Journal of Statistical Physics, Vol. 116, Nos. 1/4, August 2004 (© 2004)

Total Positivity Properties of Generalized Hypergeometric Functions of Matrix Argument

Donald St. P. Richards[1]

Received July 27, 2003; accepted February 13, 2004

In multivariate statistical analysis, several authors have studied the total positivity properties of the generalized (${}_0F_1$) hypergeometric function of two real symmetric matrix arguments. In this paper, we make use of zonal polynomial expansions to obtain a new proof of a result that these ${}_0F_1$ functions fail to satisfy certain pairwise total positivity properties; this proof extends both to arbitrary generalized (${}_rF_s$) functions of two matrix arguments and to the generalized hypergeometric functions of Hermitian matrix arguments. In the case of the generalized hypergeometric functions of two Hermitian matrix arguments, we prove that these functions satisfy certain modified pairwise TP_2 properties; the proofs of these results are based on Sylvester's formula for compound determinants and the condensation formula of C. L. Dodgson [Lewis Carroll] (1866).

KEY WORDS: Compound determinant; condensation formula; FKG inequality; likelihood ratio test statistics; monotone power function; random matrix; total positivity; noncentral Wishart distribution; zonal polynomial.

1. INTRODUCTION

Suppose that $X_1,\ldots, X_n \in \mathbb{R}^p$ are mutually independent, normally distributed (i.e., Gaussian), column random vectors, with common positive-definite (symmetric) covariance matrix Σ. Form the $p \times n$ random matrix $X = [X_1,\ldots, X_n]$ having columns $X_1,\ldots, X_n$, and let $\mu = E(X)$, the expectation of X. We assume throughout that $n \geqslant p$, so that the random matrix XX' is positive-definite almost surely. It is well-known that XX' has a *noncentral Wishart distribution* with *n degrees of freedom* and *noncentral matrix*

[1] Department of Statistics, Penn State University, University Park, Pennsylvania 16802-2111; e-mail: richards@stat.psu.edu

0022-4715/04/0800-0907/0

parameter $\mu\mu'\Sigma^{-1}$; cf. James (1964), Muirhead (1982). Let $l_1, \ldots, l_p$ denote the eigenvalues of XX' and set $L = \mathrm{diag}(l_1, \ldots, l_p)$. Further, let $\lambda_1, \ldots, \lambda_p$ denote the eigenvalues of the noncentral matrix parameter $\mu\mu'\Sigma^{-1}$, and set $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$.

In statistical inference about the mean matrix μ, it is of interest to test the null hypothesis H_0: $\mu = 0$ against the alternative hypothesis H_1: $\mu \neq 0$. Classical statistical test procedures for testing H_0 are based on the eigenvalues $l_1, \ldots, l_p$, and usually are of the form $h(l_1, \ldots, l_p)$ where the real-valued function h satisfies various invariance properties. We refer to Muirhead (1982) and Anderson (2003) for extensive accounts of these hypothesis testing problems.

In general, a viable test statistic is required to satisfy various probabilistic properties. In particular, it is important that a test statistic have *monotone power function.* In basic terms, the monotone power function property necessitates that, as the null hypothesis becomes increasingly implausible, the test statistic under consideration be increasingly able to detect this rising implausibility.

As a consequence of results of James (1964), it is well-known that ϕ, the probability density function of the random matrix L, can be expressed in the form

$$\phi(L) = \phi_0(L)\, \phi_1(\Lambda)\, f(\Lambda, L) \tag{1.1}$$

where ϕ_0 is the probability density function of the eigenvalues of a central Wishart random matrix (Muirhead, 1982), ϕ_1 is a "nuisance function" which plays no role in the subsequent analysis, and the function $f(\Lambda, L)$ is expressible in terms of a generalized hypergeometric function (${}_0F_1$) of two matrix arguments; cf. Perlman and Olkin (1980), Eq. (3.1).

In a study of the monotonicity properties of some likelihood ratio test statistics, Perlman and Olkin (1980) established a remarkable connection with the theory of total positivity. Perlman and Olkin (1980) proved that if the function $f(\Lambda, L)$ satisfied certain total positivity properties then the power function of the underlying test statistic has some desirable monotonicity properties.

In investigating the statistical inference problem about μ, Perlman and Olkin (1980) raised the question of whether or not the function $f(\Lambda, L)$ is *totally positive of order* 2 (TP_2) in each pair (l_i, l_j), $1 \leqslant i \neq j \leqslant p$ and in each pair (λ_i, l_j), $1 \leqslant i, j \leqslant p$. Groeneboom and Truax (2002) subsequently proved that the function f does not satisfy these pairwise TP_2 conditions; however, they also proved that f satisfies a weaker class of TP_2 criteria and they deduced the corresponding statistical implications of their result.

The purpose of the present paper is two-fold. First, we provide an alternative proof of the result of Groeneboom and Truax (2002) concerning the failure of the function f to satisfy the full class of pairwise TP_2 conditions. In prior work on these problems, the basic approach was to study the function $f(\Lambda, L)$ through an integral representation of the generalized hypergeometric function ${}_0F_1$, an approach which seems difficult to extend to more general generalized hypergeometric functions of matrix argument. On the other hand, the approach given here is based on expansions of the hypergeometric functions in series of zonal polynomials, a method which will be seen to apply readily to any generalized hypergeometric function of matrix argument. These results are given in Section 2.

The second purpose of the present paper relates to the results of Bondar (1988). As noted above, the function $f(\Lambda, L)$ is known not to satisfy the full set of TP_2 conditions. Nevertheless, Bondar (1988) observed that the program formulated by Perlman and Olkin (1980) will remain viable if, for certain functions $\psi(\Lambda, L)$ which are symmetric in $\lambda_1,..., \lambda_p$ and also symmetric in $l_1,..., l_p$, the function $\psi(\Lambda, L)\, f(\Lambda, L)$ is TP_2 in each pair (l_i, l_j), $1 \leqslant i \neq j \leqslant p$ and in each pair (λ_i, l_j), $1 \leqslant i, j \leqslant p$, i.e., if $\psi(\Lambda, L)\, f(\Lambda, L)$ satisfies the full set of pairwise TP_2 conditions. In the case of Gaussian random vectors $X_1,..., X_n \in \mathbb{R}^p$, which is the case of primary interest to statisticians, we have not been able to determine any function ψ for which the full class of TP_2 conditions are valid.

In the case of mutually independent, *complex* random vectors $X_1,..., X_n \in \mathbb{C}^p$ having complex normal distributions with a common positive definite Hermitian covariance matrix $\tilde{\Sigma}$, we form the random matrix $X = [X_1,..., X_n]$ and let $\mu = E(X)$. We denote by $X^* = \bar{X}'$ and $\mu^* = \bar{\mu}'$ the transpose of the complex conjugate of X and μ, respectively. As before, we assume that $n \geqslant p$ in order to ensure that XX^* is positive-definite, almost surely. We again denote by $l_1,..., l_p$ the eigenvalues of XX^* and set $L = \mathrm{diag}(l_1,..., l_p)$; we also denote by $\lambda_1,..., \lambda_p$ the eigenvalues of the non-central matrix parameter $\mu\mu^*\tilde{\Sigma}^{-1}$, and set $\Lambda = \mathrm{diag}(\lambda_1,..., \lambda_p)$. Again from results of James (1964), the probability density function of the random matrix L is known to be of the form (1.1). In a straightforward analogy with the results of Section 2, the corresponding function f also fails to satisfy the full set of pairwise TP_2 conditions, i.e., TP_2 in each pair (l_i, l_j), $1 \leqslant i \neq j \leqslant p$ and in each pair (λ_i, l_j), $1 \leqslant i, j \leqslant p$. Nevertheless, and this is the major result of the present paper, we establish that for the case in which

$$\psi(\Lambda, L) = \prod_{1 \leqslant i < j \leqslant p} |(\lambda_i - \lambda_j)(l_i - l_j)|, \tag{1.2}$$

the full set of pairwise TP_2 conditions holds for the function $\psi(\Lambda, L)\, f(\Lambda, L)$.

The proofs of these results are noteworthy, for they involve Sylvester's formula for compound determinants (Karlin, 1968) and the famous condensation formula of C. L. Dodgson [Lewis Carroll] (1866). More generally, these techniques are applicable to all of the complex-case noncentral eigenvalue probability density functions listed by James (1964), Section 8.

2. THE REAL CASE

A *partition* $\kappa = (k_1, \ldots, k_n)$ is an n-tuple of nonnegative integers $k_1, \ldots, k_n$ satisfying $k_1 \geqslant \cdots \geqslant k_n$. The *length* of κ is defined to be $|\kappa| := k_1 + \cdots + k_n$. For any $a \in \mathbb{R}$, the *partitional rising factorial* is defined by

$$[a]_\kappa := \prod_{j=1}^{n} (a - \tfrac{1}{2}(j-1))_{k_j},$$

where $(a)_k = a(a+1)\cdots(a+k-1)$, $k = 0, 1, 2, \ldots$, is the classical rising factorial.

Corresponding to each partition κ is a *zonal polynomial*, $C_\kappa(\Lambda)$ (cf. Muirhead, 1982, Chapter 7). The polynomials $C_\kappa(\Lambda)$ have a rich theory and satisfy many remarkable properties; in particular, $C_\kappa(\Lambda)$ is homogeneous of degree $|\kappa|$ in Λ and is positive if Λ is positive-definite. It suffices for our purposes to note that for the case in which the partition κ is of length no more than two, the zonal polynomials $C_\kappa(\Lambda)$ are given explicitly as follows (cf. Muirhead, 1982, p. 232 ff.). Define the monomial symmetric functions

$$M_{(2)}(\Lambda) = \mathrm{tr}\,(\Lambda^2) = \sum_{j=1}^{p} \lambda_j^2$$

and

$$M_{(1,1)}(\Lambda) = \sum_{1 \leqslant i < j \leqslant p} \lambda_i \lambda_j = \tfrac{1}{2}[(\mathrm{tr}\,\Lambda)^2 - \mathrm{tr}\,(\Lambda^2)];$$

then the zonal polynomials of degree up to two are

$$C_\kappa(\Lambda) = \begin{cases} 1, & \kappa = (0) \\ \mathrm{tr}(\Lambda), & \kappa = (1) \\ M_{(2)}(\Lambda) + \tfrac{2}{3} M_{(1,1)}(\Lambda), & \kappa = (2) \\ \tfrac{4}{3} M_{(1,1)}(\Lambda), & \kappa = (1,1) \end{cases} \tag{2.1}$$

For nonnegative integers r and s, numerator parameters $a_1,\ldots, a_r \in \mathbb{C}$ and denominator parameters $b_1,\ldots, b_s \in \mathbb{C}$, the generalized hypergeometric function, ${}_rF_s$, of two matrix arguments is defined by the zonal polynomial expansion

$$ {}_rF_s(a_1,\ldots, a_r; b_1,\ldots, b_s; \Lambda, L) = \sum_{k=0}^{\infty} \sum_{|\kappa|=k} \frac{[a_1]_\kappa \cdots [a_r]_\kappa}{[b_1]_\kappa \cdots [b_s]_\kappa} \frac{C_\kappa(\Lambda)\, C_\kappa(L)}{k!\, C_\kappa(I_p)} \tag{2.2} $$

where I_p denotes the $p \times p$ identity matrix and $b_1,\ldots, b_s$ are such that, for all partitions κ, $[b_j]_\kappa \neq 0$ for all $j = 1,\ldots, s$.

We refer to Muirhead (1982) or Gross and Richards (1987) for the general theory of these generalized hypergeometric functions. For the case in which $r = 0$ and $s = 1$, (2.2) reduces to

$$ {}_0F_1(b; \Lambda, L) = \sum_{k=0}^{\infty} \sum_{|\kappa|=k} \frac{1}{[b]_\kappa} \frac{C_\kappa(\Lambda)\, C_\kappa(L)}{k!\, C_\kappa(I_p)}, \tag{2.3} $$

an everywhere convergent series.

Suppose that $X = [X_1,\ldots, X_n]$ is a $p \times n$ random matrix whose columns are mutually independent, Gaussian random vectors with common covariance matrix Σ. Then, with the notation in (1.1), we have $\phi_1(\Lambda) = \exp(-\mathrm{tr}\, \Lambda/2)$, $f(\Lambda, L) = {}_0F_1(n/2; \Lambda, L)$, and

$$ \phi_0(L) = k(p, n) \prod_{j=1}^{p} l_i^{(n-p-1)/2} \exp(-l_i/2) \prod_{1 \leqslant i < j \leqslant p} (l_i - l_j)_+, $$

where t_+ denotes the positive part of t and $k(p, n)$ is a normalizing constant such that ϕ_0 is itself a probability density function; cf. Perlman and Olkin (1980), Eq. (3.1).

Let us also recall (cf. Karlin, 1968) that a nonnegative function $K: \mathbb{R}^2 \to \mathbb{R}$ is *totally positive* of order p (TP_p) if, for all $u_1 > \cdots > u_p$ and $v_1 > \cdots > v_p$, the $r \times r$ determinant

$$ \det(K(u_i, v_j)) $$

is nonnegative for all $r = 1,\ldots, p$. For the case in which the function K is sufficiently smooth and strictly positive, it is well-known that the TP_2 property is equivalent to the inequality

$$ \frac{\partial^2}{\partial u\, \partial v} \log K(u, v) \geqslant 0 $$

for all u, v.

Now we consider the TP_2 properties of the function $f(\Lambda, L)$. Because the variables $\lambda_1, \ldots, \lambda_p$ are nonnegative and $l_1, \ldots, l_p$ are positive (almost surely), we have $C_\kappa(\Lambda) \geqslant 0$ and $C_\kappa(L) \geqslant 0$, almost surely, for all κ. Then $f(\Lambda, L) > 0$ almost everywhere, so its TP_2 properties may be determined by studying the sign of the function

$$\frac{\partial^2}{\partial l_1 \partial l_2} \log {}_0F_1(n/2; \Lambda, L). \tag{2.4}$$

We shall use the zonal polynomial expansion (2.3) to prove that this partial derivative is negative in a neighborhood of $L = 0$ when Λ is sufficiently large.

Using the zonal polynomial expansion in (2.3), we obtain

$$\begin{aligned} {}_0F_1(n/2; \Lambda, L) &= \sum_{k=0}^{2} \sum_{|\kappa|=k} \frac{C_\kappa(\Lambda)\, C_\kappa(L)}{k!\, [n/2]_\kappa\, C_\kappa(I_p)} + O(L^3) \\ &= 1 + \frac{2}{np} (\operatorname{tr} \Lambda)(\operatorname{tr} L) + \frac{6}{np(n+2)(p+2)} C_{(2)}(\Lambda)\, C_{(2)}(L) \\ &\quad + \frac{3}{np(n-1)(p-1)} C_{(1,1)}(\Lambda)\, C_{(1,1)}(L) + O(L^3). \end{aligned}$$

On applying (2.1) to express each $C_\kappa(L)$ in terms of the $M_\kappa(L)$, and differentiating with respect to l_1 and l_2, we obtain

$$\begin{aligned} &\frac{\partial}{\partial l_j} {}_0F_1(n/2; \Lambda, L) \\ &\quad = \frac{2}{np} \operatorname{tr} \Lambda + \frac{4}{np(n+2)(p+2)} (2l_j + \operatorname{tr} L)\, C_{(2)}(\Lambda) \\ &\quad\quad + \frac{4}{np(n-1)(p-1)} (-l_j + \operatorname{tr} L)\, C_{(1,1)}(\Lambda) + O(L^2), \end{aligned} \tag{2.5}$$

for $j = 1, 2$, and

$$\begin{aligned} \frac{\partial^2}{\partial l_1\, \partial l_2} {}_0F_1(n/2; \Lambda, L) &= \frac{4}{np(n+2)(p+2)} C_{(2)}(\Lambda) \\ &\quad + \frac{4}{np(n-1)(p-1)} C_{(1,1)}(\Lambda) + O(L). \end{aligned} \tag{2.6}$$

Define

$$G(\Lambda, L) := {}_0F_1(n/2; \Lambda, L) \frac{\partial^2}{\partial l_1\, \partial l_2} {}_0F_1(n/2; \Lambda, L) - \frac{\partial}{\partial l_1} {}_0F_1(n/2; \Lambda, L) \cdot \frac{\partial}{\partial l_2} {}_0F_1(n/2; \Lambda, L).$$

Clearly,

$$\frac{\partial^2}{\partial l_1\, \partial l_2} \log {}_0F_1(n/2; \Lambda, L) = \frac{G(\Lambda, L)}{[{}_0F_1(n/2; \Lambda, L)]^2},$$

and since ${}_0F_1(n/2; \Lambda, L) \to 1$ as $L \to 0$ we then have

$$\lim_{L \to 0} \frac{\partial^2}{\partial l_1\, \partial l_2} \log {}_0F_1(n/2; \Lambda, L) = \lim_{L \to 0} G(\Lambda, L).$$

By straightforward algebraic computations using (2.5) and (2.6) we obtain

$$G(\Lambda, L) = \frac{4}{np(n+2)(p+2)} C_{(2)}(\Lambda) + \frac{4}{np(n-1)(p-1)} C_{(1,1)}(\Lambda) - \frac{4}{n^2 p^2} (\operatorname{tr} \Lambda)^2 + O(L).$$

Again applying (2.1) to express each $C_\kappa(\Lambda)$ in terms of the $M_\kappa(\Lambda)$, we obtain

$$G(\Lambda, L) = -\frac{8(n+p+2)}{n^2 p^2 (n+2)(p+2)} M_{(2)}(\Lambda) + \frac{16[n(n+p+1)+(p-1)(p+2)]}{n^2 p^2 (n+2)(p-1)(p+2)} M_{(1,1)}(\Lambda) + O(L).$$

Therefore

$$\lim_{L \to 0} \frac{\partial^2}{\partial l_1\, \partial l_2} \log {}_0F_1(n/2; \Lambda, L) = \lim_{L \to 0} G(\Lambda, L) = -a_1 \sum_{j=1}^{p} \lambda_j^2 + a_2 \sum_{1 \leqslant i < j \leqslant p} \lambda_i \lambda_j$$

with constants $a_j > 0$, $j = 1, 2$. For fixed $\lambda_2, \ldots, \lambda_p$, the right-hand side of this last equation is a quadratic polynomial in λ_1 in which the coefficient of λ_1^2 is negative. Hence, for fixed $\lambda_2, \ldots, \lambda_p$, this polynomial attains negative values for sufficiently large values of λ_1.

In conclusion,

$$\frac{\partial^2}{\partial l_1 \, \partial l_2} \log {}_0F_1(n/2; \Lambda, L) < 0$$

for sufficiently small L and sufficiently large Λ.

To conclude this section, we note that the arguments utilized above may be applied to any generalized hypergeometric function ${}_rF_s$ of two matrix arguments; in short, these functions do not generally satisfy pairwise TP_2 properties.

3. THE COMPLEX CASE

We now consider the complex analog of the problems studied in the previous section. We suppose now that we begin with mutually independent, complex random vectors $X_1, \ldots, X_n$ having complex Gaussian distributions (Goodman, 1963). As before, we assume that the vectors $X_1, \ldots, X_n$ have a common positive-definite Hermitian covariance matrix $\tilde{\Sigma}$. After forming the $n \times p$ matrix $X = [X_1 2, \ldots, X_n]$, we wish to study the total positivity properties of $l_1, \ldots, l_p$, the eigenvalues of XX^*. Let $\tilde{\mu} := E(X)$, and let Λ denote the diagonal matrix whose diagonal entries are the eigenvalues of the noncentrality parameter matrix $\tilde{\mu}\tilde{\mu}^* \tilde{\Sigma}^{-1}$. By the results of James (1964), Eq. (102), it follows that the probability density function of L is of the form (1.1) where $\phi_1(L) = \exp(-\mathrm{tr}\, L)$, and

$$\phi_0(L) = k(p, n) \prod_{j=1}^{p} l_i^{n-p} \exp(-l_i) \prod_{1 \leqslant i < j \leqslant p} (l_i - l_j)^2$$

with $k(p, n)$ a normalizing constant. Moreover, we have $f(\Lambda, L) = {}_0\tilde{F}_1(n; \Lambda, L)$ where ${}_0\tilde{F}_1$ is a generalized hypergeometric function of two *Hermitian* matrix arguments.

Analogous to (2.3), there is an expansion for ${}_0\tilde{F}_1$ in terms of the *complex* zonal polynomials, cf. James (1964), Eq. (85). First, for each partition $\kappa = (k_1, \ldots, k_p)$, the partitional rising factorial is now defined as

$$[a]_\kappa := \prod_{j=1}^{n} (a - j + 1)_{k_j}.$$

Next, for nonnegative integers r and s, numerator parameters $a_1,..., a_r \in \mathbb{C}$ and denominator parameters $b_1,..., b_s \in \mathbb{C}$, the generalized hypergeometric function, ${}_r\tilde{F}_s$, of two Hermitian matrix arguments is defined by the zonal polynomial expansion

$$ {}_r\tilde{F}_s(a_1,..., a_r; b_1,..., b_s; \Lambda, L) = \sum_{k=0}^{\infty} \sum_{|\kappa|=k} \frac{[a_1]_\kappa \cdots [a_r]_\kappa}{[b_1]_\kappa \cdots [b_s]_\kappa} \frac{\tilde{C}_\kappa(\Lambda)\, \tilde{C}_\kappa(L)}{k!\, \tilde{C}_\kappa(I_p)}, \quad (3.1) $$

where $\tilde{C}_\kappa$ is the complex zonal polynomial; cf. James (1964).

Using explicit formulas for the low-degree complex zonal polynomials and following the arguments of Section 2, it is straightforward to establish that

$$ \frac{\partial^2}{\partial l_1\, \partial l_2} \log {}_0\tilde{F}_1(n; \Lambda, L) < 0 $$

for sufficiently small L and sufficiently large Λ. Therefore, the probability density function ϕ also generally fails to satisfy the pairwise TP_2 properties.

Nevertheless, motivated by comments of Bondar (1988), we now establish the full class of pairwise TP_2 properties for a modified form of the function ϕ. We first prove the following result.

Theorem 3.1. For $p \geqslant 1$, $a > p-1$, $\lambda_1 > \cdots > \lambda_p > 0$ and $l_1 > \cdots > l_p > 0$,

$$ \frac{\partial^2}{\partial l_1\, \partial l_2} \log \left[\prod_{1 \leqslant i < j \leqslant p} (\lambda_i - \lambda_j)(l_i - l_j) \cdot {}_0\tilde{F}_1(a; \Lambda, L) \right] \geqslant 0. \quad (3.2) $$

The proof of this result rests on an explicit determinantal formula for ${}_0\tilde{F}_1(a; \Lambda, L)$ in terms of the classical, scalar-valued, generalized hypergeometric functions, and on a consequence of Sylvester's formula for compound determinants. To begin, we state the following result.

Theorem 3.2 (Gross and Richards, 1989). Suppose that for each $i = 1,..., s$, $b_i - j + 1$ is not a non-positive integer for any $j = 1,..., p$. Then

$$ {}_r\tilde{F}_s(a_1,..., a_r; b_1,..., b_s; \Lambda, L) = c_{r,s} \frac{\det({}_rF_s(a_1 - p+1,..., a_r - p+1; b_1 - p+1,..., b_s - p+1; \lambda_i l_j))}{\prod_{1 \leqslant i < j \leqslant p} (\lambda_i - \lambda_j)(l_i - l_j)} \quad (3.3) $$

where $c_{r,s}$ is a positive constant.

Note that the ${}_rF_s$ functions in the determinant on the right-hand side of (3.3) are the classical scalar-valued generalized hypergeometric functions. The explicit formula for the constant $c_{r,s}$ is given by Gross and Richards (1989); however, its value is not needed here.

Assume that $l_1 > \cdots > l_p$. For the rest of the paper, we will use the notation

$$K(\lambda, l) = {}_0F_1(a-p+1; \lambda l),$$

$\lambda, l \in \mathbb{R}$; however, it will be clear from the context that K generally can be chosen as an arbitrary totally positive function of suitable order. By Theorem 3.2,

$$\begin{aligned}
&\frac{\partial^2}{\partial l_1\, \partial l_2} \log[(\lambda_1-\lambda_2)(l_1-l_2)\cdot {}_0F_1(a; \Lambda, L)] \\
&\quad = \frac{\partial^2}{\partial l_1\, \partial l_2} \log\left[\prod_{1 \leqslant i<j \leqslant p} (\lambda_i-\lambda_j)(l_i-l_j)\cdot {}_0F_1(a; \Lambda, L)\right] \\
&\quad = \frac{\partial^2}{\partial l_1\, \partial l_2} \log \det(K(\lambda_i, l_j)).
\end{aligned}$$

Let us define

$$\begin{aligned}
G_1(\Lambda, L) &:= [\det(K(\lambda_i, l_j))]^2 \frac{\partial^2}{\partial l_1\, \partial l_2} \log \det(K(\lambda_i, l_j)) \\
&= \det(K(\lambda_i, l_j)) \cdot \frac{\partial^2}{\partial l_1\, \partial l_2} \det(K(\lambda_i, l_j)) \\
&\quad - \frac{\partial}{\partial l_1} \det(K(\lambda_i, l_j)) \cdot \frac{\partial}{\partial l_2} \det(K(\lambda_i, l_j)).
\end{aligned} \tag{3.4}$$

We recall a "generalized relation of second-order determinants," given by Karlin (1968), p. 7, Eq. (0.16): For p-dimensional vectors

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix}, \quad \mathbf{f}^{(1)} = \begin{pmatrix} f_1^{(1)} \\ f_2^{(1)} \\ \vdots \\ f_p^{(1)} \end{pmatrix}, \quad \mathbf{f}^{(2)} = \begin{pmatrix} f_1^{(2)} \\ f_2^{(2)} \\ \vdots \\ f_p^{(2)} \end{pmatrix}, \ldots, \quad \mathbf{f}^{(p-2)} = \begin{pmatrix} f_1^{(p-2)} \\ f_2^{(p-2)} \\ \vdots \\ f_p^{(p-2)} \end{pmatrix},$$

define the determinant

$$D(\mathbf{a}, \mathbf{b}, \mathbf{f}^{(1)}, \ldots, \mathbf{f}^{(p-2)}) = \begin{vmatrix} a_1 & b_1 & f_1^{(1)} & \cdots & f_1^{(p-2)} \\ a_2 & b_2 & f_2^{(1)} & \cdots & f_2^{(p-2)} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_p & b_p & f_p^{(1)} & \cdots & f_p^{(p-2)} \end{vmatrix}.$$

By an application of Sylvester's formula for compound determinants, Karlin (loc. cit.) proves that for any $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathbb{R}^p$,

$$\begin{vmatrix} D(\mathbf{a}, \mathbf{c}, \mathbf{f}^{(1)}, \ldots, \mathbf{f}^{(p-2)}) & D(\mathbf{a}, \mathbf{d}, \mathbf{f}^{(1)}, \ldots, \mathbf{f}^{(p-2)}) \\ D(\mathbf{b}, \mathbf{c}, \mathbf{f}^{(1)}, \ldots, \mathbf{f}^{(p-2)}) & D(\mathbf{b}, \mathbf{d}, \mathbf{f}^{(1)}, \ldots, \mathbf{f}^{(p-2)}) \end{vmatrix}$$
$$= D(\mathbf{a}, \mathbf{b}, \mathbf{f}^{(1)}, \ldots, \mathbf{f}^{(p-2)})\, D(\mathbf{c}, \mathbf{d}, \mathbf{f}^{(1)}, \ldots, \mathbf{f}^{(p-2)}). \tag{3.5}$$

We now set

$$\mathbf{a} = \begin{pmatrix} K(\lambda_1, l_1) \\ K(\lambda_2, l_1) \\ \vdots \\ K(\lambda_p, l_1) \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} \frac{\partial}{\partial l_1} K(\lambda_1, l_1) \\ \frac{\partial}{\partial l_1} K(\lambda_2, l_1) \\ \vdots \\ \frac{\partial}{\partial l_1} K(\lambda_p, l_1) \end{pmatrix},$$

$$\mathbf{c} = \begin{pmatrix} K(\lambda_1, l_2) \\ K(\lambda_2, l_2) \\ \vdots \\ K(\lambda_p, l_2) \end{pmatrix}, \quad \mathbf{d} = \begin{pmatrix} \frac{\partial}{\partial l_2} K(\lambda_1, l_2) \\ \frac{\partial}{\partial l_2} K(\lambda_2, l_2) \\ \vdots \\ \frac{\partial}{\partial l_2} (\lambda_p, l_2) \end{pmatrix},$$

and, for $j = 1, \ldots, p-2$, set

$$\mathbf{f}^{(j)} = \begin{pmatrix} K(\lambda_1, l_{j+2}) \\ K(\lambda_2, l_{j+2}) \\ \vdots \\ K(\lambda_p, l_{j+2}) \end{pmatrix}.$$

With these substitutions, the right-hand side of (3.4) is precisely the determinant on the left-hand side of (3.5). Therefore by (3.5), we have

$$G_1(\Lambda, L) = D(\mathbf{a}, \mathbf{b}, \mathbf{f}^{(1)}, \ldots, \mathbf{f}^{(p-2)})\, D(\mathbf{c}, \mathbf{d}, \mathbf{f}^{(1)}, \ldots, \mathbf{f}^{(p-2)}).$$

Now

$$D(\mathbf{a}, \mathbf{b}, \mathbf{f}^{(1)}, \ldots, \mathbf{f}^{(p-2)}) \equiv \begin{vmatrix} K(\lambda_1, l_1) & \frac{\partial}{\partial l_1} K(\lambda_1, l_1) & K(\lambda_1, l_3) & \cdots & K(\lambda_1, l_p) \\ K(\lambda_2, l_1) & \frac{\partial}{\partial l_1} K(\lambda_2, l_1) & K(\lambda_2, l_3) & \cdots & K(\lambda_2, l_p) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ K(\lambda_p, l_1) & \frac{\partial}{\partial l_1} K(\lambda_p, l_1) & K(\lambda_p, l_3) & \cdots & K(\lambda_p, l_p) \end{vmatrix} \tag{3.6}$$

and

$$D(\mathbf{c}, \mathbf{d}, \mathbf{f}^{(1)}, \ldots, \mathbf{f}^{(p-2)}) \equiv \begin{vmatrix} K(\lambda_1, l_2) & \frac{\partial}{\partial l_2} K(\lambda_1, l_2) & K(\lambda_1, l_3) & \cdots & K(\lambda_1, l_p) \\ K(\lambda_2, l_2) & \frac{\partial}{\partial l_2} K(\lambda_2, l_2) & K(\lambda_2, l_3) & \cdots & K(\lambda_2, l_p) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ K(\lambda_p, l_2) & \frac{\partial}{\partial l_2} K(\lambda_p, l_2) & K(\lambda_p, l_3) & \cdots & K(\lambda_p, l_p) \end{vmatrix}. \tag{3.7}$$

By appeal to the theory of total positivity, if the kernel K is TP_p then each of these determinants is nonnegative for $\lambda_1 > \cdots > \lambda_p$ and $l_1 > \cdots > l_p$. Indeed, the determinant (3.6) may be expressed as the limiting value,

$$\lim_{l_2 \to l_1} \frac{\det(K(\lambda_i, l_j))}{l_1 - l_2}; \tag{3.8}$$

since both numerator and denominator in this limit are nonnegative, it is now clear that (3.6) is nonnegative. A similar argument applies to show that (3.7) is also nonnegative.

In the case of the kernel $K(\lambda, l) = {}_0F_1(a-p+1, \lambda l)$, $\lambda, l > 0$, it is well-known that K is TP_p for $a > p-1$; cf. Karlin (1968), Gross and Richards (1989). Consequently, we deduce that $G_1(\Lambda, L) \geqslant 0$.

We next establish the following result.

Theorem 3.3. For $a > p-1$, $\lambda_1 > \cdots > \lambda_p > 0$, and $l_1 > \cdots > l_p > 0$,

$$\frac{\partial^2}{\partial \lambda_1 \, \partial l_1} \log \left[\prod_{1 \leqslant i < j \leqslant p} (\lambda_i - \lambda_j)(l_i - l_j) \cdot {}_0F_1(a; \Lambda, L) \right] \geqslant 0. \tag{3.9}$$

Let us define the function

$$G_2(\Lambda, L) := [\det(K(\lambda_i, l_j))]^2 \times \frac{\partial^2}{\partial \lambda_1 \, \partial l_1} \log \left[\prod_{1 \leqslant i < j \leqslant p} (\lambda_i - \lambda_j)(l_i - l_j) \cdot {}_0F_1(a; \Lambda, L) \right]. \tag{3.10}$$

We seek conditions on a for which $G_2(\Lambda, L) \geqslant 0$ for all $\Lambda, L > 0$. Using the explicit determinant formula (3.3) in Theorem 3.2, we have

$$\begin{aligned} G_2(\Lambda, L) = \det(K(\lambda_i, l_j)) \cdot \frac{\partial^2}{\partial \lambda_1 \, \partial l_1} \det(K(\lambda_i, l_j)) \\ - \frac{\partial}{\partial \lambda_1} \det(K(\lambda_i, l_j)) \cdot \frac{\partial}{\partial l_1} \det(K(\lambda_i, l_j)). \end{aligned} \tag{3.11}$$

Theorem 3.3 will be proved by an application of the famous *condensation formula* of C. L. Dodgson [Lewis Carroll] (1866). Dodgson's formula, long familiar to specialists in the theory of determinants (Dwyer, 1951, p. 147), has recently regained prominence due to its reappearance in various combinatorial problems, including the remarkable alternating-sign matrix conjecture; cf., Mills, Robbins, and Rumsey (1983), Robbins and Rumsey (1986), Zeilberger (1997), and Bressoud and Propp (1999).

Let $A = (a_{ij})$ be an $n \times n$ matrix, and denote by $A_r(i, j)$ the $r \times r$ minor of A consisting of r *consecutive* rows and columns of A starting with row i and column j. Then Dodgson's condensation formula is that

$$A_n(1, 1)\, A_{n-2}(2, 2) = A_{n-1}(1, 1)\, A_{n-1}(2, 2) - A_{n-1}(1, 2)\, A_{n-1}(2, 1).$$

Written another way, Dodgson's formula provides that

$$\begin{aligned} &\det(a_{i,j})_{1 \leqslant i, j \leqslant n} \cdot \det(a_{i,j})_{2 \leqslant i, j \leqslant n-1} \\ &\quad = \det(a_{i,j})_{1 \leqslant i, j \leqslant n-1} \cdot \det(a_{i,j})_{2 \leqslant i, j \leqslant n} - \det(a_{i,j})_{\substack{1 \leqslant i \leqslant n-1 \\ 2 \leqslant j \leqslant n}} \cdot \det(a_{i,j})_{\substack{2 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n-1}}. \end{aligned} \tag{3.12}$$

Theorem 3.4. Suppose that a kernel $K: \mathbb{R}^2 \to \mathbb{R}$ is totally positive of order $p+1$. For fixed $\lambda_2 > \cdots > \lambda_p$ and $l_2 > \cdots > l_p$ define the kernel $M: \mathbb{R}^2 \to \mathbb{R}$ by

$$M(x, y) = \begin{vmatrix} K(x, y) & K(x, l_2) & \cdots & K(x, l_p) \\ K(\lambda_2, y) & K(\lambda_2, l_2) & \cdots & K(\lambda_2, l_p) \\ \vdots & \vdots & \vdots & \vdots \\ K(\lambda_p, y) & K(\lambda_p, l_2) & \cdots & K(\lambda_p, l_p) \end{vmatrix}. \tag{3.13}$$

Then M is TP_2 on the region $(\lambda_2, \infty) \times (l_2, \infty)$.

Proof. For $x_1 > x_2 > \lambda_2 > \cdots > \lambda_p$ and $y_1 > y_2 > l_2 > \cdots > l_p$,

$$\begin{vmatrix} M(x_1, y_1) & M(x_1, y_2) \\ M(x_2, y_1) & M(x_2, y_2) \end{vmatrix}$$

$$= \left\| \begin{matrix} \begin{vmatrix} K(x_1, y_1) & K(x_1, l_2) & \cdots & K(x_1, l_p) \\ K(\lambda_2, y_1) & K(\lambda_2, l_2) & \cdots & K(\lambda_2, l_p) \\ \vdots & \vdots & \vdots & \vdots \\ K(\lambda_p, y_1) & K(\lambda_p, l_2) & \cdots & K(\lambda_p, l_p) \end{vmatrix} & \begin{vmatrix} K(x_1, y_2) & K(x_1, l_2) & \cdots & K(x_1, l_p) \\ K(\lambda_2, y_2) & K(\lambda_2, l_2) & \cdots & K(\lambda_2, l_p) \\ \vdots & \vdots & \vdots & \vdots \\ K(\lambda_p, y_2) & K(\lambda_p, l_2) & \cdots & K(\lambda_p, l_p) \end{vmatrix} \\ \begin{vmatrix} K(x_2, y_1) & K(x_2, l_2) & \cdots & K(x_2, l_p) \\ K(\lambda_2, y_1) & K(\lambda_2, l_2) & \cdots & K(\lambda_2, l_p) \\ \vdots & \vdots & \vdots & \vdots \\ K(\lambda_p, y_1) & K(\lambda_p, l_2) & \cdots & K(\lambda_p, l_p) \end{vmatrix} & \begin{vmatrix} K(x_2, y_2) & K(x_2, l_2) & \cdots & K(x_2, l_p) \\ K(\lambda_2, y_2) & K(\lambda_2, l_2) & \cdots & K(\lambda_2, l_p) \\ \vdots & \vdots & \vdots & \vdots \\ K(\lambda_p, y_2) & K(\lambda_p, l_2) & \cdots & K(\lambda_p, l_p) \end{vmatrix} \end{matrix} \right\|.$$

For the three determinants $M(x_1, y_2)$, $M(x_2, y_1)$, and $M(x_2, y_2)$, we perform a sequence of row-column interchanges as follows: In the determinant $M(x_1, y_2)$, we interchange the first column in succession with all other columns, giving us

$$M(x_1, y_2) = (-1)^{p-1} \begin{vmatrix} K(x_1, l_2) & \cdots & K(x_1, l_p) & K(x_1, y_2) \\ K(\lambda_2, l_2) & \cdots & K(\lambda_2, l_p) & K(\lambda_2, y_2) \\ \vdots & \vdots & \vdots & \vdots \\ K(\lambda_p, l_2) & \cdots & K(\lambda_p, l_p) & K(\lambda_p, y_2) \end{vmatrix}.$$

In the determinant $M(x_2, y_1)$ we interchange the first row in succession with all other rows, giving us

$$M(x_2, y_1) = (-1)^{p-1} \begin{vmatrix} K(\lambda_2, y_1) & K(\lambda_2, l_2) & \cdots & K(\lambda_2, l_p) \\ \vdots & \vdots & \vdots & \vdots \\ K(\lambda_p, y_1) & K(\lambda_p, l_2) & \cdots & K(\lambda_p, l_p) \\ K(x_2, y_1) & K(x_2, l_2) & \cdots & K(x_2, l_p) \end{vmatrix}.$$

Finally, in the determinant $M(x_2, y_2)$ we interchange the first column in succession with all other columns and then we follow this by interchanging the resulting first row in succession with all other rows. This gives us the result

$$M(x_2, y_2) = \begin{vmatrix} K(\lambda_2, l_2) & \cdots & K(\lambda_2, l_p) & K(\lambda_2, y_2) \\ \vdots & \vdots & \vdots & \vdots \\ K(\lambda_p, l_2) & \cdots & K(\lambda_p, l_p) & K(\lambda_p, y_2) \\ K(x_2, l_2) & \cdots & K(x_2, l_p) & K(x_2, y_2) \end{vmatrix}.$$

Hence we have

$$
\begin{vmatrix} M(x_1, y_1) & M(x_1, y_2) \\ M(x_2, y_1) & M(x_2, y_2) \end{vmatrix}
= \begin{Vmatrix}
\begin{vmatrix} K(x_1, y_1) & K(x_1, l_2) & \cdots & K(x_1, l_p) \\ K(\lambda_2, y_1) & K(\lambda_2, l_2) & \cdots & K(\lambda_2, l_p) \\ \vdots & \vdots & \vdots & \vdots \\ K(\lambda_p, y_1) & K(\lambda_p, l_2) & \cdots & K(\lambda_p, l_p) \end{vmatrix} & \begin{vmatrix} K(x_1, l_2) & \cdots & K(x_1, l_p) & K(x_1, y_2) \\ K(\lambda_2, l_2) & \cdots & K(\lambda_2, l_p) & K(\lambda_2, y_2) \\ \vdots & \vdots & \vdots & \vdots \\ K(\lambda_p, l_2) & \cdots & K(\lambda_p, l_p) & K(\lambda_p, y_2) \end{vmatrix} \\
\begin{vmatrix} K(\lambda_2, y_1) & K(\lambda_2, l_2) & \cdots & K(\lambda_2, l_p) \\ \vdots & \vdots & \vdots & \vdots \\ K(\lambda_p, y_1) & K(\lambda_p, l_2) & \cdots & K(\lambda_p, l_p) \\ K(x_2, y_1) & K(x_2, l_2) & \cdots & K(x_2, l_p) \end{vmatrix} & \begin{vmatrix} K(\lambda_2, l_2) & \cdots & K(\lambda_2, l_p) & K(\lambda_2, y_2) \\ \vdots & \vdots & \vdots & \vdots \\ K(\lambda_p, l_2) & \cdots & K(\lambda_p, l_p) & K(\lambda_p, y_2) \\ K(x_2, l_2) & \cdots & K(x_2, l_p) & K(x_2, y_2) \end{vmatrix}
\end{Vmatrix}.
$$

A sufficiently attentive reader will have observed by now that these row-column operations are designed to bring the four determinants into a format for application of Dodgson's formula. Indeed, it now follows immediately from (3.12) that the last determinant equals

$$
\begin{vmatrix} K(\lambda_2, l_2) & \cdots & K(\lambda_2, l_p) \\ \vdots & \vdots & \vdots \\ K(\lambda_p, l_2) & \cdots & K(\lambda_p, l_p) \end{vmatrix} \cdot \begin{vmatrix} K(x_1, y_1) & K(x_1, l_2) & \cdots & K(x_1, l_p) & K(x_1, y_2) \\ K(\lambda_2, y_1) & K(\lambda_2, l_2) & \cdots & K(\lambda_2, l_p) & K(\lambda_2, y_2) \\ \vdots & \vdots & \vdots & \vdots & \\ K(\lambda_p, y_1) & K(\lambda_p, l_2) & \cdots & K(\lambda_p, l_p) & K(\lambda_p, y_1) \\ K(x_2, y_1) & K(x_2, l_2) & \cdots & K(x_2, l_p) & K(x_2, y_2) \end{vmatrix}. \tag{3.14}
$$

Since K is TP_p then the first determinant in (3.14) is nonnegative. Notice that the second determinant in (3.14) is of order $p+1$; we interchange its last column in succession with columns $2,\ldots,p$ and, in the resulting determinant, we interchange its last row in succession with rows $2,\ldots,p$. Then the second determinant in (3.14) equals

$$
\begin{vmatrix} K(x_1, y_1) & K(x_1, y_2) & K(x_1, l_2) & \cdots & K(x_1, l_p) \\ K(x_2, y_1) & K(x_2, y_2) & K(x_2, l_2) & \cdots & K(x_2, l_p) \\ K(\lambda_2, y_1) & K(\lambda_2, y_2) & K(\lambda_2, l_2) & \cdots & K(\lambda_2, l_p) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ K(\lambda_p, y_1) & K(\lambda_p, y_1) & K(\lambda_p, l_2) & \cdots & K(\lambda_p, l_p) \end{vmatrix}.
$$

Noting that $x_1 > x_2 > \lambda_2 > \cdots > \lambda_p$ and $y_1 > y_2 > l_2 > \cdots > l_p$ then, since K is TP_{p+1}, it follows that this latter determinant is also nonnegative. Therefore the function M in (3.13) is TP_2 on $(\lambda_2, \infty) \times (l_2, \infty)$.

As a consequence of the preceding result, we obtain the positivity of the function G_2 in (3.10) or (3.11) as a limiting case of Theorem 3.13; this is done by taking limits similar to what was done in (3.8).

Finally, we remark that the results of Theorems 3.1 and 3.3 clearly extend in a straightforward manner to the generalized hypergeometric functions ${}_r\tilde{F}_s$ of two Hermitian matrix arguments.

REFERENCES

1. T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 3rd edn. (Wiley, New York, 2003).
2. J. V. Bondar, *The Monotone Power Function for Tests in MANOVA—The* 2×2 *Case*, unpublished manuscript (Carleton University, Ottawa, 1988).
3. D. Bressoud and J. Propp, How the alternating sign matrix conjecture was solved, *Notices Amer. Math. Soc.* **46**:637–646 (1999).
4. C. L. Dodgson, Condensation of determinants, *Proc. Roy. Soc. London* **15**:150–155 (1866).
5. P. S. Dwyer, *Linear Computations* (Wiley, New York, 1951).
6. N. R. Goodman, Statistical analysis based on a certain multivariate complex Gaussian distribution. (An introduction), *Ann. Math. Statist.* **34**:152–177 (1963).
7. P. Groeneboom and D. R. Truax, A monotonicity property of the power function of multivariate tests, *Indag. Math.* **11**:209–218 (2000).
8. K. I. Gross and D. S. P. Richards, Special functions of matrix argument. I. Algebraic induction, zonal polynomials, and hypergeometric functions, *Trans. Amer. Math. Soc.* **301**:781–811 (1987).
9. K. I. Gross and D. St. P. Richards, Total positivity, spherical series, and hypergeometric functions of matrix argument, *J. Approx. Theory* **59**:224–246 (1989).
10. A. T. James, Distributions of matrix variates and latent roots derived from normal samples, *Ann. Math. Statist.* **35**:475–501 (1964).
11. S. Karlin, *Total Positivity*, Vol. 1 (Stanford University Press, Stanford, CA, 1968).
12. W. H. Mills, D. P. Robbins, and H. Rumsey, Alternating sign matrices and descending plane partitions, *J. Combin. Theory Ser. A* **34**:340–359 (1983).
13. R. J. Muirhead, *Aspects of Multivariate Statistical Theory* (Wiley, New York, 1982).
14. M. D. Perlman and I. Olkin, Unbiasedness of invariant tests for MANOVA and other multivariate problems, *Ann. Statist.* **8**:1326–1341 (1980).
15. D. P. Robbins and H. Rumsey, Determinants and alternating sign matrices, *Adv. in Math.* **62**:169–184 (1986).
16. D. Zeilberger, Dodgson's determinant-evaluation rule proved by two-timing men and women, *Electron. J. Combin.* **4**, Research Paper 22, approx. 2 pp. (1997).