JOURNAL OF MULTIVARIATE ANALYSIS 54, 18–31 (1995)

Large Sample Asymptotic Theory of Tests for Uniformity on the Grassmann Manifold

YASUKO CHIKUSE

Kagawa University, Japan

AND

GEOFFREY S. WATSON

Princeton University

The Grassmann manifold $G_{k,m-k}$ consists of k-dimensional linear subspaces $\mathscr{V}$ in R^m. To each $\mathscr{V}$ in $G_{k,m-k}$, corresponds a unique $m \times m$ orthogonal projection matrix P idempotent of rank k. Let $P_{k,m-k}$ denote the set of all such orthogonal projection matrices. We discuss distribution theory on $P_{k,m-k}$, presenting the differential form for the invariant measure and properties of the uniform distribution, and suggest a general family $F^{(P)}$ of non-uniform distributions. We are mainly concerned with large sample asymptotic theory of tests for uniformity on $P_{k,m-k}$. We investigate the asymptotic distribution of the standardized sample mean matrix U taken from the family $F^{(P)}$ under a sequence of local alternatives for large sample size n. For tests of uniformity versus the matrix Langevin distribution which belongs to the family $F^{(P)}$, we consider three optimal tests—the Rayleigh-style, the likelihood ratio, and the locally best invariant tests. They are discussed in relation to the statistic U, and are shown to be approximately, near uniformity, equivalent to one another. Zonal and invariant polynomials in matrix arguments are utilized in derivations.

1. INTRODUCTION

The Grassmann manifold $G_{k,m-k}$ consists of k-planes, i.e., k-dimensional linear subspaces in R^m. To each k-plane $\mathscr{V}$ in $G_{k,m-k}$, corresponds a unique $m \times m$ orthogonal projection matrix P idempotent of rank k. If k

Received January 4, 1992; revised December 19, 1994.

AMS subject classifications: primary 62H15, 62E20; secondary 58C35, 33C20.

Key words and phrases: Grassmann manifolds, orthogonal projection matrices, invariant measures, Rayleigh-style test, likelihood ratio test, locally best invariant test, zonal and invariant polynomials in matrix arguments.

0047-259X/95 $12.00

column vectors of an $m \times k$ matrix Y are orthonormal ($Y'Y = I_k$), i.e., Y belongs to the Stiefel manifold $V_{k,m}$, and span $\mathscr{V}$, then we have

$$YY' = P, \tag{1.1}$$

which is invariant under the transformation $Y \to YQ$, for $Q \in O(k)$, where $O(k)$ denotes the orthogonal group, i.e., $V_{k,k}$. Let $P_{k,m-k}$ denote the set of all such orthogonal projection matrices.

In this paper, we present some results of distribution theory on $P_{k,m-k}$, and discuss large sample asymptotic theory for the problem of testing uniformity of distributions on $P_{k,m-k}$.

In Section 2, we derive the differential form for the invariant measure (namely the uniform distribution) on $P_{k,m-k}$, based on that on $G_{k,m-k}$ due to James [15], and discuss properties of the uniform distribution and methods for generating non-uniform distributions. We suggest "a general family $F^{(P)}$ of non-uniform distributions", which includes "the matrix Langevin $L^{(P)}(m, k; F)$ distribution" (see (2.13)), having properties analogous to the Downs' [13] distribution on the Stiefel manifold, and the uniform distribution ($F = 0$).

In Section 3, we consider the test for uniformity of distributions on $P_{k,m-k}$, and investigate, for large sample size n, asymptotic behavior of the sample mean matrix $\bar{P} = \sum_{j=1}^{n} P_j/n$ taken from the general family $F^{(P)}$; we note that $\bar{P}$ is a sufficient statistic for the $L^{(P)}(m, k; F)$ distribution. Under a sequence of local alternatives for large n, we derive an asymptotic distribution of the elements of the "standardized" sample mean matrix U, which is a degenerate r-variate normal distribution ($r = m(m+1)/2$).

Section 4 is concerned with testing the uniformity ($\Lambda = 0$) for the $L^{(P)}(m, k; F)$ distribution, with the spectral decomposition (sd) $F = \Gamma \Lambda \Gamma'$. We consider three optimal tests, the Rayleigh-style, the likelihood ratio, and the locally best invariant tests. They are discussed in relation to the statistic U considered in Section 3, and are shown to be approximately, near uniformity, equivalent to one another and distributed asymptotically, for large n, as χ^2_{r-1}. The maximum likelihood estimates (mle's) of the parameters Γ and Λ are also discussed.

A brief discussion is given in the Appendix of the zonal $C_\lambda(A)$ and invariant $C_\phi^{\sigma,\tau}(A, B)$ polynomials in matrix arguments together with the hypergeometric functions ${}_pF_q(a_{[p]}; b_{[q]}; A)$ and ${}_pF_q^{(m)}(a_{[p]}; b_{[q]}; A, B)$ of matrix arguments, which are utilized in derivations throughout the paper.

Before closing this section, we note related works. For asymptotic theory for large concentration Λ in connection with tests on the $L^{(P)}(m, k; F)$ distributions, see Chikuse [7]. On the problem of testing uniformity on Stiefel manifolds, there exists some literature. See Mardia [18], Watson [21], and many others on the hypersphere $V_{1,m}$, and Mardia and Khatri [19], Chikuse [5, 6], and the references in these and the present article on the general $V_{k,m}$.

2. Distribution Theory on the Grassmann Manifold

The Invariant Measure

Let P be an $m \times m$ orthogonal projection matrix in $P_{k,m-k}$, corresponding to a k-plane $\mathscr{V}$ in the Grassmann manifold $G_{k,m-k}$. We write $Y = (y_1, ..., y_k)$ for the matrix $Y \in V_{k,m}$ satisfying (1.1). An invariant measure on $G_{k,m-k}$ is given by the differential form (James [15])

$$(dY) = \bigwedge_{j=1}^{m-k} \bigwedge_{i=1}^{k} y'_{k+j}\, dy_i, \tag{2.1}$$

where we choose an $m \times (m-k)$ matrix $Y_1 = (y_{k+1}, ..., y_m)$ such that $[Y \,\vdots\, Y_1] \in O(m)$. Here, for any matrix $X(= (x_{ij}))$, $dX(= (dx_{ij}))$ denotes the matrix of differentials. The volume of $G_{k,m-k}$ is given, for $\Gamma_k(a) = \pi^{k(k-1)/4} \prod_{j=1}^{k} \Gamma(a - (j-1)/2)$, by

$$g(k, m) = \pi^{k(m-k)/2} \Gamma_k(k/2)/\Gamma_k(m/2). \tag{2.2}$$

Let us derive the invariant measure on $P_{k,m-k}$. Differentiating $P = YY' = \sum_{j=1}^{k} y_j y'_j$ yields

$$dP = \sum_{j=1}^{k} (y_j\, dy'_j + dy_j\, y'_j), \tag{2.3}$$

which is now used to express (2.1) in terms of dP. We thus obtain the *differential form for the invariant measure on* $P_{k,m-k}$ as

$$(dP) = \bigwedge_{j=1}^{m-k} \bigwedge_{i=1}^{k} y'_i\, dP\, y_{k+j}, \tag{2.4}$$

which certainly has the invariance property. Let $[dP]$ denote the normalized invariant measure $(dP)/g(k, m)$ of unit mass on $P_{k,m-k}$; namely, $[dP]$ is the uniform distribution on $P_{k,m-k}$. Probability density functions (Pdf's) of distributions on $P_{k,m-k}$ are expressed with respect to $[dP]$.

The Uniform Distribution

We present some results on the uniform distribution on $P_{k,m-k}$.

Proposition 2.1. (i) *If P is uniformly distributed on $P_{k,m-k}$, so is HPH' for all $H \in O(m)$, fixed or random independent of P, and hence $E(P) = HE(P)H'$, so that we have*

$$E(P) = (k/m)\, I_m. \tag{2.5}$$

(ii) *From the definitions of invariant measures, it follows that*

$$P = YY' \text{ is uniform on } P_{k,m-k}, \quad \text{iff} \quad Y \text{ is uniform on } V_{k,m}, \tag{2.6}$$

and, hence that any uniformly distributed P *on* $P_{k,m-k}$ *is written as*

$$P = Z(Z'Z)^{-1} Z', \tag{2.7}$$

where all the elements of an $m \times k$ *random matrix* Z *are i.i.d. normal* $N(0, 1)$ *(see e.g., Muirhead [20]).*

Proof. The results (i) and (ii) can be proved readily by the known results, and alternatively via the characteristic functions (c.f.'s). We present a formal proof of (ii). The c.f. $\Phi_P(T) = E(\text{etr}\, iPT)$ of P when P is uniform on $P_{k,m-k}$, for an $m \times m$ symmetric matrix T, is given by ${}_1F_1(k/2; m/2; iT)$, which is the same as the c.f. $\Phi_{YY'}(T) = E(\text{etr}\, iYY'T)$ of YY' when Y is uniform on $V_{k,m}$, where $\text{etr}\, A = \exp(\text{tr}\, A)$.

Non-uniform Distributions

Non-uniform distributions may be generated by several methods. Let us give a motivation before suggesting a general family of distributions. Let Z be an $m \times k$ random matrix having the pdf of the form proportional to

$${}_pF_q(a_{[p]}; b_{[q]}; Z'AZ)\, \text{etr}(-Z'BZ), \tag{2.8}$$

with the hypergeometric function ${}_pF_q$ of a matrix argument, A and B being symmetric constant matrices. We shall consider the distribution of the "orientation" $H_Z = Z(Z'Z)^{-1/2}$ $(\in V_{k,m})$ of Z (e.g., Chikuse [4]). In view of Herz [14, Lemma 1.4] (see also Muirhead [20, Theorem 2.1.14]) and the Laplace transform (James [16, (28)]) of the ${}_pF_q$ function, we obtain the pdf of H_Z, which is proportional to

$$|H'_Z B H_Z|^{-m/2}\, {}_{p+1}F_q(a_{[p]}, m/2; b_{[q]}; H'_Z A H_Z (H'_Z B H_Z)^{-1}).$$

It is noted that $A = 0$ gives the matrix angular central Gaussian distribution (Chikuse [4]), and $B = I_m$ gives the pdf proportional to ${}_{p+1}F_q(a_{[p]}, m/2; b_{[q]}; AH_Z H'_Z)$.

The above argument may suggest "a general family $F^{(P)}$ of distributions" on $P_{k,m-k}$ with the pdf being a symmetric homogeneous function, i.e.,

$$g_0^{-1}(F)\, g(FP), \qquad F \text{ being an } m \times m \text{ symmetric matrix}, \tag{2.9}$$

where

$$g(FP) = \sum_{l=0}^{\infty} \sum_{\lambda} d_\lambda C_\lambda(FP)/l!, \quad \text{with } d_{(0)} = 1, \tag{2.10}$$

and the normalizing constant, which is obtained in view of (A.2),

$$g_0(F)=\sum_{l=0}^{\infty}\sum_{\lambda}[d_\lambda(k/2)_\lambda/(m/2)_\lambda\, l!]\; C_\lambda(F), \tag{2.11}$$

where the C_λ are zonal polynomials (see the Appendix). We may be interested, in particular, in the distribution whose pdf is of the form

$${}_pF_q(a_{[p]};b_{[q]};FP)/{}_{p+1}F_{q+1}(a_{[p]},k/2;b_{[q]},m/2;F), \tag{2.12}$$

whose simple special cases are

$$\operatorname{etr}(FP)/{}_1F_1(k/2;m/2;F), \tag{2.13}$$

$$\begin{aligned}&{}_0F_1(k/2;FP)/{}_0F_1(m/2;F),\\ &{}_1F_0(m/2;FP)/{}_0F_1(k/2;F),\end{aligned} \tag{2.14}$$

and

$${}_1F_1(m/2;k/2;FP)/\operatorname{etr}F.$$

We note that, if $Y\in V_{k,m}$ has the matrix Langevin distribution (Downs [13],), then $YY'\in P_{k,m-k}$ has the distribution (2.14). The distribution (2.13) is a slight modification of Downs' [13] distribution on the Stiefel manifold, and may be called "the matrix Langevin distribution on $P_{k,m-k}$," which is denoted by $L^{(P)}(m,k;F)$. It is noted that the $L^{(P)}(m,k;F)$ distribution is derived as the distribution of $P=XX'$ where X is distributed as matrix Bingham; see Jupp and Mardia [17] for the matrix Bingham distribution. We let the sd of F be

$$F=\Gamma\Lambda\Gamma'=\sum_{j=1}^{m}\lambda_j\underset{\sim}{\gamma}_j\underset{\sim}{\gamma}_j', \quad\text{with}\quad \Gamma=(\underset{\sim}{\gamma}_1,\dots,\underset{\sim}{\gamma}_m)\in O(m),$$
$$\text{and}\qquad \Lambda=\operatorname{diag}(\lambda_1,\dots,\lambda_m),\ \lambda_1\geqslant\cdots\geqslant\lambda_m. \tag{2.15}$$

Writing the sd of $D=F_1-F_2$ for $m\times m$ symmetric matrices F_1 and F_2 as $D=\sum_{j=1}^{m}\delta_j\underset{\sim}{d}_j\underset{\sim}{d}_j'$, we have that $\operatorname{tr}DP=\sum_{j=1}^{m}\delta_j(P\underset{\sim}{d}_j)'(P\underset{\sim}{d}_j)$, where $\sum_{j=1}^{m}(P\underset{\sim}{d}_j)'(P\underset{\sim}{d}_j)=k$, which is constant for all $P\in P_{k,m-k}$ iff $\delta_1=\cdots=\delta_m$. Thus, the restriction $\operatorname{tr}F=\operatorname{tr}\Lambda=0$ is imposed to ensure the identifiability of F. We will be concerned with testing the uniformity $\Lambda=0$.

The distribution $L^{(P)}(m,k;F)$ has a unique mode $P_0=\sum_{j=1}^{k}\underset{\sim}{\gamma}_j\underset{\sim}{\gamma}_j'$ which is the closest idempotent of rank k to F, iff

$$\operatorname{rank}F\geqslant k \qquad\text{and}\qquad \lambda_k\gneqq\lambda_{k+1}, \tag{2.16}$$

and we have $\max_P\operatorname{tr}FP=\operatorname{tr}FP_0=\sum_{j=1}^{k}\lambda_j$.

It might be possible to suggest other families of non-uniform distributions on $P_{k,m-k}$. We could make use of (2.6) and (2.7), where Y has some suitable non-uniform distributions on $V_{k,m}$ and the elements of Z are not necessarily i.i.d. normal $N(0, 1)$.

3. Asymptotic Distributions for the Sample Mean Matrix

Given a random sample $P_1, ..., P_n$ from the general family $F^{(P)}$ (2.9) of distributions on $P_{k,m-k}$, we put the "standardized" sample mean matrix

$$U = (n/\alpha)^{1/2} [\bar{P} - (k/m) I_m], \quad \text{with } \bar{P} = \sum_{j=1}^{n} P_j/n$$

$$\text{and} \qquad \alpha = 2k(m-k)/m(m-1)(m+2), \tag{3.1}$$

where we have $\operatorname{tr} U = 0$, and, in particular, $E(U) = 0$ for the uniform distribution $F = 0$ (see (2.5)). Here, we note that the sample mean matrix $\bar{P}$ is a sufficient statistic for the $L^{(P)}(m, k; F)$ distribution (2.13). We are concerned with testing the null hypothesis of uniformity $H_0: F = 0$, against a sequence of local alternatives,

$$H_1: F = F_0/n, \qquad \text{with } F_0 \text{ an } m \times m \text{ symmetric constant matrix.} \tag{3.2}$$

We shall investigate asymptotic behavior of the statistic U under the alternative hypothesis H_1 (3.2) for large n.

Rearrange the $r(= m(m+1)/2)$ distinct elements u_{ij}, $i \leqslant j$, of $U = (u_{ij})$ as

$$\underset{\sim}{u} = (u_{11}, ..., u_{mm}, u_{12}, ..., u_{1m}, u_{23}, ..., u_{m-1,m})'. \tag{3.3}$$

For an $r \times 1$ vector

$$\underset{\sim}{t} = (t_{11}, ..., t_{mm}, t_{12}, ..., t_{1m}, t_{23}, ..., t_{m-1,m})', \tag{3.4}$$

the c.f. $\Phi_u(\underset{\sim}{t}) (= E(\exp i\underset{\sim}{t}'\underset{\sim}{u}))$ of $\underset{\sim}{u}$ is expressed as the c.f. of U

$$\Phi_U(T) = E \operatorname{etr}(iTU) = \operatorname{etr}[-i(kn^{1/2}/m\alpha^{1/2}) T] \{E \operatorname{etr}[i(n\alpha)^{-1/2} TP]\}^n, \tag{3.5}$$

where the elements T_{ij} of the $m \times m$ symmetric matrix $T = (T_{ij})$ are given by

$$T_{ij} = (1 + \delta_{ij}) t_{ij}/2, \qquad \text{with } \delta \text{ being Kronecker's delta.}$$

From (2.9), we have

$$E\,\mathrm{etr}[i(n\alpha)^{-1/2}TP]$$

$$=g_0^{-1}(F)\int_{V_{k,m}} \mathrm{etr}[i(n\alpha)^{-1/2}TYY']\,g(FYY')[dY]$$

$$=g_0^{-1}(F)\Bigg[1+\sum_{\substack{s=0\\(s+t\geq 1)}}^{\infty}\sum_{\sigma}\sum_{t=0}^{\infty}\sum_{\tau}\sum_{\phi\in\sigma\cdot\tau} n^{-(s/2+t)}\alpha^{-s/2}$$

$$\times d_\tau \theta_\phi^{\sigma,\tau} b_\phi C_\phi^{\sigma,\tau}(iT, F_0)/s!\,t!\Bigg], \tag{3.6}$$

with $b_\phi=(k/2)_\phi/(m/2)_\phi$,where $[dY]$ denotes the normalized invariant measure of unit mass on $V_{k,m}$, and the second equality is obtained by utilizing the expansion (A.3) for the ${}_0F_0$ ($=\mathrm{etr}$) function and (2.10), on making the transformation $Y\to HY$, $H\in O(m)$, and integrating over $O(m)$ in view of (A.6). Utilizing the discussion of the zonal C_ϕ and invariant $C_\phi^{\sigma,\tau}$ polynomials in matrix arguments given in the Appendix and referring to the table of zonal polynomials (James [16]), we obtain (though the detailed calculation is omitted)

$$\Phi_U(T)=\exp(\underset{\sim}{t}'\Sigma\underset{\sim}{t}/2)[1+O(n^{-1/2})], \tag{3.7}$$

where

$$\Sigma=\mathrm{diag}(\Sigma_1, I_{r-m}/2), \qquad \text{with}\quad \Sigma_1=I_m-m^{-1}\underset{\sim}{1}_m\underset{\sim}{1}_m' \text{ being of rank } m-1,$$

$$\text{for}\quad \underset{\sim}{1}_m=(1,\dots,1)'\in R^m. \tag{3.8}$$

We note that Σ_1 is diagonalized as

$$Q_0'\Sigma_1Q_0=\mathrm{diag}(0, I_{m-1}), \qquad \text{for}\quad Q_0=[m^{-1/2}\underset{\sim}{1}_m \,\vdots\, Q_1]\in O(m), \tag{3.9}$$

namely, $Q_1Q_1'=I_m-m^{-1}\underset{\sim}{1}_m\underset{\sim}{1}_m'(=\Sigma_1)$. We now consider the $(r-1)\times 1$ random vector

$$\underset{\sim}{w}=G\underset{\sim}{u}, \qquad \text{with}\quad G=\begin{bmatrix} Q_1' & 0 \\ 0 & 2^{1/2}I_{r-m}\end{bmatrix}. \tag{3.10}$$

The c.f. of $\underset{\sim}{w}$, $\Phi_w(\underset{\sim}{s})=E\exp[i(G'\underset{\sim}{s})'\underset{\sim}{u}]$, for an $(r-1)\times 1$ vector $\underset{\sim}{s}=(s_1,\dots,s_{r-1})'$, is expressed as $\Phi_U(T)$, where, for the $r\times 1$ vector

$$G'\underset{\sim}{s}=\underset{\sim}{t}=(t_{11},\dots,t_{mm},t_{12},\dots,t_{1m},t_{23},\dots,t_{m-1,m})', \tag{3.11}$$

we put the $m \times m$ matrix $T = (T_{ij})$, with $T_{ij} = (1 + \delta_{ij}) t_{ij}/2$. Hence, $\Phi_w(\underset{\sim}{s})$ is given by the right-hand side of (3.7) with $\underset{\sim}{t}'\Sigma \underset{\sim}{t} = \underset{\sim}{s}' G \Sigma G' \underset{\sim}{s} = \underset{\sim}{s}'\underset{\sim}{s}$.

Thus, we establish

THEOREM 3.1. *We consider the standardized sample mean matrix U, defined by* (3.1) *constructed from a random sample of size n from the general family* $F^{(P)}$ ((2.9)) *of distributions on* $P_{k,m-k}$, *under the local alternative hypothesis* (3.2). *The vector* $\underset{\sim}{u}$, *defined by* (3.3) *consisting of the* $r(=m(m+1)/2)$ *distinct elements of U, is distributed as degenerate multivariate normal* $N_r(\underset{\sim}{0}, \Sigma)$, *asymptotically for large n, where* Σ *is given in* (3.8). *Therefore, the* $(r-1) \times 1$ *vector* $\underset{\sim}{w}$ *defined by* (3.10) *has the multivariate normal* $N_{r-1}(\underset{\sim}{0}, I_{r-1})$ *distribution, asymptotically for large n.*

For certain distributions (including the $L^{(P)}(m, k; F)$ distribution) on $P_{k,m-k}$, the best critical region by the Neyman-Pearson lemma for testing uniformity may be given based on tr $F\bar{P}$. We give the following

COROLLARY 3.1. *Under the condition of Theorem* 3.1, *the statistic* $w = c^{-1/2}\,\mathrm{tr}(BU)$, *where B is an arbitrary* $m \times m$ *symmetric and non-spherical constant matrix, and* $c = \mathrm{tr}\, B^2 - (\mathrm{tr}\, B)^2/m$, *is distributed as normal* $N(0, 1)$, *asymptotically for large n.*

Remark 3.1. We employed the local alternatives of order $O(n^{-1})$; the case of the local alternatives of order $O(n^{-1/2})$ may be valid also, though. This fact may be compared with the local alternative hypothesis of order $O(n^{-1/2})$ for the uniformity test on the Stiefel manifold (see Chikuse [6]).

Remark 3.2. The correction terms of $O(n^{-j/2})$, $j = 1, 2$, in Theorem 3.1 and Corollary 3.1 are obtained by utilizing the theory of zonal and invariant polynomials in matrix arguments. These are expressed in terms of the zonal and invariant polynomials and the classical Hermite polynomials, but are omitted here, since our interest is the asymptotic property of the matrix variate U which plays important roles in the tests for uniformity on $P_{k,m-k}$ in Section 4.

Remark 3.3. We can give an alternative proof of Theorem 3.1, following the method of Anderson and Stephens [1]. Deriving the pdf of the matrix variate

$$V = U + m^{-1/2} v I_m,$$

where v is a normal $N(0, 1)$ variate independent of U, transforming the distinct elements of V to $(\underset{\sim}{w}, v)$, and then integrating v out from the pdf of $(\underset{\sim}{w}, v)$, we obtain the desired asymptotic pdf of $\underset{\sim}{w}$.

4. Approximate Properties of Optimal Tests for Uniformity

Rayleigh-Style Test

It has been established, from Theorem 3.1, that the $(r-1)\times 1$ vector $\underset{\sim}{w}$ defined by (3.10) is asymptotically distributed as multivariate normal $N_{r-1}(\underset{\sim}{0}, I_{r-1})$ for large n under the null hypothesis H_0 of uniformity. Thus, the statistic $\underset{\sim}{w}'\underset{\sim}{w}$ can be regarded as a "Raleigh-style test statistic" for testing uniformity on $P_{k,m-k}$, and has, for large n, the asymptotic χ^2_{r-1} distribution under H_0.

Following the notation in Section 3, with $\underset{\sim}{u}=(\underset{\sim}{u}'_1, \underset{\sim}{u}'_2)'$ for $\underset{\sim}{u}_1$ being $m\times 1$, we can express $\underset{\sim}{w}'\underset{\sim}{w}=\underset{\sim}{u}'_1\Sigma_1\underset{\sim}{u}_1+2\underset{\sim}{u}'_2\underset{\sim}{u}_2$, and hence

$$\operatorname{tr} U^2=\underset{\sim}{u}'_1\underset{\sim}{u}_1+2\underset{\sim}{u}'_2\underset{\sim}{u}_2=\underset{\sim}{w}'\underset{\sim}{w}+z_0, \qquad \text{with} \quad z_0=\underset{\sim}{u}'_1(I_m-\Sigma_1)\,\underset{\sim}{u}_1. \tag{4.1}$$

From the definitions of $\underset{\sim}{u}_1$ and Σ_1, it is readily seen that $z_0=0$, yielding that

$$\underset{\sim}{w}'\underset{\sim}{w}=\operatorname{tr} U^2. \tag{4.2}$$

Maximum Likelihood Estimates and Likelihood Ratio Test

Letting $P_1, ..., P_n$ be a random sample from the $L^{(P)}(m, k; F)$ distribution with the sd $F=\Gamma\Lambda\Gamma'$, the log-likelihood is given by

$$\log L=n \operatorname{tr} \Lambda\Gamma'\bar{P}\Gamma-n\log {}_1F_1(\Lambda),$$

$$\text{with the notation } {}_1F_1(\Lambda)={}_1F_1(k/2; m/2; \Lambda) \text{ and } \bar{P}=\sum_{j=1}^{n} P_j/n. \tag{4.3}$$

Using the sd $\bar{P}=RDR'$, where $R\in O(m)$ and $D=\operatorname{diag}(d_1, ..., d_m)$, $d_1>\cdots>d_m>0$ (with probability one), it is seen that $\max_\Gamma \operatorname{tr} \Lambda\Gamma'\bar{P}\Gamma=\operatorname{tr}\Lambda R'\bar{P}R=\sum_{j=1}^{m}\lambda_j d_j$ and we have the mle $\hat{\Gamma}=R$ of Γ. Next, differentiating

$$n\sum_{j=1}^{m} d_j\lambda_j-n\log {}_1F_1(\Lambda)+n\eta\sum_{j=1}^{m}\lambda_j, \qquad \text{with } \eta \text{ the Lagrange's multiplier,} \tag{4.4}$$

with respect to λ_j, $j=1, ..., m$, we can show that the mle $\hat{\Lambda}=\mathrm{diag}(\hat{\lambda}_1, ..., \hat{\lambda}_m)$ of Λ satisfies

$$\partial \log {}_1F_1(\hat{\Lambda})/\partial\hat{\lambda}_j = d_j + \eta, \qquad j=1, ..., m, \quad \text{with} \sum_{j=1}^{m} \hat{\lambda}_j = 0. \tag{4.5}$$

We are concerned with testing the null hypothesis of uniformity $H_0\colon \Lambda = 0$ against the alternative $H_1 = \Lambda \neq 0$, when Γ is unknown. We have the likelihood ratio test statistic

$$-2\log L^* = 2n\left[-\log {}_1F_1(\hat{\Lambda}) + \sum_{j=1}^{m} \hat{\lambda}_j d_j\right], \tag{4.6}$$

which is known to be distributed approximately as χ^2_{r-1} for large n. For small Λ, that is, for a slight departure from the null hypothesis H_0, on expanding the ${}_1F_1$ function in terms of the λ_j's referring to the table of zonal polynomials (James [16]) and then differentiating, (4.5) becomes

$$\alpha\hat{\lambda}_j + O(\hat{\Lambda}^2) = d_j + \eta, \qquad j=1, ..., m, \quad \text{with} \sum_{j=1}^{m} \hat{\lambda}_j = 0, \tag{4.7}$$

where α is given in (3.1). Ignoring the terms of $O(\hat{\Lambda}^2)$, we have the solution of (4.7) as

$$\hat{\underset{\sim}{\lambda}} \doteqdot [\underset{\sim}{d} - (k/m)\,\underset{\sim}{1}_m]/\alpha, \tag{4.8}$$

where $\hat{\underset{\sim}{\lambda}} = (\hat{\lambda}_1, ..., \hat{\lambda}_m)'$ and $\underset{\sim}{d} = (d_1, ..., d_m)'$. Substituting the approximate solution (4.8) into (4.6) we have

$$-2\log L^* \doteqdot (n/\alpha)[\underset{\sim}{d} - (k/m)\,\underset{\sim}{1}_m]'\,[\underset{\sim}{d} - (k/m)\,\underset{\sim}{1}_m] = \mathrm{tr}\, U^2, \tag{4.9}$$

where the second identity is obtained from the fact that the elements of $(n/\alpha)^{1/2}\,[\underset{\sim}{d} - (k/m)\,\underset{\sim}{1}_m]$ are the latent roots of U in the notation of Section 3.

Locally Best Invariant Test

We consider the locally best invariant test (Beran [2]) for the null hypothesis H_0 of uniformity against the invariant hypothesis H_1 of the $L^{(P)}(m, k; F)$ distribution with the sd $F = \Gamma\Lambda\Gamma'$ with the restriction $\mathrm{tr}\,\Lambda = 0$. Given a random sample $P_1, ..., P_n$, the likelihood ratio $L(P_1, ..., P_n)$ is, with the statistic U defined by (3.1),

$$L(P_1, ..., P_n) \propto \mathrm{etr}\, nF\bar{P} = \mathrm{etr}[(nk/m)F]\;\mathrm{etr}(n\alpha)^{1/2}\, FU. \tag{4.10}$$

The Neyman-Pearson lemma shows that the best invariant test is to reject H_0 for large values of

$$\int_{O(m)} L(HP_1H', ..., HP_nH')[dH] \propto \int_{O(m)} \mathrm{etr}(n\alpha)^{1/2} FHUH'[dH]$$
$$= {}_0F_0^{(m)}((n\alpha)^{1/2} \Lambda, U) = 1 + n\alpha \sum_{\phi \vdash 2} C_\phi(\Lambda)\, C_\phi(U)/2C_\phi(I_m) + O(\Lambda^3), \tag{4.11}$$

from (A.5) and the fact tr $U = 0$.

Thus, the locally best invariant test statistic must be a suitable increasing function of

$$\sum_{\phi \vdash 2} C_\phi(\Lambda)\, C_\phi(U)/C_\phi(I_m), \qquad \text{with} \quad \phi = (2), (1^2),$$
$$= 2(\mathrm{tr}\, \Lambda^2)\, \mathrm{tr}\, U^2/(m-1)(m+2), \tag{4.12}$$

where we refer to the table (James [16]) of zonal polynomials, utilizing the fact tr $\Lambda = 0$.

Summarizing the above results, we establish

THEOREM 4.1. *Three optimal tests can be proposed for testing the uniformity* ($\Lambda = 0$) *for the* $L^{(P)}(m, k; F)$ *distribution, that is, the Rayleigh-style, the likelihood ratio, and the locally best invariant tests. It is shown that the latter two test statistics are approximately, near the uniformity* (*i.e., for small* Λ), *equivalent to the former test statistic* tr U^2, *which is distributed asymptotically, for large n, as* χ^2_{r-1}.

We are, in the present paper, concerned with the tests for uniformity ($\Lambda = 0$) without any knowledge of the other parameters than Λ (e.g., Γ for the $L^{(P)}(m, k; F)$ distribution with $F = \Gamma \Lambda \Gamma'$). We shall consider, in the other paper, inferential problems on Λ including the test for uniformity ($\Lambda = 0$), when knowledge of the other parameters than Λ is given or when the form of the population distribution is partially specified.

APPENDIX: ZONAL AND INVARIANT POLYNOMIALS IN MATRIX ARGUMENTS

The zonal polynomials $C_\lambda(A)$ in an $m \times m$ symmetric matrix A were defined by the theory of group representations of the real linear group $GL(m, R)$ of $m \times m$ nonsingular matrices on the vector space of homogeneous polynomials of degree l on the space of $m \times m$ symmetric matrices (James [16]). Here, $[2\lambda]$ indexes each irreducible representation, where $\lambda \vdash l$, that is, λ is an ordered partition of l, i.e., $\lambda = (l_1, ..., l_m)$, $l_1 \geqslant \cdots \geqslant l_m \geqslant 0$, $\sum_{j=1}^{m} l_j = l$. The polynomials $C_\lambda(A)$, $\lambda \vdash l = 0, 1, ...$, are invariant under the transformation $A \to HAH'$, $H \in O(m)$. We note that, when the rank of A is $r(\leqslant m)$, then $C_\lambda(A) = 0$ if $l_{r+1} \neq 0$. The polynomial $(\mathrm{tr}\, A)^l$ has a unique decomposition

$$(\mathrm{tr}\, A)^l = \sum_\lambda C_\lambda(A), \tag{A.1}$$

where the sum $\sum_\lambda$ ranges over all $\lambda \vdash l$. The following is a basic property:

$$\int_{O(m)} C_\lambda(AHBH')[dH] = C_\lambda(A)\, C_\lambda(B)/C_\lambda(I_m). \tag{A.2}$$

The hypergeometric function ${}_pF_q(a_{[p]}; b_{[q]}; A)$ of an $m \times m$ symmetric matrix A has a series representation (Constantine [9])

$$\sum_{l=0}^{\infty} \sum_\lambda \left[\prod_{j=1}^{p} (a_j)_\lambda \Big/ \prod_{j=1}^{q} (b_j)_\lambda\, l! \right] C_\lambda(A), \tag{A.3}$$

where the a_j and b_j are real or complex numbers, and $(a)_\lambda = \prod_{j=1}^{m} (a-(j-1)/2)_{l_j}$, with $(a)_l = a(a+1)\cdots(a+l-1)$. The hypergeometric function of two symmetric matrix arguments A and B is defined by

$${}_pF_q^{(m)}(a_{[p]}; b_{[h]}; A, B) = \int_{O(m)} {}_pF_q(a_{[p]}; b_{[q]}; AHBH')[dH], \tag{A.4}$$

and has a series representation

$$\sum_{l=0}^{\infty} \sum_\lambda \left[\prod_{j=1}^{p} (a_j)_\lambda \Big/ \prod_{j=1}^{q} (b_j)_\lambda\, l!\, C_\lambda(I_m) \right] C_\lambda(A)\, C_\lambda(B). \tag{A.5}$$

The invariant polynomials $C_\phi^{\sigma,\tau}(A, B)$ in two $m \times m$ symmetric matrix arguments A and B were defined (Davis [10,11]), extending the zonal polynomials, by the theory of group representations of $GL(m, R)$ on the vector space of polynomials in two arguments, homogeneous of degree s, t on the space of $m \times m$ symmetric matrices. We write $\phi \in \sigma \cdot \tau$ to indicate that the irreducible representation indexed by $[2\phi](\phi \vdash (s+t))$ occurs (possibly with multiplicity greater than one) in the decomposition of the Kronecker product $[2\sigma] \otimes [2\tau]$ of the irreducible representations indexed by $[2\sigma]$ and $[2\tau]$ into a direct sum of irreducible representations of $GL(m, R)$. The $C_\phi^{\sigma,\tau}(A, B)$ are invariant under the simultaneous transformations $A \to HAH'$, $B \to HBH'$, $H \in O(m)$. The $C_\phi^{\sigma,\tau}$ polynomials were extended to the invariant polynomials $C_\phi^{\sigma_1, ..., \sigma_r}(A_1, ..., A_r)$ in more than two matrix arguments $A_1, ..., A_r$ by Chikuse [3]; also see Chikuse and Davis [8]. The polynomials have been tabulate up to three matrices and the first few degrees in Davis [10, 12].

We have a basic result (Davis [10, (1.2)]),

$$\int_{O(m)} C_\sigma(AHSH')\, C_\tau(BHTH')[dH] = \sum_{\phi \in \sigma \cdot \tau} C_\phi^{\sigma, \tau}(A, B)\, C_\phi^{\sigma, \tau}(S, T)/C_\phi(I_m), \tag{A.6}$$

where the sum $\sum_{\phi \in \sigma \cdot \tau}$ ranges over all $[2\phi]$ occurring in $[2\sigma] \otimes [2\tau]$, and

$$C_\phi^{\sigma, \tau}(A, A) = \theta_\phi^{\sigma, \tau} C_\phi(A), \qquad \text{Davis [10, (2.1)]}). \tag{A.7}$$

Acknowledgments

The authors are grateful to the referees for some helpful comments. The work of G. S. Watson is supported by NSF Grant DMS 9212415.

References

[1] Anderson, A. W., and Stephens, M. A. (1972). Tests for randomness of directions against equatorial and bimodal alternatives. *Biometrika* **59** 613–621.

[2] Beran, R. J. (1968). Testing for uniformity on a compact homogeneous space. *J. Appl. Prob.* **5** 177–195.

[3] Chikuse, Y. (1980). Invariant polynomials with matrix arguments and their applications. In *Multivariate Stastical Analysis* (R. P. Gupta, Ed.), pp. 53–68. North-Holland, Amsterdam.

[4] Chikuse, Y. (1990). The matrix angular central Gaussian distribution. *J. Multivariate Anal.* **33** 265–274.

[5] Chikuse, Y. (1991). High dimensional limit theorems and matrix decompositions on the Stiefel manifold. *J. Multivariate Anal.* **36** 145–162.

[6] Chikuse, Y. (1991). Asymptotic expansions for distributions of the large sample matrix resultant and related statistics on the Stiefel manifold. *J. Multivariate Anal.* **39** 270–283.

[7] Chikuse, Y. (1993). Asymptotic theory for the concentrated Langevin distributions on the Grassmann manifold. In *Statistical Sciences and Data Analysis* (K. Matsusita *et al.*, Eds.), pp. 237–245. VSP, Zeist.

[8] Chikuse, Y. and Davis, A. W. (1986). Some properties of invariant polynomials with matrix arguments and their applications in econometrics. *Ann. Inst. Statist. Math. A* **38** 109–122.

[9] Constantine, A. G. (1963). Some non-central distribution problems in multivariate analysis. *Ann. Math. Statist.* **34** 1270–1285.

[10] Davis, A. W. (1979). Invariant polynomials with two matrix arguments extending the zonal polynomials: Applications to multivariate distribution theory. *Ann. Inst. Statist. Math. A* **31** 465–485.

[11] DAVIS, A. W. (1980). Invariant polynomials with two matrix arguments, extending the zonal polynomials. In *Multivariate Analysis*, Vol. 5 (P. R. Krishnaiah, Ed.), pp. 287–299. North-Holland, Amsterdam.

[12] DAVIS, A. W. (1981). On the construction of a class of invariant polynomials in several matrices, extending the zonal polynomials. *Ann. Inst. Statist. Math. A* **33** 297–313.

[13] DOWNS, T. D. (1972). Orientation statistics. *Biometrika* **59** 665–676.

[14] HERZ, C. S. (1955). Bessel functions of matrix argument. *Ann. Math.* **61** 474–523.

[15] JAMES, A. T. (1954). Normal multivariate analysis and the orthogonal group. *Ann. Math. Statist.* **25** 40–75.

[16] JAMES, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.* **35** 475–501.

[17] JUPP, P. E., AND MARDIA, K. V. (1976). Maximum likelihood estimators for the matrix von Mises-Fisher and Bingham distributions. *Ann. Statist.* **7** 599–606.

[18] MARDIA, K.V. (1972). *Statistics of Directional Data*. Academic Press, New York.

[19] MARDIA, K. V., AND KHATRI, C. G. (1977). Uniform distribution on a Stiefel manifold. *J. Multivariate Anal.* **7** 468–473.

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

[21] WATSON, G. S. (1983). *Statistics on Spheres*. Lecture Notes in Mathematics, Vol. 6. Wiley, New York.