Sankhyā : The Indian Journal of Statistics
Special issue in memory of D. Basu
2002, Volume 64, Series A, Pt. 3, pp 653-677

RANK-BASED PROCEDURES FOR STRUCTURAL HYPOTHESES ON COVARIANCE MATRICES

By JOHN MARDEN
University of Illinois, Urbana-Champaign, USA
and
YONGHONG GAO
University of Missouri, Kansas City, USA

SUMMARY. Multivariate sign- and rank-based procedures are developed for structural hypotheses on covariance matrices defined through group symmetries. We show these test statistics are asymptotically chi-square distributed, and calculate their asymptotic relative efficiencies with respect to conventional parametric test statistics. The efficiency calculations and a simulation study suggest that the rank procedures are competitive with the usual normal-theory procedures under normality, and can be quite a bit better under heavy-tailed distributions. The sign procedures are somewhat less efficient, but better at very heavy tails.

1. Introduction

A wide class of structural models on multivariate normal covariance matrices can be expressed using symmetries. For example, a multivariate normal vector has the intraclass correlation structure (where all variances are equal, and all covariances are equal) if the covariance matrix is invariant under permutation of the components, and the components are independent if the covariance matrix in invariant under sign changes of the components. In general, symmetry models are given by requiring the covariance to be invariant under the action of a subgroup of the orthogonal group. Such models include the independent and identically distributed (iid) model, compound symmetry models (which are extensions of the intraclass correlation model),

---

Paper received March 2002.
AMS (1991) *subject classification.* Primary 62H15, 62F35; secondary 62F03.
Keywords and phrases. Multivariate signs, multivariate ranks, structural hypotheses, covariance matrices, intraclass correlation.

spherical symmetry, and models for equality of covariance matrices, independence of blocks of variables, complex normal structure, and quaternion normal structure. Andersson (1975) presents a unified theory for these group symmetry models.

The likelihood procedures given in Andersson and elsewhere in the literature depend crucially on the multivariate normality of the observations. Deviations from normality can significantly affect the level and power of the procedures. In addition, one may not even wish to assume approximate normality, but still consider symmetry models. Marden (1999) presents a generalization of the model beyond normality, and suggests an approach to hypothesis testing based on multivariate signs and ranks. These procedures are relatively simple and can be considered types of "rank transform" methods. In this paper, we present the statistics in detail, show that these statistics are asymptotically chi-squared under the null, find the Pitman efficiency in certain situations, and do a small simulation study. We find that the sign and rank procedures are more robust than the normal-theory procedures, and reasonably efficient, especially the rank procedures.

The (mean zero) invariant normal models are defined as follows. Suppose $X \sim N_p(0, \Sigma)$ (multivariate p-dimensional normal with mean 0 and covariance matrix Σ), and $\mathcal{G}$ is a subgroup of $\mathcal{O}(p)$, the group of $p \times p$ orthogonal matrices. Then the invariant normal model given by $\mathcal{G}$ is

$$\{N_p(0, \Sigma) | \Sigma \text{ such that } \Sigma = g\Sigma g' \text{ for all } g \in \mathcal{G}\}. \tag{1}$$

Different $\mathcal{G}$'s yield different models. Section 2 reviews a number of examples found in the literature.

Andersson (1975) focuses on hypothesis testing problems in which both hypotheses are invariant normal models, with the null model being contained in the alternative model. That is, one tests

$$H_0 : \mathcal{G} = \mathcal{G}_0 \text{ versus } H_1 : \mathcal{G} = \mathcal{G}_1, \tag{2}$$

based on model (1) for $\mathcal{G}_1 \subset \mathcal{G}_0 \subset \mathcal{O}(p)$. Andersson analyses all possible hypothesis testing problems for the normal case. The interesting result is that any such testing problem can be decomposed into a series of smaller testing problems, each one being one of a set of ten types of problems.

Note that the model (1) implies that gX and X have the same distribution for any $g \in \mathcal{G}$. The extension of the normal model we consider depends on $\mathcal{G}$ as above and a $p \times 1$ vector γ. Let $F(\mathcal{G}, \gamma)$ consist of all distribution functions F such that

$$X \sim F \Longrightarrow g(X - \gamma) \text{ and } (X - \gamma) \text{ have the same distribution for any } g \in \mathcal{G}.$$

Then the general group symmetry model is

$$X_1, \ldots, X_n \text{ are iid} \sim \ F \text{ for } F \in F(\mathcal{G}, \gamma). \tag{3}$$

For examples, if the components of $X_i - \eta$ are iid, or just exchangeable, then (3) holds with $\mathcal{G}$ being the group of permutation matrices; if the X_i's are spherically symmetric about η, then $\mathcal{G}$ would be the group of orthogonal matrices; and if the components of $X_i - \eta$ are symmetric about 0, then $\mathcal{G}$ would be the group of diagonal matrices with ± 1's on the diagonal. Note that if model (3) holds, and the covariance matrix of X exists, then it is $\mathcal{G}$-invariant:

$$Cov(X) = Cov(gX) = gCov(X)g' \text{ for any } g \in \mathcal{G}. \tag{4}$$

Marden (1999) proposes using procedures based on multivariate spatial signs and ranks for testing (2) under (3). Section 3 presents the test statistics and their asymptotic distribution under the null hypothesis. Section 4 looks at the local asymptotic power of the test statistics. These results can be used as large-sample approximations to the power, but we use them for calculating relative efficiencies. We present some simulations for two illustrative problems: Testing for zero correlation between two blocks of variables, and testing for intraclass correlation structure. See Sections 5 and 6. The technical results are collected in Section 7.

2. Review of Some Models

Structural models on covariance matrices (and means) go way back, although they were not always presented in the general form (1). Examples include independence of variables, or sets of variables, and equality of variances. In the former, with K sets of variables, the covariance matrix would have structure

$$\Sigma = \begin{pmatrix} \Sigma_{11} & 0 & \ldots & 0 \\ 0 & \Sigma_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \Sigma_{KK} \end{pmatrix}, \tag{5}$$

where the Σ_{kk}'s are unrestricted covariance matrices. The group $\mathcal{G}$ in (1) consists of the matrices

$$\Gamma = \begin{pmatrix} \pm I^{(1)} & 0 & \ldots & 0 \\ 0 & \pm I^{(2)} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \pm I^{(K)} \end{pmatrix}, \tag{6}$$

where $I^{(k)}$ is the identity matrix of the same dimension as Σ_{kk}. When the Σ_{kk}'s are all of the same dimension, Bartlett's problem tests whether the covariances are equal in (5). It uses the group that permutes the blocks of variables as well as changes signs as in (6).

Wilks (1946) considered the intraclass correlation structure, given by (1) with $\mathcal{G}$ being the group of $p \times p$ permutation matrices. For $p = 3$, we would have

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix}. \tag{7}$$

This structure is often assumed for repeated measure designs, including growth curve models. It also arises from the balanced one-way random effect model.

Votaw (1948) looks at an extended model, called compound symmetry, where within sets of variables there is intraclass correlation, and between sets the correlations are equal. For example, with one set of three variables and one of two, the covariance matrix would have the form

$$\Sigma = \begin{pmatrix} a & b & b & c & c \\ b & a & b & c & c \\ b & b & a & c & c \\ c & c & c & d & e \\ c & c & c & e & d \end{pmatrix}. \tag{8}$$

Such models might arise, e.g., if students are given three interchangeable batteries of math questions, and two interchangeable batteries of verbal questions.

Goodmann (1963) looked at the covariance matrix of multivariate complex normals, which arise in spectral analysis of multiple time series. A p-dimensional complex normal is $X^{(1)} + iX^{(2)}$, where $X^{(1)}$ and $X^{(2)}$ are p-dimensional real normals with joint covariance of the form

$$Cov((X^{(1)\prime}, X^{(2)\prime})') = \begin{pmatrix} \Sigma & F \\ -F & \Sigma \end{pmatrix}, \tag{9}$$

i.e., $Cov(X^{(1)}) = Cov(X^{(2)})$ and $Cov(X^{(1)}, X^{(2)})$ is skew-symmetric. The corresponding group is

$$\mathcal{G} = \left\{ I_{2p}, \begin{pmatrix} 0 & -I_p \\ I_p & 0 \end{pmatrix}, \begin{pmatrix} 0 & I_p \\ -I_p & 0 \end{pmatrix} \right\}, \tag{10}$$

where I_q is the $q \times q$ identity matrix.

Olkin and Press (1969) consider a circular stationary model, where variables are thought of as being equally spaced around a circle, and the covariance between two variables depends on just their distance. E.g., for $p = 6$, the covariance matrix looks like

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & a & b & c & b & a \\ a & 1 & a & b & c & b \\ b & a & 1 & a & b & c \\ c & b & a & 1 & a & b \\ b & c & b & a & 1 & a \\ a & b & c & b & a & 1 \end{pmatrix}. \tag{11}$$

The group here is the circular group, generated (when $p = 6$) by

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{12}$$

Some papers, besides those already mentioned, that consider testing problems (2) on the above models include Krishnaiah and Pathak (1967) and Srivastava (1965) on the intraclass correlation models, Lee et al. (1976) on compound symmetry, and Khatri (1965a, 1965b), Krishnaiah et al. (1983), and Andersson and Perlman (1984) on the complex normal. Arnold (1973, 1976) and Andersson (1975) take a more general view by looking at "products" of symmetry problems.

The purpose of the model (3) and procedures introduced in the subsequent sections is two-fold: To present easy-to-use procedures in the normal model that are robust against some deviations from normality, and to extend the symmetry models beyond normality. In the latter case, some of the models change meaning, e.g., the independence model (5) in the normal case is the "zero covariance" case in general, and the complex normal does not necessarily have a non-normal counterpart. Other models, e.g., equality of covariances, intraclass correlation, and compound symmetry retain their meaning outside of normality.

3. Test Statistics

Sign and rank procedures are very popular in univariate statistics, as they often provide robust and easy-to-use procedures. For the structural hypotheses considered here, the usual coordinatewise signs or ranks are not appropriate, because the group structure is destroyed. Instead, rotation-equivariant quantities are necessary. The simplest seem to be the spatial signs and ranks. One could also use the affine-equivariant rank-based estimators as in Visuri, Koivunen, and Oja (2000)

In the univariate setting, the sign function $S(x)$ is $-1, 0, 1$, as $x < 0, x = 0, x > 0$, respectively. Given a sample $x_1, x_2, \ldots, x_n$, the rank of observation x_i among those in the sample is k if x_i is the k^{th} smallest. The rank can be written as

$$rank(x_i) = \frac{1}{2}\left(\sum_{j=1}^{n} S(x_i - x_j) + n + 1\right). \tag{13}$$

A natural multivariate generalization of the sign function is

$$S(x) = \begin{cases} x/\|x\| & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}. \tag{14}$$

In analogy with (13), focusing on just the sign part and dividing by n, the spatial rank of a vector x relative to the data is defined to be

$$R(x; x_1, \ldots, x_n) = \frac{1}{n}\sum_{j=1}^{n} S(x - x_j). \tag{15}$$

Note that the spatial rank of x is the average of the unit vectors pointing from the data vectors to x. The rank of x relative to a distribution function F is defined by

$$R(x; F) = E[S(x - Y)], Y \sim F. \tag{16}$$

Then we can write $R(x; x_1, \ldots, x_n) = R(x; \widehat{F}_n)$, where $\widehat{F}_n$ is the empirical distribution function of the data. These ranks are introduced in Möttönen and Oja (1995) and Chaudhuri (1996).

The test statistics we use are based on the sample covariance matrices of the spatial signs or ranks of the data. The reason such statistics are effective in testing the hypotheses of interest is that if the distribution of the data is invariant under $\mathcal{G}$, then so are the distributions of the (shifted) signs and the ranks.

If γ is known, then the sign tests will be based on the $S(X_i - \gamma)$'s. Because γ is typically not known, we also consider the $S(X_i - \widehat{\gamma})$'s, where

$\widehat{\gamma}$ is the *spatial median* of the data. That is, $\widehat{\gamma}$ minimizes $\sum_{i=1}^{n} \|X_i - \gamma\|$ over γ. See Small (1990) for an excellent review of multivariate medians. This minimizer is unique unless all points lie on a line. The spatial ranks are shift-invariant, so we need not center them. The following table lists the general testing procedures we will be dealing with. We will denote the generic statistic U_i, an appropriately centered quantity for each observation, from which we estimate the relevant parameter:

Procedure	Parameter Σ_U	U_i
Normal-theory	$Cov(X_i)$	$X_i - \overline{X}$
Sign, γ known	$Cov(S(X_i - \gamma))$	$S(X_i - \gamma)$
Sign, γ unknown	$Cov(S(X_i - \gamma))$	$S(X_i - \widehat{\gamma})$
Rank	$Cov(R(X_i; F))$	$R(X_i; \widehat{F}_n)$

The estimate of the parameter in each case is then $\widehat{\Sigma}_U$,

$$\widehat{\Sigma}_U = \frac{1}{n} \sum_{i=1}^{n} U_i U_i'. \tag{17}$$

For the normal-theory test, sign test with γ unknown, and the rank test, the sample mean of the U_i's is 0. For the sign test with known γ, $E(U_i) = 0$.

The main result we need is the following from Marden (1999). It guarantees that the parameters and the expected values of their estimates are $\mathcal{G}$ invariant under (3).

LEMMA 1 *Suppose model (3) holds. Then for each $g \in \mathcal{G}$,*

$$\begin{aligned} (S(g(X_1 - \gamma)), \ldots, S(g(X_n - \gamma))) &=^{\mathcal{D}} (S(X_1 - \gamma), \ldots, S(X_n - \gamma)), \\ (S(gX_1 - \widehat{\gamma}^{(g)}), \ldots, S(gX_n - \widehat{\gamma}^{(g)})) &=^{\mathcal{D}} (S(X_1 - \widehat{\gamma}), \ldots, S(X_n - \widehat{\gamma})), \quad \textit{and} \\ (R(gX_1; \widehat{F}_n^{(g)}), \ldots, R(gX_n; \widehat{F}_n^{(g)})) &=^{\mathcal{D}} (R(X_1; \widehat{F}_n), \ldots, R(X_n; \widehat{F}_n)), \end{aligned} \tag{18}$$

where $\widehat{\gamma}^{(g)}$ is the spatial median and $\widehat{F}_n^{(g)}$ is the empirical distribution function of $gX_1, \ldots, gX_n$.

Consider the hypothesis testing problem (2). The estimates of Σ_U under H_0 and H_1 we use are

$$\widehat{\Sigma}_U(\mathcal{G}_0) = \int_{\mathcal{G}_0} g\widehat{\Sigma}_U g' \nu_0(dg), \text{ and } \widehat{\Sigma}_U(\mathcal{G}_1) = \int_{\mathcal{G}_1} g\widehat{\Sigma}_U g' \nu_1(dg), \tag{19}$$

respectively, where ν_i is Haar probability measure on $\mathcal{G}_i$, $i = 0, 1$. Often, $\mathcal{G}_i$ will be a finite group, so that the integrals are straight averages over the elements $g \in \mathcal{G}_i$. These estimates are motivated by noting that under the normal model, $\widehat{\Sigma}_N(\mathcal{G}_0)$ and $\widehat{\Sigma}_N(\mathcal{G}_1)$ are the maximum likelihood estimates of $Cov(X_i)$ under the null and alternative hypotheses, respectively.

The test statistics for (2) are based on linear functions of the difference

$$\widehat{\Sigma}_U(\mathcal{G}_0) - \widehat{\Sigma}_U(\mathcal{G}_1), \tag{20}$$

where the exact linear functions depend on the testing problem. The test statistic is then a quadratic form in the linear functions, which is asymptotically chi-squared under the null hypothesis if certain conditions hold.

We will show in Section 7 that under appropriate conditions,

$$\sqrt{n}(\widehat{\Sigma}_U - \Sigma_U) \xrightarrow{\mathcal{D}} N(0, \Lambda_U) \tag{21}$$

for some covariance matrix Λ_U. We also need a consistent estimator of Λ_U: $\widehat{\Lambda}_U$ such that

$$\widehat{\Lambda}_U \xrightarrow{\mathcal{P}} \Lambda_U. \tag{22}$$

Let

$$V_i = vec(U_i U_i'), \tag{23}$$

where for a $p \times p$ symmetric matrix A, $vec(A)$ is the $p^2 \times 1$ vector obtained by stacking the columns:

$$vec(A) = (a_{11}, a_{21}, \ldots, a_{p1}, a_{12}, a_{22}, \ldots . a_{p2}, \ldots, a_{1p}, \ldots, a_{pp})'. \tag{24}$$

Now the sample mean of the V_i's equals the vectorized sample covariance matrix of the U_i's:

$$\overline{V} = vec\left(\frac{1}{n}\sum_{i=1}^{n} U_i U_i'\right) = vec(\widehat{\Sigma}_U). \tag{25}$$

For the normal and sign tests, $\widehat{\Lambda}_U$ is simply the sample covariance matrix of the V_i's:

$$\widehat{\Lambda}_U = \frac{1}{n}\sum_{I=1}^{n}(V_i - \overline{V})(V_i - \overline{V})'. \tag{26}$$

The rank case is a little more complicated. Let

$$V_{Ri} = \frac{1}{n^2}\sum_j \sum_k \; vec(S_{ij}S_{ik}' + S_{ji}S_{jk}' + S_{jk}S_{ji}'), \tag{27}$$

where

$$S_{ij} = S(X_i - X_j). \tag{28}$$

It can be shown that

$$\overline{V}_R = \frac{1}{n}\sum_{i=1}^{n} V_{Ri} = 3\ vec(\widehat{\Sigma}_R). \tag{29}$$

Then we take

$$\widehat{\Lambda}_R = \frac{1}{n}\sum_{i=1}^{n}(V_{Ri} - \overline{V}_R)(V_{Ri} - \overline{V}_R)'. \tag{30}$$

See Proposition 4 in Section 7.1.2 for this case.

Any set of linear functions of the difference in (20) can be represented by $C\overline{V}$ for some $t \times p^2$ matrix C. By Lemma 1, we have that

$$C\ vec(\Sigma_U) = 0 \text{ under the null hypothesis.} \tag{31}$$

Suppose further that

$$C\Lambda_U C' \text{ is nonsingular.} \tag{32}$$

Then the test statistic is the Hotelling T^2-like statistic

$$T = n(C\overline{V})'(C\widehat{\Lambda}_U C')^{-1}(C\overline{V}), \tag{33}$$

and we will show in Section 7.1 that under the null,

$$T \xrightarrow{\mathcal{D}} \chi^2_t. \tag{34}$$

4. Local Asymptotic Power

For assessing the power asymptotically, we look at local observations

$$X_{in} = X_i + \frac{1}{\sqrt{n}} A X_i, \tag{35}$$

where the X_i's are iid with distribution in the null hypothesis, and A is a fixed $p \times p$ matrix chosen so that the X_{in}'s have distribution outside of the null. Note that if the covariance matrix Σ_X of X_i exists, then

$$Cov_n(X_{in}) = \Sigma_X + \frac{1}{\sqrt{n}}(A\Sigma_X + \Sigma_X A') + \frac{1}{n}A'\Sigma A, \tag{36}$$

so that in terms of that parameter we are looking at contiguous alternatives of the form $\Sigma + n^{-1/2}\Delta + o(n^{-1/2})$. We find it technically convenient to work directly on the random variables rather than in terms of the parameter. In addition, the asymptotic distributions of the rank and sign tests do not assume the covariance exists.

Let $\widehat{\Sigma}_U^A$ denote the statistic $\widehat{\Sigma}_U$ in (17) but evaluated at the X_{in}'s. Then we can show that

$$\sqrt{n}(\widehat{\Sigma}_U^A - \Sigma_U) \xrightarrow{\mathcal{D}} N(\delta_U, \Lambda_U) \tag{37}$$

for some $\delta_U(= \delta_U(A))$, where U is the normal, sign (with known spatial median), or rank statistic. If C is a matrix as in (31) and (32),

$$\sqrt{n}C\ vec(\widehat{\Sigma}_U^A) \xrightarrow{\mathcal{D}} N(C\delta_U, C\Lambda_U C'), \tag{38}$$

and

$$T \xrightarrow{\mathcal{D}} \chi_t^2(\Delta_U) \quad \text{where} \quad \Delta_U(= \Delta_U(A, C)) \equiv \delta_U' C'(C\Lambda_U C')^{-1} C\delta_U. \tag{39}$$

That is, T is asymptotically noncentral chi-squared with noncentrality parameter Δ_U. See Section 7 for explicit formulae.

We will compare the test statistics on the basis of the Δ_U's, taking the relative efficiency of statistic U to statistic V to be Δ_U/Δ_V.

For presenting some efficiencies, we specialize to distributions $F \in F(\mathcal{G}_p^*, 0)$, where $\mathcal{G}_p^*$ denotes the group of $p \times p$ matrices generated by the permutation matrices combined with the diagonal matrices with ± 1's on the diagonals. The components of $X \sim F$ for such F are symmetric about 0 and permutation invariant. A special case is when $X_{[1]}, \ldots, X_{[p]}$ are independent and identically distributed, and symmetric about 0. (We use square brackets to denote elements of a vector or matrix, so that a $p \times 1$ vector $x = (x_{[1]}, \ldots, x_{[p]})'$.) These F's are contained in any null hypothesis of (2). We also consider invariance under the group of orthogonal matrices, $\mathcal{O}_p$.

All the statistics are invariant under multiplication by a scalar, $X_i \to cX_i$ for $c \neq 0$, so that from now on we will assume $Cov(X) = I_p$ (if the covariance exists), hence $E\|X\|^2 = p$.

It is easier to present the results if we make a particular permutation of the elements of $\widehat{\Sigma}_U$ and all the subsequent matrices. For a $p \times p$ matrix B, define the $p^2 \times 1$ vector $vec^*(B)$ that strings out the diagonal elements, then the elements below the diagonal columnwise, then the elements above the diagonal rowwise:

$$\begin{aligned} vec^*(B) \;=\; & (b_{11}, b_{22}, \ldots, b_{pp}; \\ & b_{21}, b_{31}, \ldots, b_{p1}, b_{32}, b_{42}, \ldots, b_{p2}, \ldots, b_{p(p-1)}; \\ & b_{12}, b_{13}, \ldots, b_{1p}, b_{23}, b_{24}, \ldots, b_{2p}, \ldots, b_{(p-1)p})'. \end{aligned} \tag{40}$$

Now (37) becomes

$$\sqrt{n}(vec^*(\widehat{\Sigma}_U^A) - vec^*(\Sigma_U)) \xrightarrow{\mathcal{D}} N(\delta_U^*, \Lambda_U^*), \tag{41}$$

where $\delta_U^* = vec^*(\delta_U)$ and Λ_U^* is the matrix Λ_U with the rows and columns appropriately permuted. Similarly, let C^* be obtained by rearranging the columns of the matrix C in (33), so that in (39),

$$\Delta_U = (C^*\delta_U^*)'(C^*\Lambda_U C^{*'})^{-1}C^*\delta_U^*. \tag{42}$$

Section 7.2 contains expressions for δ_U in general. Here, we present some special cases that yield compact expressions for the Δ_U's. First, we consider problems for which the C^* contains just two types of linear combinations (rows): Contrasts among the diagonal elements (variances) of $\widehat{\Sigma}_U$, and linear combinations involving just the elements below the diagonal (covariances). Such problems include all those of the form (2) that can be expressed with $\mathcal{G}_0$ (hence $\mathcal{G}_1$) contained in $\mathcal{G}_p^*$. See Remark 1. Most common testing problems, including those mentioned in Section 2, can be so expressed.

Collect the contrasts of the variances in the $t_1 \times p$ matrix C_1, and the linear combinations of the covariances in the $t_2 \times \binom{p}{2}$ matrix C_2, so that

$$C^* = \begin{pmatrix} C_1 & 0 & 0 \\ 0 & C_2 & 0 \end{pmatrix}. \tag{43}$$

Partition $vec^*(A + A')$ into v_1, the first p components, v_2, the next $\binom{p}{2}$ components, and note that by symmetry the final $\binom{p}{2}$ components also form v_2, so that

$$vec^*(A + A') = \begin{pmatrix} v_1 \\ v_2 \\ v_2 \end{pmatrix}. \tag{44}$$

Also, let

$$Q_i = v_i'C_i'(C_iC_i')^{-1}C_iv_i \quad \text{for } i = 1, 2. \tag{45}$$

Differences between the variances are measured by Q_1, and relationships among the covariances are measured by Q_2.

The next theorem presents the calculations when $F \in F(\mathcal{G}_p^*, 0)$ and $F \in F(\mathcal{O}_p, 0)$. The notation U becomes N to denote the normal statistic, S to denote the sign statistic with known γ, and R to denote the rank statistic.

Theorem 1 *Suppose $F \in F(\mathcal{G}_p^*, 0)$, and for the normal statistic, assume that the fourth moments exists and $Cov(X) = I_p$. Then there are constants α_U and β_U, depending on p and F, such that*

$$\Delta_U = \alpha_U Q_1 + \beta_U Q_2. \tag{46}$$

Furthermore, if $F \in F(\mathcal{O}_p, 0)$,

$$\alpha_N = \frac{1}{2}\beta_N, \quad \beta_N = \frac{p(2+p)}{E(\|X\|^4)}, \quad \alpha_S = \frac{1}{2}\beta_S, \quad \text{and} \quad \beta_S = \frac{p}{(2+p)}. \quad \Box \tag{47}$$

The proof is in Section 7.3, along with explicit formulae for the constants.

The theorem shows that if $F \in F(\mathcal{O}_p, 0)$, then the efficiency of the sign test relative to the normal test is independent of C^* and A:

$$\frac{\Delta_S}{\Delta_N} = \frac{E(\|X\|^4)}{(2+p)^2}. \tag{48}$$

If F is $N_p(0, I_p)$, then $E(\|X\|^4) = p(2+p)$, so that $\Delta_S/\Delta_N = p/(2+p)$, which for small p is fairly low, but approaches 1 as p increases. The minimum efficiency is achieved when X is uniformly distributed on the sphere in p space with radius $\sqrt{p}$ (so that $E(\|X\|^2) = p$), which yields the efficiency $p^2/(2+p)^2$.

The efficiency of the rank test relative to the normal in this case does depend on C^* and A, but at least we have the following.

Proposition 1 *Under the assumptions in Theorem 1, if $F \in F(\mathcal{O}_p, 0)$,*

$$\frac{\Delta_R}{\Delta_N} \quad \text{is between} \quad \frac{2\alpha_R E(\|X\|^4)}{p(2+p)} \quad \text{and} \quad \frac{\beta_R E(\|X\|^4)}{p(2+p)}. \tag{49}$$

Proof. From (46) and (47), we can write

$$\frac{\Delta_R}{\Delta_N} = \frac{2\alpha_R(Q_1/2) + \beta_R(Q_2)}{Q_1/2 + Q_2} \, \frac{E(\|X\|^4)}{p(2+p)}. \tag{50}$$

Note that the first ratio is a convex combination of $2\alpha_R$ and β_R, with weights proportional to $Q_1/2$ and Q_2. Thus (49) holds. □

The first bound in (49) is achieved when the testing problem involves only variances, and the second bound is achieved when the testing problem involves only covariances. We do not have a nice analytical form for these quantities, but can estimate them using Monte Carlo. See Table 2, where each value is estimated using one million simulations. It is interesting that the estimated $(2\alpha_R)$'s are very close to the estimated β_R's, so that the two bounds in (49) are quite close.

To be even more explicit, we consider a one-parameter family of spherically symmetric distributions $SS_p(\nu)$ of Randles (1989). Here, $X \sim SS_p(\nu)$ means X has density

$$f(x; p, \nu) = k(p, \nu) exp(-(\|x\|^2/c(p, \nu))^\nu), \tag{51}$$

where

$$c(p, \nu) = \frac{p\Gamma(\frac{p}{2\nu})}{\Gamma(\frac{p+2}{2\nu})}, k(p, \nu) = \frac{\nu\Gamma(\frac{p}{2})}{\Gamma(\frac{p}{2\nu})(\pi c(p, \nu))^{p/2}}. \tag{52}$$

(The constants are chosen so that $Cov(X) = I_p$.) The parameter ν controls the heaviness of the tails, where small values indicate heavy tails, and large values indicate light tails. The normal is obtained by setting $\nu = 1$. Tables 1 (for the Sign test) and 2 (for the Rank test) exhibit some efficiencies relative to the normal test for some choices of $SS_p(\nu)$.

TABLE 1. ASYMPTOTIC RELATIVE EFFICIENCIES FOR SIGN TEST

	$p = 2$	$p = 3$	$p = 5$	$p = 10$
$\nu = 1/10$	54.207	24.069	9.654	3.758
$\nu = 1/5$	3.972	3.081	2.269	1.627
$\nu = 1/2$	0.833	0.900	0.952	0.985
$\nu = 1$ (Normal)	0.500	0.600	0.714	0.833
$\nu = 2$	0.393	0.494	0.620	0.767
$\nu = 5$	0.347	0.445	0.574	0.731

TABLE 2. BOUNDS ON ASYMPTOTIC RELATIVE EFFICIENCIES (ESTIMATED) FOR RANK TEST

	$p = 2$	$p = 3$	$p = 5$	$p = 10$
$\nu = 1/10$	48.48, 49	22.08, 21.64	8.63, 8.59	3.42, 3.42
$\nu = 1/5$	4.61, 4.41	3.25, 3.17	2.28, 2.32	1.62, 1.62
$\nu = 1/2$	1.26, 1.12	1.20, 1.19	1.14, 1.14	1.09, 1.09
$\nu = 1$ (Normal)	0.90, 0.91	0.94, 0.92	0.93, 0.95	0.96, 0.96
$\nu = 2$	0.79, 0.93	0.85, 0.89	0.90, 0.88	0.91, 0.92
$\nu = 5$	0.85, 0.82	0.86, 0.86	0.86, 0.83	0.90, 0.89

Table 2 shows that the rank test is reasonably efficient relative to the normal for normal and light-tailed ($\nu \geq 1$) alternatives, for the most part exceeding 85%, and can be substantially more efficient than the normal for heavy-tailed alternatives. The sign test (Table 1) is not as efficient for the normal and light-tailed alternatives, and only beats the rank test at the heaviest-tailed alternative. Thus these calculations strongly recommend the rank tests.

Remark 1 *When $g \in \mathcal{G}_p^*$, then $g\Sigma g'$ permutes the variances of Σ, and permutes the covariances as well as possibly changing the signs of some of the covariances. Thus if $\mathcal{G} \subset \mathcal{G}_p^*$, $g\Sigma g' = \Sigma$ for all $g \in \mathcal{G}$ can only set some equalities among the variances, set some equalities among the covariances, and set some covariances to zero. So for a testing problem (2) with $\mathcal{G}_1 \subset \mathcal{G}_0 \subset \mathcal{G}_p^*$, the corresponding C^* is of the form (43) in which the rows of C_1 are contrasts.*

5. Some Illustrative Testing Problems

In this section we look more closely at a couple of testing problems: testing that two blocks of variables are uncorrelated, and testing for the intraclass correlation structure in Σ. We use these models for the simulation study in Section 6.

5.1 *Correlation.* The first problem assumes that the X is a $p \times 1$ vector, partitioned as

$$X = \begin{pmatrix} X^{(1)} \\ X^{(2)} \end{pmatrix}, \tag{53}$$

where $X^{(1)}$ is $p_1 \times 1$ and $X^{(2)}$ is $p_2 \times 1$, $p_1 + p_2 = p$. The problem under normality is to test whether $X^{(1)}$ and $X^{(2)}$ are independent. In the extended model (3), the null hypothesis takes the group in (6) with $K = 2$, that is,

$$\mathcal{G}_0 = \left\{ I_p, \begin{pmatrix} I_{p_1} & 0 \\ 0 & -I_{p_2} \end{pmatrix}, \begin{pmatrix} -I_{p_1} & 0 \\ 0 & I_{p_2} \end{pmatrix}, -I_p \right\}. \tag{54}$$

Under that hypothesis, if X has a covariance Σ, then (5) holds where $K = 2$, Σ_{11} is $p_1 \times p_1$ and Σ_{22} is $p_2 \times p_2$. The alternative is the general alternative that places no restriction on the distribution, that is, $\mathcal{G}_1 = \{I_p\}$.

Under the alternative hypothesis, the estimates of the covariance matrices are just the raw sample covariance matrices. The estimates under the null hypotheses are the same, but set the cross-product matrix to 0. That is, partition $\hat{\Sigma}_U$ as Σ:

$$\hat{\Sigma}_U = \begin{pmatrix} \hat{\Sigma}_U^{(11)} & \hat{\Sigma}_U^{(12)} \\ \hat{\Sigma}_U^{(21)} & \hat{\Sigma}_U^{(22)} \end{pmatrix}. \tag{55}$$

Then

$$\hat{\Sigma}_U(\mathcal{G}_0) = \begin{pmatrix} \hat{\Sigma}_U^{(11)} & 0 \\ 0 & \hat{\Sigma}_U^{(22)} \end{pmatrix}. \tag{56}$$

Thus the difference (20) between the estimates under the null and alternative hypotheses is equivalent to $\hat{\Sigma}_U^{(21)}$. We choose C to pick off the $\hat{\Sigma}_U^{(21)}$ part of $\hat{\Sigma}$, so that

$$C\bar{V} = vec(\hat{\Sigma}_U^{(21)}). \tag{57}$$

The C^* in (43) then does not need the C_1, and the $Q_1 = 0$ in (46).

Proposition 2 in Section 7 shows that the statistic (33) is asymptotically $\chi^2_{p_1 p_2}$ under the null hypothesis.

5.2 *Intraclass Correlation* A covariance matrix exhibits intraclass correlation structure if all the variances are equal, and all the covariances are equal, as in (7). For the normal model, that structure is equivalent to the distribution of X being invariant under permutation of its elements. For the extended model (3), we are interested in testing whether the elements of $X - \eta$ are permutation invariant. Thus we have the testing problem (2) where $\mathcal{G}_0$ is the group of $p \times p$ permutation matrices and the alternative is again the general one, $\mathcal{G}_1 = \{I_p\}$.

Using the statistic U, the estimates of Σ_U under H_1 and H_0 are respectively $\widehat{\Sigma}_U(\mathcal{G}_1) = \widehat{\Sigma}_U$ and $\widehat{\Sigma}_U(\mathcal{G}_0)$: $\widehat{\Sigma}_U(\mathcal{G}_0)$ is a matrix with diagonal elements equal to $\widehat{\sigma}^2_U$, and off-diagonal elements equal to $(\widehat{\sigma^2_U \rho_U})$, where

$$\widehat{\sigma}^2_U = tr(\widehat{\Sigma}_U)/p \quad \text{and } (\widehat{\sigma^2_U \rho_U}) = \frac{2}{p(p-1)} \sum \sum_{1 \le i < j \le p} (\widehat{\Sigma}_U)_{[ij]}. \tag{58}$$

Thus the difference of $\hat{\Sigma}_U(\mathcal{G}_0)$ and $\hat{\Sigma}_U(\mathcal{G}_1)$ has elements

$$(\widehat{\Sigma}_U)_{[ii]} - \widehat{\sigma}^2_U \text{ for } i = 1, \ldots, p, \text{ and } (\widehat{\Sigma}_U)_{[ij]} - (\widehat{\sigma^2_U \rho_U}) \text{ for } 1 \le i < j \le p. \tag{59}$$

These differences are equivalent to contrasts among the variances, and contrasts among the covariances, of $\widehat{\Sigma}_U$. Thus in C^* of (43) we can take the rows of C_1 to be any collection of $t_1 = p - 1$ linearly independent contrasts, and the rows of C_2 to be any collection of $t_2 = \binom{p}{2} - 1$ linearly independent contrasts. Then under the null hypothesis, the statistic (33) is asymptotically χ^2_ν with $\nu = p - 1 + \binom{p}{2} - 1 = p(p+1)/2 - 2$.

6. Simulating the Level and Power

We conducted a small simulation study to assess the fidelity of the actual level to the nominal one using the chi-squared approximation, and to compare powers of the tests. The sample size in this study is $n = 100$. We consider two testing situations of the type (2) as in Section 5, where the alternative hypothesis is always the general one, with $\mathcal{G}_1 = \{I_p\}$. (When we speak of correlation and variances below, we are talking about the cases in which those quantities exist.)

- *Correlation,* $p = 2$. $\mathcal{G}_0$ is as in (54) with $p_1 = p_2 = 1$, which is testing lack of correlation between X_1 and X_2.
- *Intraclass correlation,* $p = 3$. $\mathcal{G}_0$ is the set of 3×3 permutation matrices, so that this problem is testing whether the variances of X_1, X_2, and X_3 are equal, as well as whether the covariances between the three variables are equal, as in (7).

We present results for three test statistics: The normal, rank, and sign with estimated γ. The sign test with γ known shows similar results to that with estimated γ. We simulate from linear transformations of random variables with density as in (51), that is,

$$X = BZ,\ Z \sim SS_p(\nu), \tag{60}$$

for appropriate $p \times p$ matrix B. In both cases, for the null hypothesis, we take $B = I_p$. The choices for alternatives are given below:

Correlation, $p = 2$:

$$B \text{ symmetric, such that } BB' = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad \rho = 0.0, 0.05, ..., 0.45.$$

Intraclass correlation, $p = 3$:

$$B = \begin{pmatrix} 1 & c & 0 \\ 0 & 1 & 2c \\ 0 & 0 & 1 \end{pmatrix}, \qquad c = 0.0, 0.05, ..., 0.45.$$

6.1 *Simulated level.* For each of the above situations and test statistics, we estimated the level (in percent) of the test using the chi-squared cutoff point for level $\alpha = 5\%$. The degrees of freedom for the two situations are 1 and 4, respectively. Table 3 contains the results, where each value is based on 1000 simulations, so that the standard errors of the estimated levels are around 0.7%. The parameter ν in the $SS_p(\nu)$ distribution takes values 1/10, 1/5, 1/2, 1, 2, and 5.

TABLE 3. SIMULATED LEVELS

	Correlation, $p = 2$			Intraclass Correlation, $p = 3$		
	Normal	Rank	Sign	Normal	Rank	Sign
$\nu = 1/10$	1.7	4.7	5.3	2.5	5.8	5.6
$\nu = 1/5$	2.0	6.4	6.4	3.1	5.5	6.2
$\nu = 1/2$	4.2	4.8	6.0	4.4	5.9	5.4
$\nu = 1$ (Normal)	5.2	6.0	5.9	5.5	5.9	5.0
$\nu = 2$	6.6	6.4	6.1	5.1	5.0	5.2
$\nu = 5$	5.4	5.6	5.8	4.8	5.2	4.8

Table 4 summarizes the Table 3 by test statistic. “MAD” means median absolute deviation.

TABLE 4. SIMULATED LEVELS

	Normal	Rank	Sign
Minimum	1.7	4.7	4.8
Median	4.6	5.7	5.7
Maximum	6.6	6.4	6.4
MAD from 5%	0.7	0.7	0.7

Overall, the levels are fairly close to 5%, although the normal is quite conservative for heavy-tailed distributions (small ν).

REMARK 2 *An approximate randomization p-value can be easily found. Suppose that the null hypothesis holds, so that $F \in F(\mathcal{G}_0, \gamma)$. Then if $G_1, \ldots, G_n$ are iid uniformly distributed over $\mathcal{G}_0$ and independent of the X_i's,*

$$G_1(X_1 - \gamma), \ldots, G_n(X_n - \gamma) \tag{61}$$

are iid F. Thus for any shift-invariant test statistic $T(X_1, \ldots, X_n)$, fixing the data $x_1, \ldots, x_n$,

$$P(T(G_1(x_1 - \hat{\gamma}), \ldots, G_n(x_n - \hat{\gamma})) \geq T(x_1, \ldots, x_n)) \tag{62}$$

is an approximate randomization p-value, where $\hat{\gamma}$ is an estimate of γ. See Gao and Marden (2001).

6.2 *Simulated power.* To estimate the power, we use empirical cutoff points obtained from the simulations from the null hypotheses, so that the empirical levels are exactly 5%. The values of ν are as for the previous subsection, and the values of the parameters in the alternative are given just before that subsection. Each power is again estimated using 1000 simulated samples, so that standard errors are bounded by 1.6%.

Figures 1 and 2 exhibit the estimated power curves, where each figure has one plot for each value of ν. The two situations have similar sets of plots. At the normal distribution ($\nu = 1$), the normal test is slightly better than the rank test, and the sign test is a little worse than both. For the lighter-tailed distributions ($\nu > 1$), the pattern is the same, but the sign performs even less well comparatively. For heavier-tailed distributions ($\nu < 1$), the rank and sign tests tend to be better than the normal test. For very heavy tails, the normal test is quite poor. Table 5 summarizes the plots by calculating the maximum regret for each test, that is, the maximum (over the chosen alternative parameters) amount the test's power is exceeded by the best power of the three tests. E.g., for the correlation tests, the regret for test T^* is

$$Regret(T^*) = \max_{\rho=0.05,\ldots,0.45} \max_{T\in\{Normal,Rank,Sign\}} [Power_\rho(T) - Power_\rho(T^*)]. \tag{63}$$

Table 5. Maximum regrets

	Correlation, $p = 2$			Intraclass Correlation, $p = 3$		
$\nu \downarrow$	Normal	Rank	Sign	Normal	Rank	Sign
1/10	59.5	0.1	3.3	62.1	4.9	0.0
1/5	29.9	0.4	8.2	36.6	0.8	7.2
1/2	6.0	0.2	21.1	9.0	0.1	17.1
1 (Normal)	0.0	6.7	31.6	0.5	2.8	20.6
2	0.0	6.7	34.1	0.6	5.7	31.2
5	0.0	5.8	42.0	0.0	6.8	35.4

Note that in terms of regret, the rank test is always either the best or the second best, and is never larger than 6.8%. By contrast, the normal and sign tests can have very large regrets. Thus, at least for these examples, the rank tests appear to perform very well relative to the sign and normal tests.

7. Technical Results

7.1 *Asymptotic distributions under the null hypothesis* The purpose of this section is to show (34), that the test statistic is asymptotically chi-squared, under the assumptions (31) and (32). We need to show (21) for the statistic in (17), and to exhibit a consistent estimator $\widehat{\Lambda}_U$ as in (22) for $U = N(normal), S(sign)$, and $R(rank)$ statistics. We do not present all the proofs here, but have gathered them in the technical report Gao and Marden (2002).

We will take the tests based on the normal, sign, and rank statistics separately. The result for the normal statistic is straightforward, but we provide a proof that follows immediately from the proof for the rank statistic. In all that follows, we assume that $p \geq 2$, and the support of the distribution of X is not restricted to a lower-dimensional space.

7.1.1. *Sign statistic.* Define

$$h(x;\lambda) = S(x-\lambda)S(x-\lambda)'. \tag{64}$$

If the true spatial median of X, γ, is known, then

$$\widehat{\Sigma}_S = \frac{1}{n}\sum_{i=1}^{n} h(X_i;\gamma). \tag{65}$$

Because h is bounded, the central limit theorem automatically gives us that (21) holds for $U = S$, where $\Lambda_S = Cov(h(X;\gamma))$. Also, (22) holds with

$$\widehat{\Lambda}_S = \frac{1}{n}\sum_{i=1}^{n}(V_i - \overline{V})(V_i - \overline{V}) \tag{66}$$

where $V_i = vec(h(X_i;\gamma))$.

If γ is unknown, we replace it with the estimate $\hat{\gamma}$. We need to make a couple of assumptions: There exists an $L < \infty$ such that

$$E\left[\frac{1}{\|X-\lambda\|}\right] \quad \leq \quad L \text{ for any } \lambda;\ \text{and} \tag{67}$$

$$(X-\gamma) \quad =^{\mathcal{D}} \quad -(X-\gamma). \tag{68}$$

Equation (67) holds if the density of X is bounded. (See, e.g., Chaudhuri, 1996, Section 3.1.)

PROPOSITION 2 *Under assumptions (67) and (68),*

$$\frac{1}{\sqrt{n}}\sum(h(X_i;\hat{\gamma}) - E[h(X;\gamma)]) \xrightarrow{\mathcal{D}} N(0,\Lambda_S). \tag{69}$$

The proof uses Theorem 2.13 in Randles (1982), which gives convenient conditions for proving that the asymptotics are the same whether estimating the spatial median or not. See Gao and Marden (2002) for details.

Consider (22) using the estimate in (66). If the spatial median is known, then the result follows easily by the law of large numbers. The next proposition deals with the spatial median estimated.

PROPOSITION 3 *Suppose $E\|X\|^{-2} < \infty$, which holds in particular if $p \geq 3$ and the density of X is bounded. Then (22) holds for*

$$\widehat{\Lambda}_S = \frac{1}{n}\sum_{i=1}^{n} vec(S(X_i - \hat{\gamma})S(X_i - \hat{\gamma})')vec(S(X_i - \hat{\gamma})S(X_i - \hat{\gamma})')'. \quad (70)$$

7.1.2. *Rank test.*

PROPOSITION 4 *Let*

$$\Sigma_R = E[R(X;F)R(X;F)'] \quad \text{and} \quad \widehat{\Sigma}_R = \frac{1}{n}\sum R(X_i;\hat{F}_n)R(X_i;\hat{F}_n)'. \quad (71)$$

Then

$$\sqrt{n}(\widehat{\Sigma}_R - \Sigma_R) \xrightarrow{\mathcal{D}} N(0, \Lambda_R), \quad (72)$$

where

$$\Lambda_R = Cov(B(X;F)), \quad (73)$$

$$B(X;F) = R(X;F)R(X;F)' + W(X;F) + W(X;F)', \quad (74)$$

and

$$W(y;F) = E[S(X-y)R(X;F)']. \quad (75)$$

PROOF. Start by writing the statistic as a V-statistic:

$$\widehat{\Sigma}_R = \frac{1}{n^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} S_{ij}S'_{ik}. \quad (76)$$

(Recall equation 28.) The corresponding U-statistic is then

$$H_n = \frac{1}{n(n-1)(n-2)}\sum\sum\sum\nolimits_{i,j,k \text{ distinct}} S_{ij}S'_{ik}, \quad (77)$$

so that $\sqrt{n}(H_n - \widehat{\Sigma}_R) \to 0$. Now symmetrize the kernel by letting

$$h_{ijk} = \frac{1}{6}(S_{ij}S'_{ik} + S_{ik}S'_{ij} + S_{ji}S'_{jk} + S_{jk}S'_{ji} + S_{ki}S'_{kj} + S_{kj}S'_{ki}) \quad (78)$$

so that

$$H_n = \frac{1}{\binom{n}{3}} \sum\sum\sum_{i<j<k} h_{ijk}. \tag{79}$$

Then from Lemma A on page 183 of Serfling (1980), with $m = 3$, $Cov(H_n) = (9/n)Cov(h_1(X_1)) + O(n^{-2})$, where $h_1(x_i) = E[h_{ijk} \mid i]$. From (78), noting that $E[S_{ij}S'_{ik} \mid i] = R(x_i;F)R(x_i;F)'$, $E[S_{ji}S'_{jk} \mid i] = W(x_i;F)$, and likewise for the other terms, $h_1(x_i) = (1/3)B(x_i;F)$ by (74). Hence $Cov(h_1(X_1)) = (1/9)Cov(B(X_1;F))$, and

$$Cov(H_n) = Cov(B(X_1;F)) + O(n^{-2}) = \Lambda_R + O(n^{-2}). \tag{80}$$

Then (72) follows from Theorem A on page 192 of Serfling (1980). □

To show (22) using (30), note that

$$\begin{aligned}\widehat{\Lambda}_R &= \frac{1}{n^5} \sum_i \sum_j \sum_k \sum_l \sum_m vec(S_{ij}S'_{ik} + S_{ji}S'_{jk} + S_{jk}S'_{ji}) \\ &\quad vec(S_{ij}S'_{ik} + S_{ji}S'_{jk} + S_{jk}S'_{ji})' - 9vec(\hat{\Sigma}_U)vec(\hat{\Sigma}_U)'. \end{aligned} \tag{81}$$

Using U-statistic concepts, and noting that the summands are bounded, it can be shown that $\widehat{\Lambda}_R$ approaches $Cov(B(X_1;F))$.

7.1.3. *Normal test.*

PROPOSITION 5 *Suppose X_i has finite fourth moments. Let $\Sigma_N = Var(X_i)$, and $\widehat{\Sigma}_N = (1/n)\sum_{i=1}^n (X_i - \overline{X})(X_i - \overline{X})'$. Then*

$$\sqrt{n}(\widehat{\Sigma}_N - \Sigma_N) \xrightarrow{\mathcal{D}} N(0, \Lambda_N), \tag{82}$$

where

$$\Lambda_N = Cov((X_i - \mu)(X_i - \mu)'). \tag{83}$$

PROOF. Straightforward, although it is interesting that the proof for the rank case can be used, where S_{ij} is replaced by $X_i - X_j$ because $X_i - \overline{X} = (1/n)\sum_{j=1}^n (X_i - X_j)$. □

The consistency (22) is straightforward, still assuming the X_i's have finite fourth moments.

7.2. *Asymptotics under the alternative.* Consider X_{in}'s defined in (35). Recall from (37) that $\widehat{\Sigma}_U^A$ is the statistic $\widehat{\Sigma}_U$ in (17) evaluated at the X_{in}'s. We wish to show

$$\sqrt{n}(\widehat{\Sigma}_U^A - \widehat{\Sigma}_U) \xrightarrow{\mathcal{P}} \delta_U \tag{84}$$

for some δ_U. The next three propositions consider U being the normal, sign (with known spatial median), and rank statistic, respectively. The proofs are in Gao and Marden (2002).

PROPOSITION 6 *If U is the rank statistic, then*

$$\delta_R = E[S_{12}S_{13}'A'(I_p - S_{13}S_{13}') + (I_p - S_{12}S_{12}')AS_{12}S_{13}']. \tag{85}$$

PROPOSITION 7 *If U is the sign statistic with spatial median 0, then*

$$\delta_S = E[S_1S_1'A'(I_p - S_1S_1') + (I_p - S_1S_1')AS_1S_1']. \tag{86}$$

PROPOSITION 8 *If U is the normal statistic and X has finite second moments, then*

$$\delta_N = A\Sigma_X + \Sigma_X A'. \tag{87}$$

7.3 *Calculations under $\mathcal{G}_p^*$ and $\mathcal{O}_p$.* In this subsection we assume that $F \in F(\mathcal{G}_p^*, 0)$. We start with a lemma that helps in finding some of the δ_U's and Λ_U's. Recall vec^* from (40).

LEMMA 2 *Suppose U and V are random $p \times p$ matrices such that $(U, V) =^{\mathcal{D}} (gUg', gVg')$ for all $g \in \mathcal{G}_p^*$. Then if the expectation exists,*

$$E(vec^*(U)vec^*(V)') = \begin{pmatrix} aI_p + bJ_p & 0 & 0 \\ 0 & cI_{\binom{p}{2}} & dI_{\binom{p}{2}} \\ 0 & dI_{\binom{p}{2}} & cI_{\binom{p}{2}} \end{pmatrix} \equiv \Pi_p^*(a, b, c, d), \tag{88}$$

where J_p is the $p \times p$ matrix of all 1's,

$$a+b = E(U_{[11]}V_{[11]}), b = E(U_{[11]}V_{[22]}), c = E(U_{[12]}V_{[12]}) \text{ and } d = E(U_{[12]}V_{[21]}). \tag{89}$$

Also,

- *If either U or V is symmetric, then $c = d$;*
- *If $U = V = YY'$ for some random $p \times 1$ vector Y, then $b = c = d$.* □

The proof follows by considering the invariance under permutations and sign changes.

Next we exhibit the $\delta_U^*(\equiv vec^*(\delta_U))$ and Λ_U^*'s for the three statistics, where we set $\mu_U^* = vec^*(\Sigma_U) = vec^*(UU')$. See Gao and Marden (2002) for the proof.

PROPOSITION 9 *Suppose $F \in F(\mathcal{G}_p^*, 0)$, and for the Normal statistic case, that $Cov(X) = I_p$ and $E(X_{[11]}^4) < \infty$. Then*

$$\begin{aligned}
&\textit{Normal:} && \delta_N^* = vec^*(A + A'), \\
& && \Lambda_N^* = \Pi_p^*(a_N, c_N, c_N, c_N) - \mu_N^* \mu_N^{*\prime}, \\
& && a_N + b_N = E(X_{[1]}^4), c_N = E(X_{[1]}^2 X_{[2]}^2); \\
&\textit{Sign:} && \delta_S^* = (\tfrac{1}{p} I_{p^2} - \Pi_p^*(a_S, c_S, c_S, c_S)) vec^*(A + A'), \\
& && \Lambda_S^* = \Pi_p^*(a_S, c_S, c_S, c_S) - \mu_S^* \mu_S^{*\prime}, S = S(X), \\
& && a_S + b_S = E(S_{[1]}^4), c_S = E(S_{[1]}^2 S_{[2]}^2); \\
&\textit{Rank:} && \delta_R^* = (w I_{p^2} - \Pi_p^*(q, r, r, r)) vec^*(A + A'), \\
& && q = E[S_{12[1]} S_{13[1]}^3], r = E[S_{12[1]} S_{13[1]} S_{13[2]}^2], w = q + pr, \\
& && \Lambda_R^* = \Pi_p^*(a_R, b_R, c_R, c_R) - 9 \mu_R^* \mu_R^{*\prime}, B = B(X; F), \\
& && a_R + b_R = E(B_{[11]}^2), b_R = E(B_{[11]} B_{[22]}), c_R = E(B_{[12]}^2).
\end{aligned} \tag{90}$$

We finish with the proof of the main result in Section 4.

PROOF OF THEOREM 1. First, note that $C^* \mu_U^* = 0$ by (31). Because the rows of C_1 are contrasts, $C_1 J = 0$, hence

$$C^* \Pi_p^*(a, b, c, c) = \begin{pmatrix} aC_1 & 0 & 0 \\ 0 & cC_2 & cC_2 \end{pmatrix} \quad \text{and} \quad C^* \Pi_p^* C^{*\prime} = \begin{pmatrix} aC_1 C_1' & 0 \\ 0 & cC_2 C_2' \end{pmatrix}. \tag{91}$$

Using Proposition 9, and equations (42), (44), and (45), it is straightforward to show that

$$\begin{aligned}
\Delta_N &= \frac{1}{a_N} Q_1 + \frac{1}{c_N} Q_2 \\
\Delta_S &= \frac{1}{a_S} \left(\frac{1}{p} - a_S \right)^2 Q_1 + \frac{1}{c_S} \left(\frac{1}{p} - 2c_s \right)^2 Q_2 \\
\Delta_R &= \frac{p^2 r^2}{a_R} Q_1 + \frac{(q - (p-2)r)^2}{c_R} Q_2.
\end{aligned} \tag{92}$$

Now further restrict X to have a spherically symmetric distribution, so that $\|X\|$ and $S = S(X)$ are independent, and S is uniformly distributed over the unit sphere in p-space. For any such X,

$$\begin{aligned}
E(vec^*(XX') vec^*(XX')') &= E(\|X\|^4) E(vec^*(SS') vec^*(SS')') \\
&= E(\|X\|^4) \Pi_p^*(a_S, c_S, c_S, c_S),
\end{aligned} \tag{93}$$

and especially for $Z \sim N_p(0, I_p)$,

$$E(vec^*(ZZ')vec^*(ZZ')') = \Pi_p^*(2,1,1,1) \quad \text{and} \quad E(\|Z\|^4) = p(2+p). \quad (94)$$

Thus $a_s = 2/(p(2+p))$ and $c_S = 1/(p(2+p))$, hence (47) follows using (92). □

Acknowledgement. Support for this research was provided by National Science Foundation Grant NSF DMS 00-71757.

References

Andersson, S. (1975). Invariant normal models, *Ann. Statist.*, **3**, 132-154.

Andersson, S.A. and Perlman, M.D. (1984). Two testing problems relating the real and complex multivariate normal distributions, *J. Multivariate Anal.*, **15**, 21-51.

Arnold, S.F. (1973). Application of the theory of products of problems to certain patterned covariance matrices, *Ann. Statist.*, **1**, 682-699.

Arnold, S.F. (1976). Applications of products to the generalized compound symmetry problem, *Ann. Statist.*, **4**, 227-233.

Chaudhuri, P. (1996). On a geometric notion of quantiles for multivariate data, *J. Amer. Statist. Assoc.*, **91**, 862-872.

Gao, Y. and Marden, J.I. (2001). Some rank-based hypothesis tests for covariance structure and conditional independence, *Contemp. Math.*, **287**, 97-109.

Gao, Y. and Marden, J.I. (2002). Rank-based procedures for structural hypotheses on covariance matrices, with complete proofs. Technical report, Department of Statistics, University of Illinois at Urbana-Champaign.

Goodman, N.R. (1963). Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction), *Ann. Math. Statist.*, **34**, 152-177.

Khatri, C.G. (1965a). Classical statistical analysis based on a certain multivariate complex Gaussian distribution, *Ann. Math. Statist.*, **36**, 98-114.

Khatri, C.G. (1965b). A test for reality of a covariance matrix in a certain complex Gaussian distribution, *Ann. Math. Statist.*, **36**, 115-119.

Krishnaiah, P.R., Lee, J.C. and Chang, T.C. (1983). Likelihood ratio tests on covariance matrices and mean vectors of complex multivariate normal populations and their applications in time series. In *Handbook of Statistics 3: Time Series in the Frequency Domain*, D.R. Brillinger and P.R. Krishnaiah, eds., North-Holland, Amsterdam, 439-476.

Krishnaiah, P.R. and Pathak, P.K. (1967). Tests for the equality of covariance matrices under the intraclass correlation model, *Ann. Math. Statist.*, **38**, 1286-1288 (Correction: **39**, p.1358).

Lee, J.C., Krishnaiah, P.R. and Chang, T.C. (1976). On the distribution of the likelihood ratio test statistic for compound symmetry, *South African Statist. J.*, **10**, 49–62.

MARDEN, J.I. (1999). Multivariate rank tests. In *Multivariate Analysis, Design of Experiments, and Survey Sampling*, S. Ghosh, ed., Marcel Dekker, New York, 401-432.

MÖTTÖNEN, J. and OJA, H. (1995). Multivariate spatial sign and rank methods, *Nonparam. Statist.*, **5**, 201-213.

OLKIN, I. and PRESS, S.J. (1969). Testing and estimation for a circular stationary model, *Ann. Math. Statist.*, **40**, 1358-1373.

RANDLES, R.H. (1982). On the asymptotic normality of statistics with estimated parameters, *Ann. Statist.*, **10**, 462-474.

RANDLES, R.H. (1989). A distribution-free multivariate sign test based on interdirections, *J. Amer. Statist. Assoc.*, **84**, 1045-1050.

SERFLING, R.J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

SMALL, C.G. (1990). A survey of multidimensional medians, *Internat. Statist. Rev.*, **58**, 263-277.

SRIVASTAVA, M.S. (1965). Some tests for the intraclass correlation model, *Ann. Math. Statist.*, **36**, 1802-1806.

VISURI, S., KOIVUNEN, V. and OJA, H. (2000). Sign and rank covariance matrices, *J. Statist. Plann. Inference*, **91**, 557-575.

VOTAW, D.F. (1948). Testing compound symmetry in a normal multivariate distribution, *Ann. Math. Statist.*, **19**, 447-473.

WILKS, S.S. (1946). Sample criteria for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distribution, *Ann. Math. Statist.*, **17**, 257-281.

JOHN MARDEN
DEPARTMENT OF STATISTICS
UNIVERSITY OF ILLINOIS, URBANA-CHAMPAIGN
725 SOUTH WRIGHT STREET
CHAMPAIGN, IL 61820
USA
E-mail: marden@stat.uiuc.edu

YONGHONG GAO
DEPARTMENT OF STATISTICS
UNIVERSITY OF MISSOURI
5100 ROCKHILL ROAD
KANSAS CITY, MO 64110
USA
E-mail: gaoy@umkc.edu