Development of a statistical procedure for detecting the number of signals in a radar measurement

P.Chen, M.C.Wicks and R.S.Adve

Abstract: Ranking and selection theory is applied to the eigenvalue problem. Of concern is the development of a procedure for computing the number of signals in a measurement data vector. In the authors' approach, the multiplicity of the noise eigenvalue is computed, and used in calculating the number of non-noise (signal) eigenvalues.

1 Introduction

In the analysis of measured data, an approach that is often used involves modelling observations as the superposition of a finite number of signals embedded in additive Gaussian noise. This is especially true in phased array signal processing, time-harmonic analysis, computing the natural response of a system by estimating the number of poles from measurement data, and in detecting overlapping target echoes from radar backscatter. Practical space–time adaptive processing for airborne radar requires effective utilisation of available degrees of freedom. The question therefore arises how many degrees of freedom are required in a given interference scenario. A fundamental issue in solving these problems is correct estimation of the number of signals present.

One approach to solving this problem is based on the observation that the number of signals present can be determined via eigen-analysis of measured data. To do so, an accurate estimate of the covariance matrix of the observed data vector is essential. Once this estimate is formulated, many different techniques are available for eigendecomposition. Bartlett [1] and Lawley [2] developed a multiple hypothesis test for multiplicity of the smallest eigenvalue (latent root) and applied this approach to the analysis of measured agriculture data. Schmidt [3] applied the multiple signal classification (MUSIC) algorithm to estimate the number of incident wavefronts present in an electromagnetic signal, based upon the eigenstructure of the covariance matrix of received data. Other hypothesis testing and estimation methods based on eigenstructure analysis have been proposed by Wax and Kailath [4] and Zhao, Krishnaiah and Bai [5].

This paper uses statistical selection theory to detect the multiplicity of the smallest eigenvalue of the covariance matrix, computed using measured multichannel multipulse radar data. A unique aspect of the proposed methodology is that it predicts a confidence level in the estimated number of signals. As described in the aforementioned articles, the number of signals present is the difference between the total number of components in the observed data vector and the multiplicity of the smallest eigenvalue. In the analysis of measured data, the smallest eigenvalues may be grouped about some nominal value, as opposed to being identically equal. We propose a selection procedure to estimate the multiplicity and value of the smallest eigenvalue(s), which are significantly smaller than the other eigenvalues. We derive the probability of a correct selection, P(CS), and the least favourable configuration (LFC) for our procedures. Under the LFC, the P(CS) attains its minimum over the vector space of all eigenstructures. Therefore, a minimum sample size can be determined from the probability of CS under the LFC, P(CS|LFC), in order to implement our new procedure with a guaranteed probability requirement. Numerical examples are presented in order to illustrate our proposed procedure.

The techniques described above can be applied to the analysis of measured data collected from any multichannel/multipulse radar. As such, a new solution to the adaptive beam-forming problem arises out of the application of ranking and selection theory to the radar problem. First, the number of interfering signals present in a data vector is estimated using our new procedure. Then, optimal rank reduction can be achieved given this knowledge. And finally, adaptive processing for interference rejection and target detection can be performed using any of the standard techniques published in the literature (Reed, Kelly). This technique for estimating the number of signals in noise using statistical selection theory has applications to many other areas where eigenanalysis is useful. Note that, in this paper, correct selection includes overestimating the number of signals. This is of particular importance in the radar signal processing problem [6]. The techniques discussed in this paper, presented within the context of the radar problem, may be generalised. Targeted approaches include multiple discriminant analysis, simultaneous inferencing, principal component analysis, and canonical correlation analysis and multivariate analysis of variance. As such, the analysis of economic, educational, industrial, population, psychological, and scientific data may all benefit from this new technique.

IEE Proceedings online no. 20010420

DOI: 10.1049/ip-rsn:20010420

Paper first received 25th July 2000 and in revised form 20th March 2001

P. Chen is with the Department of Mathematics, Syracuse University, Syracuse, NY 13244-1150, USA

M.C. Wicks is with the Radar Signal Processing Branch, Air Force Research Laboratory, Rome, NY 13440, USA

R.S. Adve is with the Department of Electronic Engineering, University of Toronto, Toronto, Ontario, Canada M5S 3G4

2 Motivation and statement of the problem

As described [4, 5], the observed vector of certain signal processing problems, denoted by the $p \times 1$ vector $\boldsymbol{x}(t)$ can be written as:

$$\boldsymbol{x}(t) = \sum_{1}^{q} \boldsymbol{A}(\Phi_i)s_i(t) + \boldsymbol{n}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{n}(t) \qquad (1)$$

where

$$\boldsymbol{A} = [\boldsymbol{A}(\Phi_1), \ldots, \boldsymbol{A}(\Phi_q)],$$
$$\boldsymbol{s}(t) = (s_1(t) \ldots s_q(t))',$$
$$\boldsymbol{n}(t) = (n_1(t), \ldots, n_p(t))'$$

and $q < p$. It can be seen in eqn. 1 of Wax and Kailath [4] and in eqn. 2.1 of Zhao, Krishnaiah, and Bai [5], that the extreme case where $q = p$ is not realistic and therefore not addressed.

In the above model, $\boldsymbol{n}(t)$ is a $p \times 1$ complex vector referred to as the additive noise distributed independently of $\boldsymbol{s}(t)$ as complex multivariate normal with mean vector $\boldsymbol{0}$ and covariance matrix $\sigma^2 I_p$ where σ^2 is unknown; $\boldsymbol{s}(t)$ is distributed as complex multivariate normal with mean vector $\boldsymbol{0}$ and nonsingular covariance matrix ψ where $s_i(t)$ is a scalar complex waveform associated with the ith signal; and $\boldsymbol{A}(\Phi_i)$ is a $p \times 1$ complex vector, characterised by an unknown parameter vector Φ_i associated with the ith signal. A crucial problem associated with the model described in eqn. 1 and considered by all the articles mentioned in Section 1 is that of determining the number of signals q from a sample $\boldsymbol{x}(t_1), \boldsymbol{x}(t_2), \ldots, \boldsymbol{x}(t_n)$. The goal of this paper is to study formulations in statistical ranking and selection theory to determine the value q.

The covariance matrix $\boldsymbol{\Sigma}$ of $x(t)$ is given by:

$$\boldsymbol{\Sigma} = \boldsymbol{A}\Psi\bar{\boldsymbol{A}}' + \sigma^2 I_p \qquad (2)$$

where $\bar{\boldsymbol{A}}'$ denotes the conjugate transpose of $\boldsymbol{A}$.

Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$ denote the positive eigenvalues of the covariance matrix $\boldsymbol{\Sigma}$ and let the hypothesis H_q: $\lambda_i := \sigma^2 + \alpha_i$ $(i = 1, 2, \ldots, q)$; λ_{q+j}: σ^2 $(j = 1, 2, \ldots, p - q)$. Therefore H_q is equivalent to the hypothesis that q signals are transmitted. Wax and Kailath [4] used Akaike's information criterion (AIC) and Schwart Rissanen's MDL criterion for model selection, while Zhao, Krishnaiah, and Bai [5] used an information theoretic criterion to estimate the value q. Haimovich [7] used asymptotic theory to estimate the convariance matrix $\boldsymbol{\Sigma}$ in eqn. 2 under H_q. In Section 3, we define a selection formulation and propose selection procedures to determine the value q.

3 Ranking and selection formulation and proposed procedure

Ranking and selection procedures are generally developed using either an indifference zone or a subset selection approach. The literature on ranking and selection theory is dominated by these two methods. Wicks [8] first proposed applying statistical ranking and selection theory to radar signal processing for covariance matrix estimation. Chen, Melvin, and Wicks [9] used a variation of the subset selection approach in developing a screening procedure for choosing secondary data in radar signal processing. Their results showed dramatically improved performance over conventional techniques.

Consider k populations $\pi_1, \pi_2, \ldots, \pi_k$ where the underlying distribution of π_i is F_{θ_i}, $i = 1, 2, \ldots, k$. In most ranking and selection studies, the k populations are independent. They could also be k-correlated components of a multivariate population, which possesses a multivariate distribution with parameters of interest. As will become clear later in this paper, we are dealing with p correlated sample eigenvalues and our parameters are the population eigenvalues. In general, the unknown real parameter, θ_i, $i = 1, 2, \ldots, k$, represents the value of a quantity of interest for the ith population. By definition, we select population π_i over π_j if θ_i is greater than θ_j. The ordered values of θ_i for all i are denoted by $\theta_{[1]} < \theta_{[2]} < \theta_{[k]}$. In general, $\theta_{[i]} \neq \theta_i$. One approach to solving the basic problem of selecting the best population, called the indifference zone formulation, was developed in Bechhofer [10]. In Bechhofer's paper, the selection of the population associated with the ranked parameter $\theta_{[k]}$ results in a correct selection (CS). For the indifference zone approach to be of value, the procedure R must establish a lower bound on the probability of a correct selection P(CS). The minimum value of P(CS) is P^*, with $1/k < P^* < 1$ whenever the separation between $\theta_{[k]}$ and $\theta_{[k-1]}$ exceeds some minimum specified value. Let $\delta(\theta_i, \theta_j)$ denote an appropriate non-negative measure of the separation between the population associated with θ_i and θ_j. For the minimum probability of correct selection, P^*, δ^* is the minimum separation distance. For any specified $\delta^* > 0$, let the preference zone, Ω_{δ^*} be the subset of the parameter space $\Omega = \{\tilde{\theta} \mid \tilde{\theta} = (\theta_1, \ldots, \theta_k)\}$ defined by $\Omega_{\delta^*} = \{\tilde{\theta} \mid \delta(\theta_{[k]}, \theta_{[k-1]}) \geq \delta^*\}$. Let P(CS|R) denote the probability of a correct selection under the procedure R. For this procedure to be valid, it should satisfy $P(CS|R) \geq P^*$ for all $\tilde{\theta} \in \Omega_{\delta^*}$. The complement of the preference zone Ω_{δ^*} is called the indifference zone, a subset of the parameter space where no requirement on P(CS) is made.

For the analysis of measured data, δ^* and P^* are specified in advance. Suppose that the procedure R is based on samples of fixed size n from each population. One problem of practical interest in radar signal processing is to determine the smallest sample size n for which the probability requirement P^* holds. In the subset selection approach of Gupta [11], a procedure was developed to guarantee a non-empty subset of the k given populations which include the desired (or best) population with a minimum probability P^*. Any subset, which includes the desired population, results in a correct selection. In case of a tie, any contender may be tagged best. Any valid procedure R should satisfy $P(CS|R) \geq P^*$ for all $\tilde{\theta} \in \Omega$. In the subset selection approach, the size of the selected subset S is not decided in advance, but is determined based on the analysis of data. The procedures developed in ranking and selection theory are designed to satisfy the requirement for a minimum probability of a correct selection P^*. Any parameter configuration $\tilde{\theta}$ which yields the infimum of the P(CS) over Ω_{δ^*} in the indifference zone approach, or Ω in the subset selection approach, is called the least favourable configuration (LFC).

Many variations and generalisations of these two basic approaches have been studied. For example, one problem involves procedures for selecting the most appropriate sample populations better than a control population π_0. These sample populations may then be used to estimate other parameters of interest such as the covariance matrix. In our study of selection procedures for analysing the eigenvalues of the covariance matrix in radar data, the control population can be taken as the smallest eigenvalue. The observations are taken and their covariance matrix is

estimated. Eigenvalues are selected from those populations (eigenvalues in our study) $\pi_1, \pi_2, \ldots, \pi_p$ (instead of using k in traditional selection theory, we use p here to represent the total number of components in a random vector) having the same or similar values as the control population.

We define two disjoint and exhaustive sets Ω_G and Ω_B of the set $\Omega = \{\lambda_1, \ldots, \lambda_p\}$ by using ratio as the distance function d. That is, we define

$$d(\lambda_i, \lambda_j) = \lambda_i/\lambda_j$$

$$\Omega_G = \{\lambda_i, i = 2, 3, \ldots, p \,|\, d(\lambda_i, \lambda_p) \geq \delta^*\}$$

and

$$\Omega_B = \Omega - \Omega_G \tag{3}$$

where $\delta^* > 1$ is a preassigned real number used to differentiate between good and bad eigenvalues. Our goal is to separate the set of eigenvalues into two disjoint subsets, S_G and S_B. The separation is correct (CS) if $S_G = \Omega_G$, meaning that all eigenvalues with values significantly larger than the smallest eigenvalue will be classified into Ω_G. Our conclusion for the value of q is the number of elements in S_G. We require a procedure R that will satisfy a predetermined probability requirement $P(CS|R) \geq P^*$.

Procedure R: Compute the covariance matrix

$$\boldsymbol{S} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x}(t_i)\boldsymbol{x}'(t_i)$$

using the samples $\boldsymbol{x}(t_1), \boldsymbol{x}(t_2), \ldots, \boldsymbol{x}(t_n)$. Let $\lambda_1 > \lambda_2 > \cdots > \lambda_p$ be the ordered eigenvalues of $\boldsymbol{S}$. Let r be the largest integer in $\{1, 2, \ldots, p-1\}$ such that $\lambda_r/\lambda_p > c$, where $c > 1$ is a real number chosen to satisfy the probability requirement $P(CS) \geq P^*$. Claim that $S_G = \{\lambda_1, \lambda_2, \ldots, \lambda_r\}$ and the number of signals is $q = r$. When $\lambda_r/\lambda_p \leq c$ for all integer in $\{1, 2, \ldots, p-1\}$, we claim that $q = 0$. We will explain how to obtain a conservative approximation for c, the procedure parameter, in the next Section.

We make the following two assumptions about the model:

Assumption 1: H_q: $\lambda_i = \sigma^2 + \theta_i$ $(i = 1, 2, \ldots, q)$; λ_{q+j}: σ^2 $(j = 1, 2, \ldots, p-q)$. That is, the multiplicity of the smallest eigenvalue is $p - q$, where p is known and q is unknown. Moreover, we assume that $\theta_1 > \theta_2 \cdots > \theta_q$. This is a reasonable assumption because $\theta_1, \theta_2, \ldots, \theta_q$ came from the first term of the covariance matrix in eqn. 2. Adding ε to an eigenvalue and then letting it go to zero will not change the form of the covariance matrix and therefore it will not change our results.

Assumption 2: When $q > 1$, the parameter configuration belongs to Ω_G, the so-called preference zone in ranking and selection theory. That is, $\lambda_q/\lambda_p \geq \delta^*$. The case $q = 0$ corresponds to the case where there is no signal. The only configuration for the parameter is the equal parameter configuration $\lambda_1/\lambda_p = 1$.

The probability of a correct selection under H_q using procedure R can be written as:

$$P(CS|R) = P(q \leq r) = \sum_{j=q}^{p-1} P(\lambda_j/\lambda_p > c, \lambda_{j+1}/\lambda_p \leq c) \tag{4}$$

To determine the sample size needed to achieve certain probability requirement, we need to minimise the $P(CS|R)$ over the parameter space $\{(\lambda_1, \lambda_2, \ldots, \lambda_p) | \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_q > \lambda_{q+1} = \cdots = \lambda_p > 0\}$. The distributions of the ordered sample eigenvalues $\lambda_1 > \lambda_2 > \cdots > \lambda_p > 0$ of the sample covariance matrix play a very important role in principal component analysis, multiple discriminant analysis, simultaneous inferences, multivariate analysis of variance, and canonical correlation analysis. Many statisticians have studied them extensively in numerous articles. Zonal polynomial expressions of the exact distributions of eigenvalues for both real and complex cases can be found in James [12]. It is clear from formulas (58) and (95) in James [12] that the exact distributions of real Wishart and complex Wishart have the same form and the distribution in both cases depend on the population covariance matrix only through their eigenvalues which can take only positive real values. Therefore, for simplicity in expressions and derivations, we will consider only the real case below. Thorough summaries of sample eigenvalues can also be found in Chapters 11 and 13 of Anderson [13] and in Chapters 3, 7, and 9 of Muirhead [14]. As one can see from the density given in theorem 13.3.2 of Anderson [13] and theorem 9.4.1 of Muirhead [14], the exact computation of the probability in (2.4) which involves the joint density of λ_i $(i = 1, 2, \ldots, p)$ is almost impossible. In the following, we will first prove an important result about the least favourable configuration (LFC) of our procedure in asymptotic theory. Then we will express P(CS) under the asymptotic LFC. In Section 4, we will describe how the results in this section can be used to determine the sample size needed for our procedure R.

The following two lemmas were shown by Anderson [15].

Lemma 1: Let V be distributed according to $W_p(\Sigma, n)$ and let $\lambda_1 > \lambda_2 > \cdots > \lambda_p > 0$ and $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \hat{\lambda}_p > 0$ be the ordered eigenvalues of V and Σ, respectively. Let $\lambda = (\lambda_1, \ldots, \lambda_p)$ and $\hat{\lambda} = (\hat{\lambda}_1, \ldots, \hat{\lambda}_p)$. If λ_i s are distinct and $n \geq p$, then $\sqrt{n}(\lambda/n - \hat{\lambda})$ is asymptotically distributed as normal with mean $\mathbf{0}$ and covariance matrix:

$$Asy - \mathrm{COV}(\sqrt{n}(\lambda/n - \hat{\lambda})) = 2D_{\hat{\lambda}}^2 = 2((\mathrm{diag}(\hat{\lambda}))^2$$

Lemma 2: Suppose that $\theta_1, \theta_2, \ldots, \theta_m$ are the eigenvalues of Σ with multiplicity $q_1, q_2, \ldots, q_m$, respectively, where the sum of q s is p. Then the eigenvalues of V belonging to different eigenvalues of Σ are asymptotically independent and the limiting joint density function of $y_j = \sqrt{(n/2)}(\lambda_j - \theta_\alpha)/\theta_\alpha$, $j = 1, 2, \ldots, q_\alpha$ for all λ_j s belonging to θ_α is given by

$$f_{q_\alpha}(y_1, y_2, \ldots, y_{q_\alpha}) = \pi^{(1/4)q_\alpha(q_\alpha-1)}(2^{(1/2)q_\alpha}\Gamma_{q_\alpha}(q_\alpha/2))^{-1} \times \exp\left(-\frac{1}{2}\sum_j y_j^2\right)\prod_{i<j}(y_i - y_j) \tag{5}$$

Theorem 1: The asymptotic least favourable configuration (ALFC) for our procedure R under the preference zone defined in assumption 2 is given by:

$$\lambda_1 = \lambda_2 = \cdots = \lambda_q > \lambda_{q+1} = \cdots = \lambda_p, \tag{6}$$

where q is an integer between 1 and p and $d(\lambda_1, \lambda_p) = \lambda_1/\lambda_p = \delta^*$.

Proof: Consider a general configuration in the preference zone, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_q > \lambda_{q+1} = \cdots = \lambda_p > 0$ where $\lambda_q/\lambda_p \geq \delta^*$. From lemma 2, $\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_q$ are independent among themselves and independent of $(\hat{\lambda}_{q+1}, \ldots, \hat{\lambda}_p)$

asymptotically. Therefore we can write P(CS|R) in eqn. 4 as

$$Asy - \mathrm{P(CS|R)} = Asy - \sum_{j=q}^{p-1} \mathrm{P}(\lambda_j/\lambda_p > c, \lambda_{j+1}/\lambda_p \leq c) \tag{7}$$

A typical term in the above sum is

$$\mathrm{P}(\lambda_1 > \lambda_2 > \cdots > \lambda_q > \lambda_{q+1} > \cdots > \lambda_j > c\lambda_p > \lambda_{j+1} > \cdots > \lambda_{p-1} > \lambda_p) \tag{8}$$

where $j \geq q$.

Case 1: $j > q$: Consider the transformation

$$z_i = \sqrt{\frac{n}{2}}\frac{\lambda_i - \lambda_i}{\lambda_i} \quad (i = 1, 2, \ldots, p)$$

Then the event $\{\lambda_\alpha > \lambda_{\alpha+1}\}$ is equivalent to

$$\left\{\frac{\sqrt{\frac{2}{n}}z_\alpha + 1}{\sqrt{\frac{2}{n}}z_{\alpha+1} + 1} > \frac{\lambda_{\alpha+1}}{\lambda_\alpha}\right\}$$

According to lemmas 1 and 2, the joint distribution of $z_1, z_2, \ldots, z_p$ is independent of the parameters $\lambda_1, \lambda_2, \ldots, \lambda_p$. The probability given in eqn. 8 can be rewritten as:

$$\mathrm{P}(\lambda_1 > \lambda_2, \lambda_2 > \lambda_3, \ldots, \lambda_q > \lambda_{q+1}, \ldots, \lambda_j > c\lambda_p, c\lambda_p > \lambda_{j+1}, \ldots, \lambda_{p-1} > \lambda_p) =$$

$$\mathrm{P}\left(\frac{\sqrt{\frac{2}{n}}z_1 + 1}{\sqrt{\frac{2}{n}}z_2 + 1} > \frac{\lambda_2}{\lambda_1}, \frac{\sqrt{\frac{2}{n}}z_2 + 1}{\sqrt{\frac{2}{n}}z_3 + 1} > \frac{\lambda_3}{\lambda_2}, \ldots, \frac{\sqrt{\frac{2}{n}}z_q + 1}{\sqrt{\frac{2}{n}}z_{q+1} + 1} > \frac{\lambda_{q+1}}{\lambda_q}, \ldots, \lambda_j > c\lambda_p, c\lambda_p > \lambda_{j+1}, \ldots, \lambda_{p-1} > \lambda_p\right) \tag{9}$$

where the second part of the event in eqn. 9, $\{\lambda_j > c\lambda_p, c\lambda_p > \lambda_{j+1}, \ldots, \lambda_{p-1} > \lambda_p\}$ does not depend on the parameters $\lambda_1, \lambda_2, \ldots, \lambda_q$. We first fix the ratios $\lambda_2/\lambda_1, \lambda_3/\lambda_2, \ldots, \lambda_q/\lambda_{q-1}$ in the right-hand side of eqn. 9. The probability decreases as we decrease the parameter λ_q to its boundary $\delta^*\lambda_p$. Next, we fix the ratios $\lambda_2/\lambda_1, \lambda_3/\lambda_2, \ldots, \lambda_{q-1}/\lambda_{q-2}$. The probability decreases as we decrease the parameter λ_{q-1} to its boundary $\delta^*\lambda_p$. We repeat the above process until all the λ_i s, $i = 1, 2, \ldots, q$ are reduced to $\delta^*\lambda_p$.

Case 2: $j = q$: The only event in this case that is different from case 1 is that $\{\lambda_q > c\lambda_p\}$. It is equivalent to

$$\left\{\frac{\sqrt{\frac{2}{n}}z_q + 1}{\sqrt{\frac{2}{n}}z_p + 1} > c\frac{\lambda_p}{\lambda_q}\right\}$$

We can proceed as in case 1 to obtain the desired result. This completes the proof of the theorem.

Corollary 1: Under the asymptotic least favourable configuration $\lambda_1 = \lambda_2 = \cdots = \lambda_q > \lambda_{q+1} = \cdots = \lambda_p$ where $\lambda_1/\lambda_p = \delta^* > 1$, *Asy*-P(CS) is a decreasing function of δ^*.

Proof: It is clear from eqn. 9 that P(CS) is a decreasing function of any of the largest q eigenvalues.

Theorem 2: Under the asymptotic least favorable configuration given in Theorem 3.1, *Asy*-P(CS) is an increasing function of n, the sample size.

Proof: As in the proof of theorem 1, we consider a typical term in P(CS) given in eqn. 7 and we consider two cases separately.

Case 1: $j > q$: A typical term in P(CS|*asy*-LFC) can be written as

$$\mathrm{P}(\lambda_1 > \lambda_2 > \cdots > \lambda_q > \lambda_{q+1} > \cdots > \lambda_j > c\lambda_p > \lambda_{j+1} > \cdots > \lambda_{p-1} > \lambda_p) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{g_3}\cdots\int_{-\infty}^{z_{j+2}}\int_{-\infty}^{g_2}\int_{-\infty}^{g_1}\cdots\int_{-\infty}^{z_{p-3}}\int_{z_p}^{z_{p-2}} \times f dz_{p-1}dz_{p-2}\ldots dz_{j+1}dz_p dz_j \ldots dz_{q+1}dz_q \ldots dz_1 \tag{10}$$

where f is the joint density function of $z_1, z_2, \ldots, z_p$ which is independent of $\lambda_1, \lambda_2, \ldots, \lambda_p$ and g s are functions of n defined as

$$g_1(n) = cz_p + \sqrt{\frac{n}{2}}(c-1) \qquad g_2(n) = \frac{1}{c}z_j + \sqrt{\frac{n}{2}}\left(1 - \frac{1}{c}\right)$$

$$g_3(n) = z_q + \sqrt{\frac{n}{2}}(\delta^* - 1) \tag{11}$$

Since $0 < c < \delta^*$, all the g s are increasing in n. Therefore, P(CS|*asy*-LFC) is increasing function of n.

Case 2: $j = q$: A typical term in P(CS|*asy*-LFC) can be written as:

$$\mathrm{P}(l_1 > l_2 > \cdots > l_q > cl_p > l_{q+1} > \cdots > l_{p-1} > l_p) = \int_{-\infty}^{\infty}\int_{-\infty}^{z_1}\cdots\int_{-\infty}^{z_{q-1}}\int_{-\infty}^{g_4}\int_{-\infty}^{g_1}\cdots\int_{-\infty}^{z_{p-3}}\int_{z_r}^{z_{p-2}} \times f dz_{p-1}dz_{p-2}\ldots dz_{q+1}dz_p dz_q \ldots dz_2 dz_1$$

where

$$g_1(n) = cz_p + \sqrt{\frac{n}{2}}(c-1)$$

as in eqn. 11 and

$$g_4(n) = \frac{\delta^*}{c}z_q + \sqrt{\frac{n}{2}}\left(\frac{\delta^*}{c} - 1\right) \tag{12}$$

It is clear that $g_4(n)$ is also an increasing function of n. Therefore, P(CS|*asy*-LFC) is increasing function of n. This completes the proof of the theorem.

From the above theorems and the corollary, we know that the probability of a correct selection for our procedure approaches 1 by either increasing δ^*, the size of our 'indifference zone' or the sample size n. It is also clear that P(CS) increases as c decreases. But, the size of the selected subset of the eigenvalues will also increase. Therefore, we may overestimate the number of signals by decreasing c. In the next Section, we will discuss the method we use to find the procedure parameter c and the role that the sample size n plays in the procedure.

4 Approximation of the procedure parameter

In this Section, we explain how to approximate and obtain a conservative bound for c, the procedure parameter. From theorem 1 and the fact that the distributions of the sample eigenvalues depend on the population covariance matrix only through their eigenvalues, the *Asy*-LFC that we use to calculate P(CS) can be written as

$$\Sigma = \text{diag}(a, \ldots, a, a\delta^*, \ldots, a\delta^*) \tag{13}$$

where a is a positive real number.

It is clear that we can take $a = 1$ in eqn. 13 since our procedure R is defined by the selection statistic λ_r/λ_p $(r = 1, \ldots, p-1)$ whose distribution is invariant under a scalar multiplication. From James [16], the probability density function of the sample eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_p$ can be written as

$$\left(\frac{n}{2}\right)^{pn/2} \frac{\pi^{p^2/2}(\det \Sigma)^{-n/2}}{\Gamma_p(\frac{1}{2}n)\Gamma_p(\frac{1}{2}p)} \prod_{i=1}^{p} \lambda_i^{(n-p-1)/2} \times \prod_{i<j}^{p} (\lambda_i - \lambda_j) {}_0F_0^{(p)}(-\tfrac{1}{2}nL, \Sigma^{-1}) \tag{14}$$

where $L = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p)$ and ${}_0F_0^{(p)}$ is a generalised hypergeometric function of a pth-order matrix argument. James [16] also expressed the PDF in terms of zonal polynomials. The computation involving density (eqn. 14) is complicated. In 1970s, many statisticians made contributions to asymptotic expansion and numerical evaluations involving the distribution of functions of the eigenvalues of covariance matrix. For example, Sugiyama [17] gave expressions for the PDF of the ratio λ_1/λ_p. Sugiura [18] gave expressions for the joint PDF of $(\lambda_1, \lambda_2, \ldots, \lambda_p)$. Clemm, Krishnaiah, and Waikar [19] tabulated the percentage points for the largest and the smallest eigenvalues. Waikar and Schuurmann [20] and Krishnaiah and Schuurmann [21] obtained alternative forms for the CDF of λ_1/λ_p to compute the percentage points of the ratio. The computations of all the tables in the aforementioned articles are based on the assumption that the covariance matrix is an identity matrix. Here, in our research, the covariance matrix is diagonal. But, the eigenvalues are slippage. That is, in addition to the value of unity that the smallest eigenvalues takes, the larger eigenvalues take a value of δ^*. Instead of finding the asymptotic expansion and evaluating of the exact probability of correct selection as in the previously mentioned articles, we calculate P(CS|R, A*sy*-LFC) directly using simulation.

In this paper, we use MATLAB version 5.0 on a PC Pentium Pro to compute the eigenvalues of a Wishart distribution generated from a random sample of multivariate Gaussian distribution. The simulation was accomplished using MVNRND and 10 000 repetitions. We first compared the percentage points of the joint distribution of the largest and the smallest obtained from our simulation to those by Clemm, Krishnaiah, and Waikar [19]. We checked 5%, 2.5%, 1%, 0.5% percentage points for $p = 2, 5, 10, 20$ (the largest p tabulated in Clemm, Krishnaiah, and Waikar [19]) and $n = 5, 10, 20, 50$ (the largest n tabulated in Clemm, Krishnaiah, and Waikar [19]). All our simulation results are within 5% of their exact values. Next, we compared the percentage points of the ratio of the smallest root to the largest root obtained in our simulation to those in Krishnaiah and Schuurmann [21]. We checked the 0.99 and 0.95 percentage points for $p = 3, 4, 5$ (the largest p tabulated in Krishnaiah and Schuurmann [21]) and $n = 10, 30, 50$ (the largest n tabulated in Krishnaiah and Schuurmann [21]). Again, all our simulation results are within 5% of their exact values. However, our simulation results can be extended to much larger p and n in both cases. We conclude that simulation methods are appropriate for calculating P(CS) in our study.

In the next example, we show how to obtain the procedure parameter c by simulation. MATLAB simulation program is used to find c value for p up to 50, n up to 5000, and any $\delta^* > 1$. It is available from the first author at pinchen@syr.edu.

Example: Finding the procedure parameter c:

Case 1: $p = 10$, $n = 200$, $\delta^* = 2$, simulation repetition = 10 000:

Table 1 shows the 5 percentiles for the given statistic λ_q/λ_{10}, $q = 0, \ldots, 9$. The underlined diagonal entries are the 5 percentiles of λ_q/λ_{10} when the q value is the correct number of signals. We choose $c = 1.24$, the 5th percentile of λ_9/λ_{10} rounded to the second decimal place. By doing so, more than 95% of the time our estimated q values are 9, 8, 7, 6, 6, 6, 6, 6, 6, 6 when the true $q = 9, 8, 7, 6, 5, 4, 3, 2, 1, 0$, respectively. $P(\lambda_9/\lambda_{10} > c, \lambda_8/\lambda_{10} \leq c) \sim 0.95$ which is only the first term in the sum of P(CS|R) given in eqn. 4. Therefore, c is clearly a conservative procedure parameter for our method and we overestimate the true number of signals, especially when q is small when c is chosen in this manner. The sample size n and the value of δ^* play important roles in the determination of c. The accuracy of our estimation of q increases as n and δ^* increase as we can see from the following two cases.

Table 1: The 5 percentiles for λ_q/λ_{10} in case 1

q	λ_9/λ_{10}	λ_8/λ_{10}	λ_7/λ_{10}	λ_6/λ_{10}	λ_5/λ_{10}	λ_4/λ_{10}	λ_3/λ_{10}	λ_2/λ_{10}	λ_1/λ_{10}
0	1.0308	1.1111	1.1906	1.2911	1.3791	1.4858	1.5950	1.7149	1.8697
1	1.0316	1.1099	1.2034	1.2867	1.3896	1.4910	1.6109	1.7766	2.5615
2	1.0272	1.1160	1.2061	1.3073	1.4037	1.5359	1.6788	2.3528	2.8001
3	1.0326	1.1137	1.2120	1.3180	1.4280	1.5790	2.1617	2.5029	2.8332
4	1.0329	1.1229	1.2193	1.3362	1.4794	1.9973	2.2989	2.5880	2.9165
5	1.0326	1.1280	1.2246	1.3739	1.8496	2.1157	2.3559	2.6030	2.9110
6	1.0335	1.1438	1.2717	1.7160	1.9542	2.1660	2.3603	2.5998	2.8688
7	1.0428	1.1627	1.5655	1.7828	1.9586	2.1677	2.3500	2.5700	2.8488
8	1.0539	1.4332	1.6123	1.7798	1.9533	2.1125	2.3034	2.4767	2.7523
9	1.2478	1.4225	1.5650	1.6958	1.8530	1.9999	2.1509	2.3423	2.5538

Table 2: The 5 percentiles for λ_q/λ_{10} in case 2

q	λ_9/λ_{10}	λ_8/λ_{10}	λ_7/λ_{10}	λ_6/λ_{10}	λ_5/λ_{10}	λ_4/λ_{10}	λ_3/λ_{10}	λ_2/λ_{10}	λ_1/λ_{10}
0	1.0202	1.0740	1.1310	1.1906	1.2504	1.3146	1.3839	1.4604	1.5546
1	1.0204	1.0731	1.1319	1.1933	1.2582	1.3254	1.4023	1.4964	2.3318
2	1.0220	1.0772	1.1382	1.2017	1.2716	1.3476	1.4412	2.1800	2.4715
3	1.0223	1.0795	1.1427	1.2128	1.2875	1.3801	2.0621	2.3031	2.5379
4	1.0224	1.0828	1.1503	1.2263	1.3162	1.9493	2.1619	2.3533	2.5614
5	1.0234	1.0886	1.1621	1.2523	1.8486	2.0432	2.2089	2.3756	2.5685
6	1.0252	1.0941	1.1819	1.7475	1.9267	2.0725	2.2169	2.3706	2.5589
7	1.0259	1.1081	1.6481	1.8098	1.9410	2.0703	2.2024	2.3474	2.5148
8	1.0330	1.5427	1.6948	1.8150	1.9276	2.0423	2.1601	2.2942	2.4553
9	1.4010	1.5348	1.6392	1.7374	1.8376	1.9372	2.0488	2.1721	2.3259

Table 3: The 5 percentiles for λ_q/λ_{10} in case 3

q	λ_9/λ_{10}	λ_8/λ_{10}	λ_7/λ_{10}	λ_6/λ_{10}	λ_5/λ_{10}	λ_4/λ_{10}	λ_3/λ_{10}	λ_2/λ_{10}	λ_1/λ_{10}
0	1.0294	1.1110	1.1959	1.2865	1.3800	1.4769	1.5868	1.7145	1.8736
1	1.0312	1.1118	1.2014	1.2922	1.3936	1.5006	1.6215	1.7762	4.9594
2	1.0317	1.1151	1.2064	1.3081	1.4126	1.5326	1.6827	4.4681	5.3721
3	1.0323	1.1173	1.2148	1.3200	1.4416	1.5861	4.1288	4.8395	5.5575
4	1.0329	1.1217	1.2239	1.3424	1.4842	3.8002	4.4334	4.9938	5.6465
5	1.0352	1.1290	1.2436	1.3782	3.5117	4.0754	4.5568	5.0733	5.6691
6	1.0361	1.1392	1.2763	3.2335	3.7432	4.1777	4.5986	5.0511	5.6148
7	1.0388	1.1620	2.9891	3.4236	3.8021	4.1696	4.5495	4.9746	5.4989
8	1.0469	2.6996	3.1019	3.4307	3.7491	4.0684	4.4222	4.8143	5.2971
9	2.3532	2.6949	2.9612	3.2386	3.4979	3.7905	4.0778	4.4106	4.8577

Case 2: $p=10$, $n=400$, $\delta^*=2$, simulation repetition = 10 000 (Table 2). Our choice for c is 1.40. The estimate for q is 9, 8, 7, 6, 5, 4, 3, 3, 2, 2, respectively for $q=9, 8, 7, 6, 5, 4, 3, 2, 1, 0$. That is, 95% of the time our estimate for q is exactly the true q value when $q=3, \ldots, 9$. We overestimate q when $q=0$, 1, and 2. Next, we consider the case:

Case 3: $p=10$, $N=200$, $\delta^*=4$, simulation repetition = 10 000 (Table 3). Following the same way as in the previous cases, we choose $c=2.35$. The estimate for q is exactly the true q value for all q. The sample size $n=200$ and $\delta^*=4$ are not unreasonable for $p=10$ in radar applications. The three examples in Wax and Kailath [4] all have $p=7$, $n=100$, and $\delta^*=10$. The three examples in Wax, Shan, and Lailath [22] have $p=9$, $n=200$, $\delta^*=10$; $p=9$, $n=100$, $\delta^*=10$; and $p=15$, $n=6400$, $\delta^*=3.98$, respectively.

5 Illustrative example

We present an example to demonstrate the performance of our method. The example adopts the same sensor array processing model assumed in all the three simulation results given in Wax and Kailath [4], Section VI. For comparison purposes, we use exactly the same model configuration as in Wax and Kailath. That is, the vector of the received signal at the array is given by

$$\boldsymbol{x}(t)=\sum_{k=1}^{q}\boldsymbol{A}(\phi_k)e^{-jh(t)}+\boldsymbol{n}(t) \tag{15}$$

where $\boldsymbol{A}(\phi_k)$ is the $p\times 1$ 'direction vector' of the kth wavefront; $\boldsymbol{A}(\phi_k)^{\mathrm{T}}=[1 \quad e^{-j\tau_k} \ldots e^{-j(q-1)\tau_k}]$ with $\tau_k=\pi\sin\phi_k$; $\eta(\cdot)=$random phase uniformally distributed on $(0, 2\pi)$; $\boldsymbol{n}(\cdot)=$vector of white noise with mean 0 and covariance $\sigma^2 I$. The signal-to-noise ratio, defined as $10\log(1/2\sigma^2)$, is 10 dB. From eqn. 15 (formula 24 in Wax and Kailath [4]), the signals have variance 1. Therefore, we assume that $\delta^*=10$ in our study. We first consider seven sensors ($p=7$) and two sources ($q=2$). Using $n=100$ samples, we simulate radar data according to eqn. 1 and the resulting eigenvalues of the sample covariance matrix are 1.0722, 0.9623, 1.1965, 0.7105, 0.5800, 7.3697, 10.3601. We next consider $p=7$ and $q=3$. Using $n=100$ samples, the eigenvalues of the sample covariance matrix of simulated radar data are 1.0688, 1.1311, 0.7159, 1.5023, 8.2028, 8.5212, 10.6351. The simulated five percentiles of λ_q/λ_7 for $q=0, 1, \ldots, 6$ are in Table 4.

Table 4: The five percentiles of λ_q/λ_7

q	λ_6/λ_7	λ_5/λ_7	λ_4/λ_7	λ_3/λ_7	λ_2/λ_7	λ_1/λ_7
0	1.0454	1.1709	1.3180	1.4908	1.6664	1.9125
1	1.0476	1.1834	1.3447	1.5262	1.7566	11.6115
2	1.0491	1.1900	1.3595	1.5759	9.8242	12.8793
3	1.0522	1.2059	1.4207	8.5851	10.9399	13.3350
4	1.0570	1.2364	7.4859	9.3746	11.1370	13.2385
5	1.0649	6.4364	7.9776	9.4016	10.9034	12.7948
6	5.2116	6.4327	7.5204	8.6702	9.9161	11.5024

Following the method in Section 4, we choose $c = 5.21$. Therefore, by our procedure, we correctly declare that there are two signals in the first case and that there are three signals in the second case. At least 95% of the time, our method will detect correctly the number of signals in this model no matter what the true q is. Moreover, for any sample size n great than or equal to 35, we have more than 95% probability of correctly detecting the number of signals no matter what the true number q is. The value $n = 35$ is obtained by running a search algorithm based on the method used to produce the table above.

6 Example using measured radar data

One of the prime motivations for this research is the application to radar signal processing. This example uses data from the Multi-Channel Airborne Radar Measurements (MCARM) program, a vast collection of airborne radar measurements over many flights with multiple acquisitions during flight. The radar antenna is a 22 (2×11) rectangular array ($p = 22$). In several acquisitions, the transmitter was off resulting in no clutter to mask signals generated by a moving target simulator (MTS) at preset Doppler frequencies.

In this example, the data cube comprising 1408 pulses with the transmitter off were analysed. The MTS transmits 10 signals overall. Fig. 1 shows the MTS signal strength as a function of Doppler frequency. As can be seen, there are 9 signals in a pattern centered at -500 Hz and a strong signal at zero Doppler. Within the MCARM database, this data set matches the model of eqn. 1 with $q = 10$. All 1408 pulses ($n = 1408$) were used to estimate the covariance matrix $\boldsymbol{S}$. As noted before, $p = 22$.

The ratio of the eigenvalues to the smallest estimated eigenvalue is shown in Fig. 2. The ratio of the first eigenvalue is 520 times that of the smallest eigenvalue and is not shown. In this case the value of c for $\delta^* = 2.0$ (3 dB) is found to be $c = 1.577$. The eigenvalues and the ratio to the smallest eigenvalue are shown in Table 5. From Fig. 2 and Table 5, this sets the minimum number of signals with signal-to-noise ratios greater than δ^* at 11. Note that, in a radar signal processing problem, it is far better to overestimate than underestimate the number of components present.

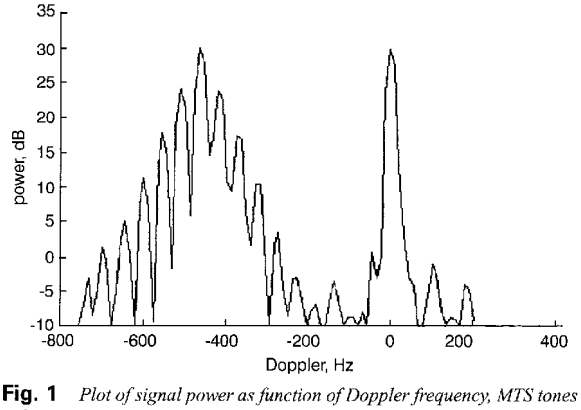

Fig. 1 *Plot of signal power as function of Doppler frequency, MTS tones only*

Angle index: 94; Range cell no. 33

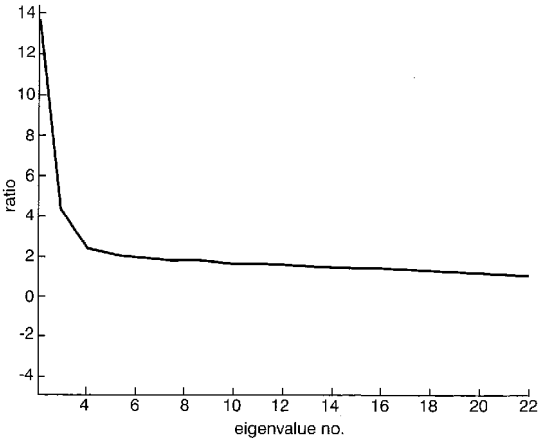

Fig. 2 *Ratio of eigenvalue to smallest estimated eigenvalue, MTS tones only*

Table 5 also lists the AIC and MDL values, used to determine the number of signals in this case, as suggested by Wax and Kailath [4]. This approach used is to determine the AIC or MDL values and the number of signals is the number where the AIC and MDL values are the lowest. As can be seen from Table 5, the AIC criterion sets the number of signals to be 19, while the MDL criterion determines the number of signals to be 4. Both values are erroneous since the true number of signals is 10. In the AIC case, even though it is better to overestimate the number of signals rather than underestimate (the MDL case), the overestimation here is extremely large. In a practical situation, this

Table 5: Statistics for determining number of signals in noise. MCARM data

Number	Eigenvalue	Ratio	AIC value	MDL value
0	4423.7	507.97	148733.9	74366.96
1	123.95	14.233	15253.19	7739.47
2	37.448	4.3002	3458.01	1949.50
3	21.403	2.4577	1636.28	1141.01
4	18.363	2.1086	1337.59	1088.79
5	16.847	1.9345	1214.14	1118.94
6	16.081	1.8466	1154.17	1175.57
7	15.859	1.8211	1113.94	1236.84
8	15.341	1.7616	1064.70	1288.34
9	14.196	1.6301	1021.56	1337.64
10	14.061	1.6147	1015.42	1400.20
11	13.543	1.5552	999.94	1452.83
12	13.199	1.5156	991.36	1503.67
13	12.621	1.4493	982.07	1548.89
14	12.263	1.4082	981.34	1593.16
15	11.944	1.3715	980.55	1632.13
16	11.705	1.4410	977.69	1664.83
17	11.030	1.2666	968.50	1689.11
18	10.708	1.2296	967.66	1712.31
19	9.9416	1.1416	962.57	1728.14
20	9.4735	1.0879	964.99	1742.48
21	8.7084	1	966.00	1750.86

would imply a significant waste of energy and other resources. This example indicates the superiority of the proposed approach over the earlier approaches.

7 Conclusions

In the analysis of measured radar data, a fundamental issue is correct estimation of the number of signals present. One approach to solving this problem is based on the observation that the number of signals present can be determined via eigenanalysis of measured data. To do so, an accurate estimate of the covariance matrix of the observed data vector is essential.

This paper uses statistical selection theory to detect the multiplicity of the smallest eigenvalue of the covariance matrix, computed using measured multi-channel multipulse radar data. The number of signals present is the difference between the total number of components in the observed data vector and the multiplicity of the smallest eigenvalue. In the analysis of measured data, the smallest eigenvalues may be grouped about some nominal value, as opposed to being identically equal. We presented a selection procedure to estimate the multiplicity and value of the smallest eigenvalue(s), which are significantly smaller than the other eigenvalues.

The techniques described above can be applied to the analysis of measured data collected from any multichannel/multipulse radar. This technique for estimating the number of signals in noise using statistical selection theory has applications to many other areas where eigenanalysis is useful. Analysis using both simulated and measured radar data illustrate this new procedure and the enhanced performance over earlier approaches.

In formulating the eigenvalue based statistical approach to determine the number of signals in noise, the noise here is assumed white. In many cases of practical interest, this assumption is not valid and the measurements are corrupted by coloured noise. The formulation of this problem in terms of coloured noise remains an open research problem. However, this does not detract from the effectiveness of the proposed formulation in many other cases of practical interest, including airborne radar, which do satisfy the assumptions of this formulation. Furthermore, as such, the techniques discussed in this paper, presented within the context of the radar problem, may be generalised. Targeted approaches include multiple discriminant analysis, simultaneous inferencing, principal component analysis, and canonical correlation analysis and multivatiate analysis of variance.

8 Acknowledgment

The authors wish to acknowledge the support and interest of Dr Joseph Guerci of the US Defense Advanced Research Projects Agency (DARPA).

9 References

1 BARTLETT, M.S.: 'A note on the multiplying factors for various χ^2 approximations', *J. Roy. Statist. Soc. B*, 1954, **16**, pp. 296–298
2 LAWLEY, D.N.: 'Tests of significance of the latent roots of the covariance and correlation matrices', *Biometrika*, 1956, **43**, pp. 128–136
3 SCHMIDT, R.O.: 'Multiple emitter location and signal parameter estimation'. Proceedings of the RADC spectrum estimation workshop, Rome, NY, 1979, pp. 243–258
4 WAX, M., and KAILATH, T.: 'Detection of signals by information theoretic criteria', *IEEE Trans. Acoust. Speech Signal Process.*, 1985, **ASSP-33**, (2), pp. 387–392
5 ZHAO, L.C., KRISHNAIAH, P.R., and BAI, Z.D.: 'On detection of the number of signals in presence of white noise', *J Multivariate Anal.*, 1986, **20**, (1), pp. 1–25
6 HALE, T., and WELSH, B.: 'Secondary data support in space time adaptive processing'. Proceedings of the 1998 IEEE radar conference, 1998, Dallas, TX, pp. 183–188
7 HAIMOVICH, A.M.: 'Asymptotic distribution of the conditional signal-to-noise ratio in an eigenanalysis-based adaptive array', *IEEE Trans. Aerosp. Electron. Syst.*, 1997, **33**, (3), pp. 988–997
8 WICKS, C.M.: 'Applications of ranking and selection theory to radar signal processing'. Technical report
9 CHEN, P., MELVIN, W.L., and WICKS, M.C.: 'Screening among multivariate normal data', *J. Multivariate Anal.*, 1999, **69**, pp. 10–29
10 BECHHOFER, R.E.: 'A single-sample multiple decision procedure for ranking means of normal populations with known variances', *Ann. Math. Stat.*, 1954, **25**, pp. 16–39
11 GUPTA, S.S.: 'On a decision rule for a problem of ranking means' (Mimeograph. Series No. 150, Institute of Statistics, University of North Carolina, Chapel Hall, NC, 1956)
12 JAMES, A.T.: 'Distributions of matrix variates and latent roots derived from normal samples', *Ann. Math. Stat.*, 1964, **35**, pp. 475–501
13 ANDERSON, T.W.: 'An introduction to multivariate statistical analysis' (Wiley, New York, 1984, 2nd edn.)
14 MUIRHEAD, R.J.: 'Aspects of multivariate statistical theory' (John Wiley Inc., 1982)
15 ANDERSON, T.W.: 'Asymptotic theory for principal component analysis', *Ann. Math. Stat.*, 1963, **34**, pp. 122–148
16 JAMES, A.T.: 'The distribution of the latent roots of the covariance matrix', *Ann. Math. Stat.*, 1960, **31**, pp. 151–158
17 SUGIYAMA, T.: 'Joint distribution of the extreme roots of a covariance matrix', *Ann. Math. Stat.*, 1970, **41**, pp. 655–657
18 SUGIURA, N.: 'Asymptotic expansions of the distributions of the latent roots and the latent vector of the Wishart and multivariate F matrices', *J. Multivariate Anal.*, **6**, pp. 500–525
19 CLEMM, D.S., KRISHNAIAH, P.R., and WAIKAR, V.B.: 'Tables for the extreme roots of the Wishart Matrix', *J. Stat. Comput. Simul.*, 1973, **2**, pp. 65–92
20 WAIKAR, V.B., and SCHUURMANN, F.J.: 'Exact joint density of the largest and smallest roots of the Wishart and MANOVA matrices', *Utilitas Math.*, 1973, **4**, pp. 253–260
21 KRISHNAIAH, P.R., and SCHUURMANN, F.J.: 'On the evaluation of some distributions that arise in simultaneous tests for the equality of the latent roots of the covariance matrix', *J. Multivariate Anal.*, 1974, **1**, pp. 308–315
22 WAX, M., SHAN, T., and KAILATH, T.: 'Spatio-temporal spectral analysis by eigenstructure methods', *IEEE Trans. Acoust. Speech Signal Process.*, 1984, **ASSP-32**, (4), pp. 817–828