Inequalities in Statistics and Probability
IMS Lecture Notes–Monograph Series Vol. 5 (1984), 173–177

STOCHASTIC MAJORIZATION OF THE LOG-EIGENVALUES OF A BIVARIATE WISHART MATRIX[1]

BY MICHAEL D. PERLMAN
University of Washington

Let $l = (l_1, l_2)$ and $\lambda = (\lambda_1, \lambda_2)$, where $\lambda_1 \geq \lambda_2 > 0$ are the ordered eigenvalues of $\mathbf{S}$ and Σ, respectively, and $\mathbf{S} \sim W_2(n, \Sigma)$ is a bivariate Wishart matrix. Let $\mathbf{m} = (m_1, m_2)$ and $\mu = (\mu_1, \mu_2)$, where $m_i = \log l_i$ and $\mu_i = \log \lambda_i$. It is shown that $P_\mu\{\mathbf{m} \notin B\}$ is Schur-convex in μ whenever B is a Schur-monotone set, i.e. $[\mathbf{x} \in B,\ \mathbf{x}$ *majorizes* $\mathbf{x}^*] \Rightarrow \mathbf{x}^* \in B$. This result implies the unbiasedness and power-monotonicity of a class of invariant tests for bivariate sphericity and other orthogonally invariant hypotheses.

1. Introduction. Let $\mathbf{S} \sim W_2(n, \Sigma)$ be a bivariate Wishart matrix with n degrees of freedom ($n \geq 2$) and expected value $n\Sigma$ (Σ positive definite). We shall be concerned with the power functions of orthogonally invariant tests for invariant testing problems such as the following:

$$
(1.1) \quad \begin{array}{ll}
H_{01}: \Sigma = \sigma^2 \mathbf{I}, \sigma^2 \textit{ arbitrary} & vs.\ K_1: \Sigma \textit{ arbitrary} \\
H_{02}: \Sigma = \mathbf{I} & vs.\ K_2: \Sigma \textit{ arbitrary} \\
H_{03}: \Sigma = \mathbf{I} & vs.\ K_3: \Sigma - \mathbf{I} \textit{ positive definite} \\
H_{04}: \Sigma = \mathbf{I} & vs.\ K_4: \Sigma - \mathbf{I} \textit{ negative definite.}
\end{array}
$$

Orthogonally invariant tests depend on $\mathbf{S}$ only through $l = (l_1, l_2)$, where $l_1 \geq l_2 (>0)$ are the ordered eigenvalues of $\mathbf{S}$. Because the power functions of such tests depend on Σ only through $\lambda = (\lambda_1, \lambda_2)$, where $\lambda_1 \geq \lambda_2$ (> 0) are the ordered eigenvalues of Σ, we may assume throughout this paper that $\Sigma = \mathbf{D}_\lambda \equiv diag(\lambda_1, \lambda_2)$.

The notions of majorization and Schur-convexity play an important role in determining such properties as unbiasedness and power monotonicity of invariant tests. To illustrate, consider the likelihood ratio test (LRT) for testing H_{01} (bivariate sphericity) *vs.* K_1. The acceptance region can be expressed in the equivalent forms

$$
(1.2) \qquad \{\mathbf{S} | tr\mathbf{S}/|\mathbf{S}|^{1/2} \leq c\} \Leftrightarrow \{l | (l_1 + l_2)/(l_1 l_2)^{1/2} \leq c\}.
$$

Since

$$
(1.3) \qquad tr\mathbf{S}/|\mathbf{S}|^{1/2} = (s_{11} + s_{22})/((s_{11}s_{22})^{1/2}|\mathbf{R}|^{1/2}) = (e^{t_1} + e^{t_2})/(e^{(t_1+t_2)/2}|\mathbf{R}|^{1/2}),
$$

where $\mathbf{S} = (s_{ij})_{i,j=1,2}$, $\mathbf{R}$ is the sample correlation matrix, and $t_i = \log s_{ii}$, and since s_{11}, s_{22}, and $\mathbf{R}$ are independent with $s_{ii} \sim \lambda_i \chi_n^2$ when $\Sigma = D_\lambda$, conditioning on $\mathbf{R}$ reduces the problem to the study of the power function of the LRT for equality of scale parameters ($\lambda_1 = \lambda_2$) based on the independent χ^2-variates s_{11} and s_{22} with equal degrees of freedom. It is easy to show that the joint density of $\mathbf{t} \equiv (t_1, t_2)$ is Schur-concave (in fact, permutation-invariant and log concave) with location parameter $\mu \equiv (\mu_1, \mu_2) \equiv (\log \lambda_1, \log \lambda_2)$, and that for fixed $\mathbf{R}$ the region

---

[1] This research was supported in part by National Science Foundation Grants MCS 80–02167 and MCS 83–01807.

AMS 1980 subject classifications. Primary 62H15; Secondary 62H10.

Key words and phrases: Bivariate Wishart distribution, log-eigenvalues, stochastic majorization, Schur function.

(1.4) $$\{\mathbf{t}\,|\,(e^{t_1}+e^{t_2})/(e^{(t_1+t_2)/2}) \leq c\,|\mathbf{R}|^{1/2}\} \equiv \{\mathbf{t}\,|\,2\cosh((t_1-t_2)/2) \leq c\,|\mathbf{R}|^{1/2}\}$$

is Schur-monotone in $\mathcal{R}^2$ (see Definition 2.1). By a well-known theorem of Marshall and Olkin (1979, 11.E.5.a, p. 299) it follows that

(1.5) $$1 - P_{\mu}\{2\cosh((t_1-t_2)/2) \leq c\,|\mathbf{R}|^{1/2}\,|\,\mathbf{R}\}$$

is a Schur-convex function of μ, which implies that the power of the LRT for bivariate sphericity is Schur-convex in μ. This in turn implies that the LRT is unbiased and that its power function increases monotonically as μ moves away from the null hypothesis line H_{01}: $\mu_1 = \mu_2$ at right angles.

The preceding conditional argument, due to Gleser (1966) for unbiasedness and to the author (cf. Marshall and Olkin (1979), pp. 387–8) for Schur-convexity, applies equally well to the LRT for p-variate sphericity, $p \geq 3$. Our goal is to extend these results to invariant tests other than the LRT, and to orthogonally invariant testing problems other than sphericity. In this note we show by means of a similar conditional argument that a quite general extension (Theorem 2.4) is possible in the bivariate case. The p-variate case appears more difficult, however, since it is not possible to express the eigenvalues of $\mathbf{S}$ directly in terms of its diagonal elements s_{ii} and correlation matrix $\mathbf{R}$, and a general result is not yet available for this case. Perlman (1982) gave a partial result in the p-variate case by a different argument.

2. Definitions and Main Result. We refer to Marshall and Olkin (1979) for the necessary general definitions and properties of majorization and Schur-convex functions. The following definitions and remarks concern related properties of regions in $\mathcal{R}^2$. We set $\mathbf{x} = (x_1, x_2)$.

Definition 2.1. A set $B \subseteq \mathcal{R}^2$ is *Schur-monotone in* $\mathcal{R}^2$ if [$\mathbf{x} \in \mathbf{B}$, $\mathbf{x}$ *majorizes* $\mathbf{x}^*$] $\Rightarrow$ $\mathbf{x}^* \in B$.

A Schur-monotone set in $\mathcal{R}^2$ is necessarily permutation-invariant ($(x_1,x_2) \in B \Rightarrow (x_2,x_1) \in B$), and a simple characterization is possible: B is Schur-monotone in $\mathcal{R}^2$ iff B is of the form

(2.1) $$B = \{\mathbf{x}\,|\,|x_1 - x_2| \leq f(x_1+x_2)\}$$

for an arbitrary function f on $(-\infty,\infty)$.

Definition 2.2. Let $\mathcal{R}_0^2 = \{\mathbf{x}\,|\,x_1 \geq x_2\}$. A set $B \subseteq \mathcal{R}_0^2$ is *Schur-monotone in* $\mathcal{R}_0^2$ *if* [$\mathbf{x} \in B$, $\mathbf{x}$ *majorizes* $\mathbf{x}^*$, $\mathbf{x}^* \in \mathcal{R}_0^2$] $\Rightarrow \mathbf{x}^* \in B$.

A set B is Schur-monotone in $\mathcal{R}_0^2$ iff its symmetric extension is Schur-monotone in $\mathcal{R}^2$. Equivalently, B is Schur-monotone in $\mathcal{R}_0^2$ iff B is of the form

(2.2) $$B = \{\mathbf{x}\,|\,0 \leq x_1 - x_2 \leq f(x_1+x_2)\}$$

for an arbitrary function f on $(-\infty,\infty)$.

The expressions (2.1) and (2.2) suggest the following characterizations of Schur-convex functions on $\mathcal{R}^2$ and $\mathcal{R}_0^2$, respectively:

FACT 2.3. A real-valued function β on $\mathcal{R}^2$ is Schur-convex on $\mathcal{R}^2$ iff β is permutation-invariant and $\beta(y+a, y-a)$ is increasing in $|a|$ for each fixed y in $(-\infty,\infty)$. A function β on $\mathcal{R}_0^2$ is Schur-convex on $\mathcal{R}_0^2$ iff $\beta(y+a, y-a)$ is increasing in $a \geq 0$ for each fixed y.

Let $m_i = \log l_i$, $i = 1,2$, where $l_1 \geq l_2$ are the eigenvalues of S. Set $\mathbf{m} = (m_1, m_2)$ and recall that $\mu = (\mu_1, \mu_2)$ where $\mu_i = \log \lambda_i$. The following is our main result.

THEOREM 2.4. *If B is a Schur-monotone region in $\mathcal{R}_0^2$, then*

$$\beta(\mu) \equiv P_\mu\{\mathbf{m} \notin B\}$$

is a Schur-convex function of μ on $\mathcal{R}_0^2$.

Proof. The proof extends the argument in the second paragraph of Section 1. We shall show that the event $\{\mathbf{m} \in B\}$, when expressed in terms of $t_1(\equiv \log s_{11})$ and $t_2(\equiv \log s_{22})$ for fixed $\mathbf{R}$, is a Schur-monotone region in $\mathcal{R}^2$, so that the conditional probability $P_\mu\{\mathbf{m} \notin B \mid \mathbf{R}\}$ is a Schur-convex function of μ on $\mathcal{R}_0^2$. This will immediately yield the desired result. By (2.2), the event $\{\mathbf{m} \in B\}$ is of the form

$$\{\mathbf{m} \mid 0 \leq m_1 - m_2 \leq f(m_1 + m_2)\} \tag{2.3}$$

for some function f on $(-\infty,\infty)$. Since $m_i = \log l_i$, (2.3) is equivalent to

$$\{l \mid 1 \leq (l_1/l_2) \leq g(l_1 l_2)\} \tag{2.4}$$

for some nonnegative function g on $[0,\infty)$. In the bivariate case, however, the ordered characteristic roots $l_1 \geq l_2$ of $\mathbf{S}$ are given by

$$\tfrac{1}{2}\{tr\mathbf{S} \pm [(tr\mathbf{S})^2 - 4|\mathbf{S}|]^{1/2}\}.$$

so that

$$\begin{aligned} l_1/l_2 &= \{tr\mathbf{S} + [(tr\mathbf{S})^2 - 4|\mathbf{S}|]^{1/2}\}^2/(4|\mathbf{S}|) \\ &= \{(s_{11} + s_{22}) + [(s_{11} + s_{22})^2 - 4\, s_{11}s_{22}|\mathbf{R}|]^{1/2}\}^2/(4s_{11}s_{22}|\mathbf{R}|) \\ &= \{\cosh(t_1 - t_2)/2) + [\cosh^2((t_1 - t_2)/2) - |\mathbf{R}|]^{1/2}\}^2/|\mathbf{R}|. \end{aligned} \tag{2.5}$$

By (2.4) and (2.5), therefore, the event $\{\mathbf{m} \in B\}$ is equivalent to

$$\{\ (\mathbf{t},\mathbf{R}) | \cosh((t_1 - t_2)/2) + [\cosh^2((t_1 - t_2)/2) - |\mathbf{R}|]^{1/2} \leq [|\mathbf{R}| g(e^{t_1+t_2}|\mathbf{R}|)]^{1/2}\ \}. \tag{2.6}$$

Since $y + [y^2 - |\mathbf{R}|]^{1/2}$ is increasing in y for $y \geq 1$ (note that $|\mathbf{R}| \leq 1$) and since $\cosh y$ is an increasing function of $|y|$, it follows that for fixed $\mathbf{R}$. (2.6) is of the form

$$\{\mathbf{t} \mid |t_1 - t_2| \leq h(t_1 + t_2)\} \tag{2.7}$$

for some function h on $(-\infty,\infty)$. By (2.1) it follows that (2.7) is a Schur-monotone region in $\mathcal{R}^2$, which completes the proof. □

3. Applications to the Power Functions of Invariant Tests. We shall apply Theorem 2.4 with B and β representing the acceptance region and power function of an orthogonally invariant test for each of the testing problems in (1.1).

The testing problem H_{01} vs K_1 in (1.1) can be re-expressed in terms of $\mu \equiv (\mu_1, \mu_2)$ as

$$H_{01}\colon \mu_1 = \mu_2 \ vs.\ K_1\colon \mu_1 \neq \mu_2.$$

The acceptance region of the likelihood ratio test (LRT) can be expressed in the equivalent forms

$$\begin{aligned} A_{01} &= \{l \mid (l_1 + l_2)/(l_1 l_2)^{1/2} \leq c\} \\ \Leftrightarrow B_{01} &= \{\mathbf{m} \mid 2\cosh((m_1 - m_2)/2) \leq c\} \end{aligned}$$

(cf. (1.2)–(1.4)). This is of the form (2.2), so B_{01} is a Schur-monotone region in $\mathcal{R}_0^2$. Thus Theorem 2.4 applies, so the power function of the LRT,

$$\beta(\mu) \equiv P_\mu\{\mathbf{m} \notin B_{01}\},$$

in Schur-convex in μ, as already seen in Section 1.

The testing problem H_{02} vs. K_2 can be re-expressed as

$$H_{02}\colon \mu_1 = \mu_2 = 0 \ vs.\ K_2\colon (\mu_1, \mu_2) \neq (0,0).$$

The acceptance region of the LRT can be written in the equivalent forms

$$A_{02} = \{l|\Sigma_{i=1}^{2}[\log(l_i/n) - (l_i/n) + 1] \geq c\}$$
$$\Leftrightarrow\ B_{02} = \{\mathbf{m}|\Sigma_{i=1}^{2}\ \gamma(m_i - \log n) \geqslant c\}.$$

where $\gamma(y) = y - e^y + 1$. Since γ is a concave function on $(-\infty,\infty)$, the symmetric extension of B_{02} to $\mathcal{R}^2$ is a convex, permutation-invariant region, hence B_{02} is a Schur-monotone region in $\mathcal{R}_0^2$. By Theorem 2.4, therefore, the power function of the LRT is Schur-convex in μ. Other invariant acceptance regions appropriate for testing H_{02} *vs* K_{02} include the regions

$$A_{(r)} = \{l\,\big|\, |\log l_1|^r + |\log l_2|^r \leqslant c\}$$
$$\Leftrightarrow B_{(r)} = \{\mathbf{m}\,\big|\, |m_1|^r + |m_2|^r \leqslant c\}$$

with $r > 0$. For $r \geqslant 1$, the symmetric extension of $B_{(r)}$ to $\mathcal{R}^2$ is convex and permutation-invariant, so the corresponding power function is Schur-convex in μ. Other acceptance regions possibly appropriate for this problem are the regions

$$A_{[r]} = \{l\,\big|\, (\log(l_1/l_2))^r + |\log l_1 l_2|^r \leqslant c\}$$
$$\Leftrightarrow B_{[r]} = \{\mathbf{m}\,\big|\, (m_1 - m_2)^r + |m_1 + m_2|^r \leqslant c\}$$

with $r > 0$. For each $r > 0$, $B_{[r]}$ is a Schur-monotone region in $\mathcal{R}_0^2$ by (2.2), so the corresponding power function is Schur-convex in μ. Note that for $0 < r < 1$, $B_{[r]}$ is not convex.

Next, we discuss a class of one-sided acceptance regions based on $tr(\mathbf{S}^r)$ appropriate for testing

$$H_{03}\colon \mu_1 = \mu_2 = 0\ vs.\ K_3\colon \mu_1 \geqslant \mu_2 \geqslant 0.$$

For $-\infty \leqslant r \leqslant \infty$ define

$$T_r = [\tfrac{1}{2}\, tr(\mathbf{S}^r)]^{1/r} = [\tfrac{1}{2}(l_1^r + l_2^r)]^{1/r} = [\tfrac{1}{2}(e^{rm_1} + e^{rm_2})]^{1/r};$$

note that

$$T_\infty = l_1 = e^{m_1}, \qquad T_{-\infty} = l_2 = e^{m_2},$$
$$T_0 = |S|^{1/2} = (l_1 l_2)^{1/2} = e^{(m_1+m_2)/2},$$

by continuity. The equivalent acceptance regions

$$A_r = \{l\,\big|\, T_r \leqslant c\} \quad \Leftrightarrow \quad B_r = \{\mathbf{m}\,\big|\, T_r \leqslant c\}$$

are appropriate for testing H_{03} *vs.* K_3. For each $r \geqslant 0$ ($r \leqslant 0$) the symmetric extension of $B_r(B_r^c)$ to $\mathcal{R}^2$ is convex and permutation-invariant, so by Theorem 2.4 the corresponding power function is Schur-convex (Schur-concave) in μ. [For $r = 0$, the distribution of $\mathbf{T}_0 = |\mathbf{S}|^{1/2}$ depends on μ only through $\mu_1 + \mu_2 \equiv \log|\Sigma|$, so the power function corresponding to B_0 is trivially both Schur-convex and Schur-concave in μ.]

Similarly, the equivalent regions $A_r^c \Leftrightarrow B_r^c$ complementary to $A_r \Leftrightarrow B_r$ are appropriate acceptance regions for testing

$$H_{04}\colon \mu_1 = \mu_2 = 0 \quad vs. \quad K_4\colon 0 \geqslant \mu_1 \geqslant \mu_2.$$

It follows from the preceding paragraph that the power function associated with the acceptance region B_r^c is Schur-convex (Schur-concave) for $r < 0$ ($r > 0$).

We conclude this section by considering the LRT's for testing H_{03} *vs.* K_3 and H_{04} *vs.* K_4. It can be shown (cf. Perlman (1967)) that the acceptance region of the LRT for H_{03} *vs.* K_3 can be expressed in the equivalent forms

$$A_{03} = \{\, l \mid \Sigma_{\{i|l_i \geqslant n\}}[\log (l_i/n) - l_i/n + 1] \geqslant c\}$$
$$\Leftrightarrow B_{03} = \{\mathbf{m} \mid \Sigma_{\{i|m_i \geqslant \log n\}} [(m_i - \log n) - e^{(m_i - \log n)} + 1] \geqslant c\}$$
$$= \{\mathbf{m} \mid \Sigma_{i=1}^{2} \varphi(m_i - \log n) \geqslant c\}.$$

where

$$\varphi(y) = \begin{cases} y - e^y + 1 & \text{if } y \geq 0 \\ 0 & \text{if } y \leq 0. \end{cases}$$

Since φ is a concave function on $(-\infty,\infty)$, B_{03} is a Schur-monotone region in $\mathcal{R}_0^2$ and the associated power function is Schur-convex in μ. Similarly, the acceptance region of the LRT for H_{04} *vs.* K_4 can be expressed as

$$A_{04} = \{\, l \mid \Sigma_{\{i|l_i \leqslant n\}}[\log (l_i/n) - (l_i/n) + 1] \geqslant c\}$$
$$\Leftrightarrow B_{04} = \{\mathbf{m} \mid \Sigma_{i=1}^{2} \psi(m_i - \log n) \geqslant c\}.$$

where

$$\varphi(y) = \begin{cases} y - e^y + 1 & \text{if } y \leq 0 \\ 0 & \text{if } y \geq 0. \end{cases}$$

Since ψ is concave on $(-\infty,\infty)$, if follows as above that the power function associated with B_{04} is Schur-convex in μ.

Other power monotonicity properties of some of the tests discussed in this section may be found in Anderson and Das Gupta (1964), Das Gupta (1969), and Das Gupta and Giri (1971).

4. Concluding Remarks. The proof of Theorem 2.4 proceeded indirectly, expressing the eigenvalues l_1, l_2 of S in terms of its diagonal elements s_{11}, s_{22} and working with the relatively simple joint distribution of (s_{11}, s_{22}). This approach may not easily extend to the p-variate case ($p \geqslant 3$), so it may be preferable to work directly with the joint distribution of the eigenvalues $l_1, \ldots, l_p$ (cf. Muirhead (1982), Theorem 9.4.1). We admit, however, that we were unable to carry through the latter approach in the bivariate case.

Theorem 2.4 can be extended from probabilities of Schur-monotone regions to expectations of Schur-convex functions: if g is a Schur-convex function on $\mathcal{R}_0^2$ such that the expectations exist, then $E_\mu[g(\mathbf{m})]$ is a Schur-convex function of μ. This follows by a standard approximation argument.

REFERENCES

ANDERSON, T. W. and DAS GUPTA, S. (1964). A monotonicity property of the power functions of some tests of equality of two covariance matrices. *Ann. Math. Statist. 35* 1059–1063. [Correction: *ibid 36* (1965), 1318.]

DAS GUPTA, S. (1969). Properties of power functions of some tests concerning dispersion matrices of multivariate normal distirubtions. *Ann. Math. Statist. 40* 697–701.

DAS GUPTA, S. and GIRI, N. (1973). Properties of tests concerning covariance matrices of normal distributions. *Ann. Statist. 1* 1222–1224.

GLESER, L. J. (1966). A note on the sphericity test. *Ann. Math. Statist. 37* 464–467.

MARSHALL, A. W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York.

MUIRHEAD, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, New York.

PERLMAN, M. D. (1967). One-sided testing problems involving covariance matrices. Unpublished notes, Department of Statistics, Stanford University, Stanford, CA.

PERLMAN, M. D. (1982). Stochastic comparison of eigenvalues of Wishart matrices with applications to unbiasedness of tests for variance and covariance matrices. Invited address presented at the Symposium on Inequalities in Statistics and Probability, University of Nebraska, October 27–30, 1982.