$.
Distributions $L(\mathbf{X}) = E_n(\mu\mathbf{1}_n, \Sigma, \phi)$ with Σ from the subsets $\Lambda(\mathbf{w})$.	$\delta_\Sigma(\mathbf{X}) = \mathbf{w'X}$.
Distributions $L(\mathbf{X}) = E_n(\mu\mathbf{1}_n, \Sigma, \phi)$ with Σ from the subsets $\Omega(n)$.	$\delta_\Sigma(\mathbf{X}) = \bar{X}$.

the intersection of $\Omega(n)$ and $\Gamma(n)$. In Table 5.1, we see that these classes and others studied here take simple linear forms which may find easy applications elsewhere. In Chapter 4 we consider estimation for ensembles and mixtures of elliptical distributions obtained as the dispersion matrices are chosen from within the subsets studied.

Table 5.2 gives the MRE (Pitman) estimator under squared error loss for elliptically symmetric distributions as the dispersion matrices are chosen from within the subsets of S_n^+ in Table 5.1. The Pitman estimator for elliptically symmetric distributions is found to be linear and unbiased. It takes the Aitkin (1934) form of $\delta(\mathbf{X}) = (\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n)^{-1}\mathbf{1}_n'\Sigma^{-1}\mathbf{X}$. This simplifies for elliptical ensembles obtained on restricting Σ to each class of positive definite matrices. The necessary and sufficient conditions for the sample average to be Pitman for elliptical random vectors is the membership of Σ in the class $\Omega(n)$. Table 5.3 gives the Pitman estimators for mixtures having second order moments. As with elliptical random vectors, the Pitman estimator continues to be an unbiased weighted average. For unrestricted mixing measures $\mathbf{G}$ over S_n^+, the Pitman estimate does depend on the data through the maximal invariant under scalar translations. We use an $(n-1)$-dimensional vector obtained as contrasts of the form $\mathbf{d} = \mathbf{Hx}$ as the maximal invariant. The Pitman estimate depends on the maximal invariant through its density $g_\phi(.)$ as in the first row of Table 5.3. When the domain of $\mathbf{G}$ is restricted, the Pitman estimate is independent of the maximal invariant. In these contexts the mixture distributions behave as if they were elliptically symmetric distributions. The parameter

$$\Psi = \int_{S_n^+} \Sigma d\mathbf{G}(\Sigma)$$

plays the same role as did the dispersion matrices for elliptically symmetric distributions. The matrix Ψ is proportional to the dispersion matrix of the mixture random vector. The last row of Table 5.3 presents exhaustive conditions under which the sample average is

Table 5.3: Pitman Estimates for Mixture Distributions.

Property of mixing Measure	*Pitman Estimate*
Unrestricted Domain for the Mixing measure **G**.	$\delta_{\mathbf{G},\phi}(\mathbf{x};\mathbf{d}) = \dfrac{\int_{S_n^+} \delta_\Sigma(\mathbf{x}) g_\phi(\mathbf{d}'(\mathbf{H}\Sigma\mathbf{H}')^{-1}\mathbf{d}) d\mathbf{G}(\Sigma)}{b(\mathbf{d};\mathbf{G},\phi)}$
Dom(**G**) = $\aleph_C$.	$\delta_\Psi(\mathbf{x}) = \overline{X} - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\mathbf{x}$. for the vector γ in the expansion of $\Psi = \int_{S_n^+} \Sigma d\mathbf{G}(\Sigma)$ as $\Psi(\gamma) = \mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\mathbf{1}_n\mathbf{1}_n'$.
Dom(**G**) = $\aleph_C$ with $\Psi \in \Lambda(\mathbf{w})$.	$\delta_\Psi(\mathbf{x}) = \mathbf{w}'\mathbf{x}$.
Dom(**G**) = $\aleph_C$ with $\Psi \in \Omega(n)$.	$\delta_\Psi(\mathbf{x}) = \overline{x}$.
Dom(**G**) = $\Lambda(\mathbf{w})$.	$\delta_\Psi(\mathbf{x}) = \mathbf{w}'\mathbf{x}$.
Dom(**G**) = $\Omega(n)$.	$\delta_\Psi(\mathbf{x}) = \overline{x}$.
Dom(**G**) = S_n^+ with $\Psi_C = \int_{\aleph_C} \Sigma d\mathbf{G}(\Sigma \mid \mathbf{C}) \in \Lambda(\mathbf{w})$ for all **C** from the class $\mathscr{C}(n)$.	$\delta_\Psi(\mathbf{x}) = \mathbf{w}'\mathbf{x}$.
Dom(**G**) = S_n^+ with $\Psi_C = \int_{\aleph_C} \Sigma d\mathbf{G}(\Sigma \mid \mathbf{C}) \in \Omega(n)$ for all **C** from the class $\mathscr{C}(n)$.	$\delta_\Psi(\mathbf{X}) = \overline{x}$.

equivariantly admissible under squared error loss. We consider conditional mixing distributions $\mathbf{G}(.|\mathbf{C})$ over each subset $\aleph_C$ with expectations

$$\Psi_C = \int_{\aleph_C} \Sigma d\mathbf{G}(\Sigma|C)$$

within the class $\Omega(n)$. Such conditional distributions have mixing measures $\mathbf{G}(.)$ over S_n^+ which generate highly irregular mixture distributions for which the sample average continues to be the Pitman estimator. Similar conclusions arise for the admissibility of each estimator $\mathbf{w}'\mathbf{x}$. The estimator for μ in the contexts listed prior to these two can be seen to be special cases of the latter.

Table 5.4 summarizes numerical examples concerning dispersion structures and the corresponding estimators. The first row contains parameters or properties of each of four subsets of positive definite matrices. The second row provides matrices with dispersion structures in each of the four subsets. The next row computes the Pitman estimate corresponding to these dispersion structures for a sample vector $\mathbf{x}' = [10\ 12\ 14]$. The last row computes variances of the Pitman estimator for elliptical random vectors. The four dispersion matrices, while appearing equally ill-conditioned, lead to varied Pitman estimates.

The next section provides avenues for further study in the conditioning of matrices as well as other topics.

Table 5.4: Summary of numerical examples for dispersion matrices in four classes and the corresponding Pitman estimates for μ when $\mathbf{x}' = \begin{bmatrix} 10 & 12 & 14 \end{bmatrix}$.

$\Sigma \in \aleph_C$ for $\mathbf{C} = \left(\frac{1}{9}\right)\begin{bmatrix} 20 & 2 & -22 \\ 2 & 11 & -13 \\ -22 & -13 & 35 \end{bmatrix}$ and $\gamma' = \begin{bmatrix} 3 & 3 & 11/3 \end{bmatrix}$	$\Sigma \in \Gamma(n)$ for $\gamma = \left(\frac{1}{3}\right)\begin{bmatrix} 10 \\ 13 \\ 16 \end{bmatrix}$	$\Sigma \in \Lambda(\mathbf{w})$ for $\mathbf{w} = \begin{bmatrix} -1/4 \\ 3/4 \\ 1/2 \end{bmatrix}$. $\mathbf{w}'\Sigma = c\mathbf{1}_n'$.	$\Sigma \in \Omega(n)$. $\mathbf{1}_n'\Sigma = c\mathbf{1}_n'$ (Equal Column Sums)
$\begin{bmatrix} 5 & 3 & 1 \\ 3 & 4 & 2 \\ 1 & 2 & 8 \end{bmatrix}$	$\begin{bmatrix} 3 & 3 & 4 \\ 3 & 5 & 5 \\ 4 & 5 & 7 \end{bmatrix}$	$\begin{bmatrix} 12 & 4 & 4 \\ 4 & 4 & 0 \\ 4 & 0 & 6 \end{bmatrix}$	$\begin{bmatrix} 3 & 3 & 1 \\ 3 & 5 & -1 \\ 1 & -1 & 7 \end{bmatrix}$
$\delta_\Sigma(\mathbf{x}) =$ $\bar{x} - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\mathbf{x}$ $= 12\text{-}2 = 10$	$\delta_\Sigma(\mathbf{x}) = \bar{x} - \gamma'\mathbf{e}$ $= 12 - 6 = 6$	$\delta_\Sigma(\mathbf{x}) = \mathbf{w}'\mathbf{x}$ $= 13.5$	$\delta_\Sigma(\mathbf{x}) = \bar{x}$ $= 12$
$\mathrm{Var}[\delta_\Sigma(\mathbf{X};\phi)] =$ $\alpha_\phi(\bar{\gamma} - \gamma'\mathbf{C}^{-1}_{\mathscr{C}(n)}\gamma)$ $= 3.22\alpha - 0.66\alpha$ $= 2.56\alpha.$	$\mathrm{Var}[\delta_\Sigma(\mathbf{X};\phi)] =$ $\alpha_\phi(\bar{\gamma} - \sum_{i=1}^{n}(\gamma_i - \bar{\gamma})^2)$ $= 4.33\alpha - 2.0\alpha$ $= 2.33\alpha.$	$\mathrm{Var}[\delta_\Sigma(\mathbf{X};\phi)] =$ $\alpha_\phi(\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n)^{-1}$ $= 2.0\alpha.$	$\mathrm{Var}[\delta_\Sigma(\mathbf{X};\phi)] =$ $\alpha_\phi(\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n)^{-1}$ $= 2.33\alpha.$

5.2: FURTHER WORK

This dissertation sets the stage for further research under three broad objectives. Firstly, the conditioning of matrices through their condition numbers could be related to the arguments $\mathbf{w} \in \mathscr{W}(n)$, $\mathbf{C} \in \mathscr{C}(n)$ and γ in the various subsets of S_n^+. Secondly, estimates of a vector location parameter for nonnormal vector location families can be studied in a manner similar to developments for the scalar location families given here. Thirdly, ideas in this dissertation, such as spectral properties of matrices and the use of Laplace loss functions in estimation, could be explored in greater detail.

Conditioning of matrices within S_n^+ can be gauged through condition numbers such as the function $C(\Sigma)$ given by

$$C(\Sigma) = \left[\sum_{i=1}^{n} \xi_i^2\right]\left[\sum_{i=1}^{n}\left(1/\xi_i^2\right)\right],$$

where $\{\xi_i;\ i = 1,2, \ldots,n\}$ are the n eigenvalues of the positive definite matrix Σ. This function is from a class of functions invariant under orthogonal transformations as in Marshall and Olkin (1965). We have shown how matrices in various subsets of S_n^+ are obtained as simple functions of $\mathbf{w} \in \mathscr{W}(n)$, $\mathbf{C} \in \mathscr{C}(n)$, vectors γ, and scalars ξ. Future research could examine how condition numbers of matrices are influenced by the arguments which generate them. This could lead to improved accuracy in computations by allowing us to replace ill-conditioned matrices with less ill-conditioned matrices where applicable. We provide an example illustrating this possibility in Section 1.2 in the context of the subsets $\Lambda(\mathbf{w})$. For every matrix Σ in a subset $\Lambda(\mathbf{w})$ the vector $\mathbf{w}$ takes the form $\mathbf{w} = (\mathbf{1}_n'\Sigma^{-1}\mathbf{1}_n)^{-1}\mathbf{1}_n'\Sigma^{-1}$, and positive definite matrices obtained as $\Sigma + \lambda(\mathbf{1}_n\mathbf{1}_n')$ for scalar λ are contained in $\Lambda(\mathbf{w})$. In Section 1.2 the addition of $\lambda(\mathbf{1}_n\mathbf{1}_n')$ to a matrix reduces its condition number considerably. Further research could examine the influence of the scalar ξ on the conditioning of a matrix $\Sigma \in \Lambda(\mathbf{w})$. The scalar ξ appears in the expression of a matrix $\Sigma \in \Lambda(\mathbf{w})$ as

$$\Sigma = \Sigma(\mathbf{C},\xi) = \mathbf{C} - \mathbf{C}\mathbf{w}\mathbf{1}_n' - \mathbf{1}_n\mathbf{w}'\mathbf{C}' + \xi\mathbf{1}_n\mathbf{1}_n'.$$

Similarly the matrices in each subset $\aleph_C$ have an expansion in terms of a vector γ such that $\gamma'\mathbf{C}_{\mathscr{C}(n)}^{-1}\gamma < \bar{\gamma}$. The influence of this vector on the conditioning of matrices in each subset $\aleph_C$ can also be studied for reasons similar to those for $\Lambda(\mathbf{w})$. For the subset $\Gamma(n)$, Jensen (1992) found that a vector γ more dispersed than another in the sense of majorization, is associated with a more ill-conditioned matrix $\Sigma(\gamma)$. It is conjectured that similar results may carry over to the subsets $\aleph_C$.

A second area for study is the estimation of a p-dimensional vector of location parameters based on a sample of $n \geq p$ vector observations. Pitman estimation using matrix loss functions as in Jensen and Foutz (1991) would appear to be germane in this context. As in the scalar parameter case studied here, applicability of linear estimation under these loss functions may extend beyond normality on relaxing the assumption of independence of the p-dimensional vector observations. In this context, tests that covariance structures belong to various classes of matrices studied may be relevant. Tests for covariance structure within the class $\Gamma(n)$ have been studied in Huynh and Feldt (1970). As noted, these find application in validating the use of F-ratios in the context of experiments with repeated measurements. Similarly, developments presented here validate the use of the sample average when dispersion structures are in the class $\Omega(n)$. Tests for covariance structures within the class $\Omega(n)$ remain to be studied.

Finally, further areas of study arise as extensions of material in this dissertation. Estimation for mixture distributions could be reexamined under Laplace loss functions. Except for mixtures exhibiting strong symmetries, the Pitman estimate under Laplace loss may be expected to differ from that under squared error loss. Spectral properties of matrices in the subsets $\aleph_C$ as well as the subsets $\Lambda(\mathbf{w})$ could be examined in future work. Spectral properties of matrices in the subsets $\Gamma(n)$ and $\Omega(n)$ are more tractable and are provided in this dissertation. Spectral properties of matrices in $\Gamma(n)$ are excerpted from Jensen (1992). Matrices in $\Omega(n)$ have the scaled unit vector as an eigenvector. These spectral properties may provide means for gauging corresponding properties for the general case of the subsets $\aleph_C$ and $\Lambda(\mathbf{w})$.

GLOSSARY

$\mathbf{1}_n$ - The n-dimensional unit vector.

$\Re^n$ - The set of n-dimensional vectors.

S_n^+ - The class of positive definite matrices.

$\mathscr{H}(n)$ - The class of all (n-1)×n-dimensional matrices obtained as orthonormal completions of the unit vector.

$\Xi(\mathbf{A}, \mathbf{H})$ - Classes of positive definite matrices defined through matrices $(\mathbf{A}, \mathbf{H})$ from the ordered set $\{S_{(n-1)}^+, \mathscr{H}(n)\}$. A matrix $\Sigma \in \Xi(\mathbf{A}, \mathbf{H})$ if $\mathbf{H}\Sigma\mathbf{H}' = \mathbf{A}$. This subset is also specified equivalently as $\aleph_{\mathbf{C}}$.

$\mathscr{C}(n)$ - The class of real symmetric matrices having a one dimensional null space spanned by the unit vector. Elements are denoted as $\mathbf{C}$.

$\mathbf{C}_{\mathscr{C}(n)}^{-1}$ - The minimum norm least squares g-inverse of $\mathbf{C} \in \mathscr{C}(n)$.

$\aleph_{\mathbf{C}}$ - Alternate notation for the classes $\Xi(\mathbf{A}, \mathbf{H})$ with $\mathbf{C} = \mathbf{H}'\mathbf{A}\mathbf{H}$. Elements take the form

$$\Sigma = \mathbf{C} + \gamma\mathbf{1}_n' + \mathbf{1}_n\gamma' - \bar{\gamma}\,\mathbf{1}_n\mathbf{1}_n',$$

for some vector γ such that $\gamma'\mathbf{C}_{\mathscr{C}(n)}^{-1}\gamma < \bar{\gamma}$.

$\mathbf{B}$ - The idempotent matrix obtained as $[\,\mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n']$. The matrix $\mathbf{B} \in \mathscr{C}(n)$.

$\Gamma(n)$ - The class of matrices with the Huynh Feldt structure. The class $\Gamma(n) = \aleph_{\mathbf{B}}$.

$\mathscr{W}(n)$ - The set of all vectors $\mathbf{w}$ such that $\sum_{i=1}^{n} w_i = 1$.

$\Lambda(\mathbf{w})$ - Classes of positive definite matrices indexed through each $\mathbf{w} \in \mathscr{W}(n)$. A matrix Σ belongs in a class $\mathscr{W}(n)$ if $\mathbf{w}'\Sigma = c\mathbf{1}_n'$ for some scalar c.

$\Omega(n)$ - The class of positive definite matrices whose elements have equal column sums. i.e. $\mathbf{1}_n'\Sigma = c\mathbf{1}_n'$. The class $\Omega(n) = \Lambda(n^{-1}\mathbf{1}_n)$.

$\beta(\mathbf{w},\mathbf{C})$ - The subset of S_n^+ obtained through $(\mathbf{w},\mathbf{C})$ from the ordered set $\{\mathscr{W}(n), \mathscr{C}(n)\}$ as the intersection $\Lambda(\mathbf{w}) \cap \aleph_{\mathbf{C}}$.

$E_n(\mu\mathbf{1}_n, \Sigma, \phi)$ - An elliptically symmetric distribution with common mean μ and dispersion structure Σ. The specific type of distribution is specified by ϕ.

$\mathrm{EM}_n(\mu, \phi, \mathbf{G})$ - A random vector $\mathbf{X}$ has this distribution when its density takes the form

$$h(\mathbf{x};\, \mu,\, \phi,\, \mathbf{G}) = \int_{S_n^+} f_\phi\left[(\mathbf{x} - \mu\mathbf{1}_n)'\Sigma^{-1}(\mathbf{x} - \mu\mathbf{1}_n)\right] d\mathbf{G}(\Sigma)$$

where $f_\phi(.)$ is the density for the elliptically symmetric components in the mixture, and $\mathbf{G}$ is a probability measure over S_n^+.

REFERENCES

1. Aitkin, A.C. (1934). "On least squares and linear combinations of observations." *Royal Soc.: Edinburgh.* A 55: 42-47.

2. Berge C. (1963). *Topological Spaces, Including a Treatment of Multivalued Functions.* Macmillan, New York.

3. Cambanis, S., Huang, S. and Simons, S. (1981). "On the theory of elliptically contoured distributions." *Journal of Multivariate Analysis* 11: 368-385.

4. Chaganty, R. N. (1992) "An invariance property of common statistical tests." Technical Report TR92-16, Department of Mathematics and Statistics, Old Dominion University, Norfolk, VA.

5. Fang, K., and Zhang, Y. (1990). *Generalized Multivariate Analysis.* Springer-Verlag, New York.

6. Huynh, H and Feldt, L. S. (1970). "Conditions under which mean square ratios in repeated measurements designs have exact F-distributions" *JASA* 65: 1582-1589.

7. Jensen, D. R. (1989a). "Misspecified location and scale II: Independence of $\overline{Y}$ and S^2."*Journal of Statistical Computation and Simulation.* 32:119-121.

8. Jensen, D. R. (1989b). "Misspecified location and scale III: The use of S^2."*Journal of Statistical Computation and Simulation.* 32:121-124.

9. --------- (1992). "Matrices having structure $\Sigma(\gamma)$." Research report R043, Department of Statistics, Virginia Polytechnic Institute, Blacksburg, VA.

10. ----------, and Foutz R. V. (1991). "Pitman estimation on $\Re^k$." *Journal of Statistical Planning and Inference.* 28: 233-240.

11. Kagan, A. M., Linnik, Y. V., and Rao, C. R. (1973). *Characterization Problems in Mathematical Statistics.* John Wiley and Sons, New York.

12. Kariya, T., and Sinha, B. K. (1989). *Robustness of Statistical Tests.* Academic Press, San Diego.

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

14. Marshall, A.W., and Olkin, I. (1965). "Norms and inequalities for condition numbers." *Pacific J. Math.* 15: 241-247.

15. Marshall, A.W., and Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications.* John Wiley and Sons, New York.

16. Pitman, E. J. G. (1948). "The estimation of location and scale parameters of a continuous population of any given form." *Biometrika* 30: 390-421.

17. Rao, C. R., and Mitra, S. K. (1971). *Generalized Inverse of Matrices and Its Applications*. John Wiley and Sons, New York.

18. Rouanet, H., and Le'pine, D. (1970). "Comparison between treatments in a repeated-measurement design." *British Journal of Mathematical and Statistical Psychology.* 23:147-163.

19. Rohatgi, V. K. (1976). *Introduction to Probability Theory and Mathematical Statistics.* John Wiley and Sons, New York.

20. Schoenberg, I. J. (1938). "Metric spaces and completely monotone functions." *Annals of Mathematics.* 39:811-841.

VITA

Shankar Swamimalai Srinivasan was born in Jamshedpur, India on 18th January 1963. He is the youngest of four brothers. He graduated from Loyola School, Jamshedpur in 1979 and then spent a year in boarding school at Calcutta before entering the Indian Institute of Technology at Madras in 1980. He graduated with a Bachelor of Engineering degree from Madras in June 1985. Following this he worked for a brief while for Tata Iron and Steel Company at Jamshedpur before accepting an invitation for graduate study at Virginia Tech. He received the degrees of Master of Business Administration (August 1990), Master of Science in Statistics (December 1990), and Doctor of Philosophy in Statistics (October 1995) from Virginia Tech.