The Annals of Statistics
1993, Vol. 21, No. 2, 848–866

MULTIVARIATE ASPECTS OF SHAPE THEORY

BY COLIN R. GOODALL[1] AND KANTI V. MARDIA[2]

Pennsylvania State University and University of Leeds

We place shape theory in the setting of noncentral multivariate analysis, and thus provide a comprehensive view of shape distributions when landmark coordinates are Gaussian distributed. This work allows the statistical analysis of shape to be carried out using standard techniques of multivariate analysis. The paper includes some new results in all dimensions and a general Gaussian approximation to the size-and-shape distribution. We also discuss some inference problems and give a numerical example.

1. Introduction. Let X: $N \times K$ be a matrix representing the geometrical figure comprising N landmark, or labeled, points in $\Re^K$. Assume that X has the isotropic matrix multivariate Gaussian distribution with mean μ_X,

$$X \sim N(\mu_X, \sigma^2 I_N \otimes I_K). \tag{1.1}$$

The *shape* of X is the figure with translation, orientation and scale removed, and the space of shapes is the quotient of $(\Re^K)^N$ by the similarity group [Kendall (1984)]. Shape coordinates $\mathbf{u}$ of X are constructed in several steps summarized in the expression

$$LX = Y = TH = rWH = rW(\mathbf{u})H, \tag{1.2}$$

which we now discuss more fully. (Note that shape coordinates of μ_X are defined analogously.) The matrix L is $(N-1) \times N$ with orthonormal rows orthogonal to $\mathbf{1} = (1, 1, \dots, 1)^t$. L may be a submatrix of the Helmert matrix. Let $\mu = L\mu_X$. Then Y: $(N-1) \times K$ is invariant to translations of the figure X,

$$Y \sim N(\mu, \sigma^2 I_{N-1} \otimes I_K). \tag{1.3}$$

Let $n = \min(N-1, K)$ and $p = \operatorname{rank} \mu$. In (1.2), $Y = TH$ is the QR decomposition where T: $(N-1) \times n$ is lower triangular with $t_{ii} \geq 0$, $i = 1, \dots, \min(n, K-1)$, and where H: $n \times K$, $H \in V_{n,k}$, the Stiefel manifold. Typically in shape analysis there are more landmarks than dimensions ($N > K$). H acts on the right to transform $\Re^K$. This is the converse to the usual

Received November 1990; revised January 1992.

[1]Partially supported by Army Research Office (Durham) Grant DAAL03-88-K-0045 to Princeton University (J. W. Tukey, P.I.), by NSF Grant DMS-92-08656 to Pennsylvania State University and by additional funds from University of Leeds.

[2]Partially supported by NSF Grant DMS-88-03207 to Princeton University (G. S. Watson, P.I.), by NSF Grant DMS-92-08656 to Pennsylvania State University and by the Center for Multivariate Analysis at Pennsylvania State University.

AMS 1991 *subject classifications*. Primary 62H10; secondary 62H11, 62E15, 62E17.

Key words and phrases. Shape, size-and-shape, zonal polynomial, Gaussian model, configuration, singular values, distribution theory, polar coordinates, small variations, QR decomposition.

case in multivariate analysis, where the transformation acts on the left. We may also imagine that each landmark is a variable, and each dimension an observation, so that X^t or Y^t is a typical multivariate data matrix. However in shape analysis there are usually more variables than observations, and the usual setup in multivariate analysis corresponds to simplices or fewer landmarks, that is, triangles in $\mathfrak{R}^2, \mathfrak{R}^3, \dots$, tetrahedra in $\mathfrak{R}^3, \mathfrak{R}^4, \dots$. Our results will include both situations.

The matrix T is invariant to translations and rotations of X, so is a *size-and-shape matrix*, and the elements of T are size-and-shape coordinates of X. When $N - 1 \geq K$ then H is $K \times K$ and we distinguish two cases.

1. H includes reflection, $H \in O(K)$, $|H| = \pm 1$, $t_{KK} \geq 0$, and T, written T^R for definiteness, contains *reflection size-and-shape coordinates*.
2. H excludes reflection, $H \in SO(K)$, $|H| = +1$, t_{KK} is not restricted, and T may be written T^{NR} for definiteness.

When $n < K$, the usual case in multivariate analysis, there is no distinction to make.

To form the *shape matrix* W we divide T by its size, the centroid size of X, $r = \|T\| = \sqrt{\operatorname{tr} T^tT} = \|Y\|$. W may include reflection, $W^R = T^R/r$, or exclude reflection, $W^{NR} = T^{NR}/r$. Since $\|W\| = 1$, the elements of W are a "direction vector" for shape, and $\mathbf{u}$ comprises $m = (N-1)K - nK + n(n+1)/2 - 1$ generalized polar coordinates.

Statistical inference for shape is complicated by the nuisance parameters for translation, scale and orientation. These parameters may be estimated explicitly, as in the full maximum-likelihood approach using Procrustes methods [Goodall (1991)]. Alternatively, the nuisance parameters may be integrated out, and inference is based on the marginal shape distribution, of W^{NR}; this is the approach we take here. The shape density has been obtained for planar figures, $K = 2$, comprising 3 or more landmarks, $N \geq 3$: Small (1981) and Kendall (1984) show that when $\mu = 0$ the shape density is uniform with respect to the natural Riemannian metric on Kendall's *shape space* Σ_2^N [see also Kendall (1989) and Le and Kendall (1993)]. Mardia (1989a, b) and Mardia and Dryden (1989a) derive the shape density with $K = 2$ and μ is unrestricted and give a Gaussian approximation for small σ [see also Bookstein (1986)]. Le (1990) uses a stochastic calculus argument, and Goodall and Mardia (1991a, 1992) the QR decomposition.

Let ρ be the Riemannian distance between the shapes of X and μ_X,

$$\cos\rho = \max_{H'} \operatorname{tr} \mu^t Y H'/\xi r, \tag{1.4}$$

where $\xi = \|\mu\|$, and the maximum is taken over $H' \in O(K)$ or $H' \in SO(K)$ [see Carne (1990)]. The Procrustes distance between the shapes of X and μ_X is $2\sin\rho/2$. Given (1.1), when $K = 2$ and $\mu = 0$, then $\cot^2\rho$ is central-F distributed [Le (1990)], and, more generally, for any K and $p \leq 1$, $\cot^2\rho$ is noncentral-F distributed, and conditional on ρ the shape density is uniform [Goodall and Mardia (1992)]. However, since Σ_2^N is homogeneous [Kendall

(1984)], the shape density is the same for all μ, $\mu \neq 0$, when $K = 2$ [see Goodall (1991) and Goodall and Mardia (1991a)], and is a function only of ρ. This is not true when $K \geq 3$, and the shape density depends on μ also. The noncentral Bartlett decomposition used by Goodall and Mardia (1992) does not generalize readily to $p \geq 2$ [see Kshirsagar (1963)]. This pattern of results recurs below, where we give a systematic and extended treatment of shape distributions using noncentral multivariate analysis.

This paper includes new results in all dimensions, a general Gaussian approximation, and some extensions to multivariate theory. Basic references are James (1964), Muirhead (1982), and Farrell (1985), on matrix hypergeometric functions. To start, when $n = N - 1 \leq K$ the cross product matrix $B = YY^t$: $n \times n$ has the noncentral Wishart distribution

$$B \sim W_n(K, \sigma^2 I; \mu\mu^t) \tag{1.5}$$

and is invariant to orientation and reflection. The density is matrix hypergeometric, type ${}_0F_1$ Bessel, and essentially (1.5) is the reflection size-and-shape distribution of X; a coordinate transformation yields the density in terms of T (Section 2). However, there are three differences between our work and standard multivariate theory.

1. The noncentral Wishart density does not exist when $N - 1 \geq K$, and we therefore use the size-and-shape matrix T.
2. To obtain the density of T^{NR}, we integrate over $SO(K)$. But a zonal polynomial is an integral over $O(K)$. These integrals are the same when $n < K$, and, for $n \geq K$ and $p < K$, the integral (of $\operatorname{etr}(\mu^t TH/\sigma^2)$) over $O(K)$ is twice the integral over $SO(K)$ (Subsection 2.1). This leaves the case $p = K$.
3. To obtain the shape density we integrate out size. The result is a familiar special function only when $p = 1$.

We give the general distribution theory in Section 2, then specialize to $p = 0$, 1 and 2 or 3 in Sections 3, 4, and 5. Two variants are briefly considered:

(i) The *affine shape* distributions (Section 6) are the marginal distributions modulo affine transformations.

(ii) The singular values of Y and $Y/\|Y\|$ are useful lower-dimensional summaries of size-and-shape and shape; their distributions are given in Section 7.

Our results extend easily (Section 8) to the anisotropic case of correlated landmarks, $\operatorname{cov} X = \Sigma \otimes I_K$. Finally (Section 9) we discuss inference using an approximation for small σ, and give an example in three dimensions.

2. Size-and-shape distributions. The Jacobian $J(Y \to T, H)$ is given by

$$(dY) = \prod_{i=1}^{n} t_{ii}^{K-i}(dT)(H\,dH^t), \tag{2.1}$$

where (HdH^t) defines the unnormalized invariant probability measure on $V_{n,K}$. Muirhead (1982) gives (2.1) for $N-1 \le K$ $(n = N-1)$, but it is also true when $N-1 > K$. The distribution of the reflection size-and-shape coordinates T is as follows.

THEOREM 1. *With the setup above, the (reflection) size-and-shape distribution is type* ${}_0F_1$ *(Bessel) hypergeometric, with density*

$$(2.2)\qquad \frac{\prod_{i=1}^{n} t_{ii}^{K-i}}{\sigma^M 2^{(M-2n)/2}\pi^{(M-nK)/2}\Gamma_n(K/2)} \operatorname{etr}\left\{-\frac{T^tT}{2\sigma^2} - \frac{\mu^t\mu}{2\sigma^2}\right\} {}_0F_1\left(\frac{K}{2}; A\right),$$

where $M = (N-1)K$, $A = \mu^t TT^t\mu/4\sigma^4$ *and*

$$\Gamma_n\left(\frac{K}{2}\right) = \pi^{n(n-1)/4}\prod_{i=0}^{n-1}\Gamma\left(\frac{K-i}{2}\right)$$

is the usual multivariate gamma function [for this and related definitions, we use Muirhead (1982)]. *The first few coefficients* $\Gamma_K(K/2)$ *are* π, $\pi^{5/2}/2$ *and* $\pi^4/2$ *for* $K = 2$, 3 *and* 4.

PROOF. The joint density of T and H is, from (1.3) and (2.1),

$$(2.3)\qquad \frac{\prod^n t_{ii}^{K-i}}{\sigma^M(2\pi)^{M/2}} \operatorname{etr}\left\{-\frac{T^tT}{2\sigma^2} - \frac{\mu^t\mu}{2\sigma^2} + \frac{\mu^tTH}{\sigma^2}\right\}(dT)(HdH^t).$$

The following integral [James (1964) or Muirhead (1982)] then gives (2.2):

$$(2.4)\qquad \int_{V_{n,K}} \operatorname{etr}\left(\frac{\mu^tTH}{\sigma^2}\right)(HdH^t) = \frac{2^n\pi^{nK/2}}{\Gamma_n(K/2)}\,{}_0F_1\left(\frac{K}{2}; A\right).$$

Alternatively, when $N-1 \le K$, the density of B in (1.5) is

$$(2.5)\qquad \frac{|B|^{(K-n-1)/2}}{\sigma^M 2^{M/2}\Gamma_n(K/2)} \operatorname{etr}\left\{-\frac{B}{2\sigma^2} - \frac{\mu\mu^t}{2\sigma^2}\right\} {}_0F_1\left(\frac{K}{2}; A\right)$$

in which we substitute $B \doteq TT^t$, $(dB) = 2^n\prod_{i=1}^n t_{ii}^{n+1-i}(dT)$ and $|B| = \prod^n t_{ii}^2$. □

Turning to (reflection) shape, we write $J(T \to r, \mathbf{u}) = r^m J(\mathbf{u})$.

THEOREM 2. *With the setup above, the (reflection) shape density is*

$$(2.6)\qquad \frac{\prod^n w_{ii}^{K-i}J(\mathbf{u})}{\pi^{(M-nK)/2}\Gamma_n(K/2)} \operatorname{etr}\left\{-\frac{\mu^t\mu}{2\sigma^2}\right\}\sum_{k=0}^{\infty}\frac{2^{n+k-1}\Gamma(M/2+k)}{k!}\sum_{\kappa}\frac{C_\kappa(A^*)}{(K/2)_\kappa},$$

where $A^* = \mu^tWW^t\mu/4\sigma^2 = \sigma^2A/r^2$, $C_\kappa(A^*)$ *is the zonal polynomial of* A^* *corresponding to the partition* $\kappa = (k_1, \dots, k_l)$ *of* k, *with* $\sum_1^l k_i = k$, *and* $(K/2)_\kappa$ *is the generalized hypergeometric coefficient, or product of Pochhammer symbols*, $\prod_{i=0}^{l-1}((K-i)/2)_{k_{i+1}} = \prod_{i=0}^{l-1}\Gamma((K-i)/2 + k_{i+1})/\Gamma((K-i)/2)$.

PROOF. Expanding ${}_0F_1$ as a sum of zonal polynomials [James (1964)], and with the change of variables, (2.2) yields the joint density of $(v, \mathbf{u})$,

$$\begin{aligned}(2.7)\qquad &\frac{r^{M-1}\prod^n W_{ii}^{K-i} J(\mathbf{u})}{\sigma^M 2^{(M-2n)/2}\pi^{(M-nK)/2}\Gamma_n(K/2)}\\ &\quad\times \exp\left\{-\frac{r^2}{2\sigma^2}\right\}\operatorname{etr}\left\{-\frac{\mu^t\mu}{2\sigma^2}\right\}\sum_{k=0}^{\infty}\sum_{\kappa}\frac{C_\kappa(r^2A^*/\sigma^2)}{(K/2)_\kappa k!}.\end{aligned}$$

But $C_\kappa(r^2A^*/\sigma^2) = C_\kappa(A^*)(r/\sigma)^{2k}$, and, collecting powers of r, (2.7) is

$$\begin{aligned}(2.8)\qquad &\frac{\prod^n W_{ii}^{K-i} J(\mathbf{u}) 2^n}{\pi^{(M-nK)/2}\Gamma_n(K/2)}\operatorname{etr}\left\{-\frac{\mu^t\mu}{2\sigma^2}\right\}\sum_{k=0}^{\infty}\frac{r^{M+2k-1}}{2^{M/2}\sigma^{M+2k}}\\ &\quad\times e^{-r^2/2\sigma^2}\sum_{\kappa}\frac{C_\kappa(A^*)}{(K/2)_\kappa k!}.\end{aligned}$$

With the substitution $v = r^2/2\sigma^2$,

$$(2.9)\qquad \int_0^\infty \frac{r^{M+2k-1}}{2^{M/2}\sigma^{M+2k}}e^{-r^2/2\sigma^2}\,dr = 2^{k-1}\Gamma\left(\frac{M}{2}+k\right),$$

and the result follows. □

The zonal polynomials C_κ in (2.6) depend only on the latent roots of A^*. The order l of any partition κ for which C_κ is nonzero is at most the number of nonzero roots, which is $p = \operatorname{rank}\mu$ or fewer. We consider the cases $p = 0$, 1, 2 and 3 in detail in Sections 3–5. Note that when $p = 0$ or $p = 1$, but not otherwise, (2.6) simplifies to a multiple of a hypergeometric function, ${}_1F_1$, because $\Gamma(M/2 + k)$ is then the appropriate generalized hypergeometric coefficient.

2.1. *Distributions excluding reflections.* When $p < K$ the results (2.2) and (2.6) extend easily to size-and-shape and shape densities, *excluding* reflection. If $n < K$ there is nothing to do. If $n = K$ and $p < K$, choose $H_0 \in SO(K)$ such that $\mu^t T H_0$ has Kth column 0. For any $H \in SO(K)$, write $H^t = (H^*|\mathbf{h})$, and define $H'^t = (H^*| - \mathbf{h})$ where H^* is $K \times (K-1)$. Then using unimodularity of $SO(K)$ and of $O(K)$,

$$\begin{aligned}&\int_{SO(K)}\operatorname{etr}(\mu^t TH)(H\,dH^t)\\ (2.10)\qquad &\quad= \int_{SO(K)}\operatorname{etr}(\mu^t TH_0H)(H\,dH^t)\\ &\quad= \tfrac{1}{2}\int_{SO(K)}\operatorname{etr}(\mu^t TH_0H)(H\,dH^t) + \tfrac{1}{2}\int_{SO(K)}\operatorname{etr}(\mu^t TH_0H')(H'\,dH'^t)\\ &\quad= \tfrac{1}{2}\int_{O(K)}\operatorname{etr}(\mu^t TH)(H\,dH^t).\end{aligned}$$

We obtain the following corollaries to Theorems 1 and 2.

COROLLARY 1. *When $N - 1 \geq K$ and $p < K$, the size-and-shape density for $T = T^{NR}$* excluding *reflections is (2.2) divided by 2, where $t_{ii} \geq 0$ for $i = 1, \dots, K-1$ and t_{KK} is unrestricted. When $N - 1 < K$ (2.2) stands, since t_{KK} is not present.*

COROLLARY 2. *When $N - 1 \geq K$ and $p < K$, the density for $W = W^{NR}$* excluding *reflections is (2.6) divided by 2, and w_{KK} is unrestricted. When $N - 1 < K$ the shape density is (2.6), since w_{KK} is not present.*

Note that these results also hold when rank $\mu = K$ and rank $T < K$, an event with probability zero.

When $p = K$, (2.10) is not true and we cannot divide the density with reflection by 2. For example, suppose $p = K = 2$ and A has eigenvalues $\lambda_1 \geq \lambda_2$, and, without loss of generality, $\mu^t T/\sigma^2$ is diagonal with entries $\sqrt{\lambda_1}$ and $\sqrt{\lambda_2}$. Then by (2.4), and writing $H = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ and $(H\,dH^t) = \theta$, we obtain

$$
\begin{aligned}
4\pi\, {}_0F_1(1; \lambda_1, \lambda_2) &= \int_{O(2)} \operatorname{etr}\left(\frac{\mu^t T H}{\sigma^2}\right)(H\,dH^t) \\
&= \int_0^{2\pi} \exp\left\{2\left(\sqrt{\lambda_1} + \sqrt{\lambda_2}\right)\cos\theta\right\} d\theta \\
&\quad + \int_0^{2\pi} \exp\left\{2\left(\sqrt{\lambda_1} - \sqrt{\lambda_2}\right)\cos\theta\right\} d\theta \qquad (2.11)\\
&= 2\pi I_0\left(2\left(\sqrt{\lambda_1} + \sqrt{\lambda_2}\right)\right) + 2\pi I_0\left(2\left(\sqrt{\lambda_1} - \sqrt{\lambda_2}\right)\right) \\
&= 2\pi\, {}_0F_1\left(1; \left(\sqrt{\lambda_1} + \sqrt{\lambda_2}\right)^2\right) + 2\pi\, {}_0F_1\left(1; \left(\sqrt{\lambda_1} - \sqrt{\lambda_2}\right)^2\right).
\end{aligned}
$$

The two modified Bessel functions $I_0(\cdot)$ are equal only when $\lambda_2 = 0$. Each individually is a term in a size-and-shape density excluding reflection [Goodall and Mardia (1991a) and Section 4]. One is obtained from the other by reflecting the figure X (or, equivalently, μ_X) in any axis. An analogous result is true when $K \geq 3$.

3. Coincident means. When the landmark means coincide, then $p = 0$, the central case.

COROLLARY 3. *When $\mu = 0$ the reflection size-and-shape density is*

$$
\frac{\prod_{i=1}^{n} t_{ii}^{K-i}}{\sigma^M 2^{(M-2n)/2} \pi^{(M-nK)/2} \Gamma_n(K/2)} \operatorname{etr}\left\{-\frac{T^t T}{2\sigma^2}\right\}. \qquad (3.1)
$$

This is the central Bartlett decomposition. The elements of T are independent,

$$(3.2)\qquad \begin{aligned} t_{ii}^2 &\sim \sigma^2\chi_{k-i+1}^2, & i &= 1,\dots,n, \\ t_{ij} &\sim N(0,\sigma^2), & 1 < i &\le N-1, \quad 1 \le j \le n, \quad i > j, \end{aligned}$$

for all choices of $N-1$ and K. On integrating (3.1) using (2.9), or from (2.6), we obtain the following:

COROLLARY 4. *When $\mu = 0$, the reflection shape density is*

$$(3.3)\qquad \frac{\prod^n w_{ii}^{K-i} J(\mathbf{u}) 2^{n-1}\Gamma(M/2)}{\pi^{(M-nK)/2}\Gamma_n(K/2)}.$$

When $\mu = 0$ the size-and-shape density (3.1) factorizes into the shape density (3.2) and the size density, $r^2 \sim \sigma^2\chi_M^2(\xi^2)$. Size and shape are independent. [When $\mu \neq 0$ the noncentral term $\exp(-\xi^2/2\sigma^2)\ {}_0F_1(K/2; A)$ does not factorize.] When $K = n = 2$, (3.1) and (3.3) become

$$(3.4)\qquad \frac{2t_{11}e^{-r^2/2\sigma^2}}{\sigma^2(2\pi\sigma^2)^{N-2}} \quad\text{and}\quad \frac{2w_{11}J(\mathbf{u})(N-2)!}{\pi^{N-2}},$$

respectively. For triangles, $N = 3$, we substitute $w_{11} = \cos\rho$ and $J(\mathbf{u}) = \sin\rho$ to get

$$(3.5)\qquad \frac{\sin 2\rho}{2\pi}\frac{r^3}{\sigma^4}e^{-r^2/2\sigma^2} \quad\text{and}\quad \frac{\sin 2\rho}{2\pi}.$$

The latter is uniform measure on $S^2(1/2)$. For any N, the shape density with $K = 2$ is twice Kendall's (1984) invariant measure on shape space (excluding reflections) $\Sigma_2^N \simeq CP^{N-2}(4)$. See also Le (1990) and W. S. Kendall's (1988) "generalized D. G. Kendall theorem" which gives the shape density for triangles in $K \geq 3$ dimensions when $p = 0$.

4. Collinear means. We now consider the case of collinear means, so that $p = 1$. For general μ, let $\eta_1, \dots, \eta_n$ be the singular values of $\mu^t T$. Including reflection, the Riemannian, or Procrustes, distance between μ and T is

$$(4.1)\qquad \min_{H\in O(K)} \|\mu - TH\|^2 = \xi^2 + r^2 - 2\sum_{i=1}^{n}\eta_i = \xi^2 + r^2 - 2\xi r\cos\rho.$$

With probability 1, exactly the first p singular values are nonzero. When $p = 1$ we choose μ such that $\mu_{11} = \xi$ and $\mu_{ij} = 0$ otherwise; then

$$(4.2)\qquad \eta_1 = \xi t_{11} = \xi r\cos\rho.$$

Substituting $\eta_1^2/4\sigma^4$ for A, the expression ${}_0F_1(K/2; A)$ in (2.2) is a scalar hypergeometric function of t_{11}, and therefore of r and ρ. However the reflection size-and-shape density is a function of r and ρ only when $K \leq 2$, and not when $K \geq 3$ — because terms in t_{ii}, $i \geq 2$, appear in the Jacobian.

COROLLARY 5. *When $\mu_{11} = \xi$ and $\mu_{ij} = 0$, $ij \neq 11$, the reflection size-and-shape density is*

$$(4.3)\qquad \frac{\prod_{i=1}^{n} t_{ii}^{K-i}}{\sigma^M 2^{(M-2n)/2} \pi^{(M-nK)/2} \Gamma_n(K/2)} e^{-(r^2+\xi^2)/2\sigma^2} {}_0F_1\left(\frac{K}{2}; \frac{\xi^2 t_{11}^2}{4\sigma^4}\right).$$

This is the noncentral Bartlett decomposition, generalizing Muirhead (1982) and Farrell (1985) to arbitrary N and K. The density factorizes as in (3.2) with the modification

$$(4.4)\qquad t_{11}^2 \sim \sigma^2 \chi_K^2(\xi^2/\sigma^2).$$

Goodall and Mardia (1992) show that the shape density follows from (3.2) and (4.4) using the fact that the marginal distribution of $\cot^2 \rho = t_{11}^2 / \Sigma\Sigma_{ij \neq 11} t_{ij}^2$ is noncentral F. Alternatively, from (2.6), we have the following:

COROLLARY 6. *When $p = 1$ and $\|\mu\| = 1$, the reflection shape density has the ${}_1F_1$ confluent hypergeometric form*

$$(4.5)\qquad \frac{\prod^n w_{ii}^{K-i} J(\mathbf{u}) 2^{n-1} \Gamma(M/2)}{\pi^{(M-nK)/2} \Gamma_n(K/2)} \times \exp\left\{-\frac{1}{2\sigma^2} + \frac{\cos^2 \rho}{2\sigma^2}\right\} {}_1F_1\left(\frac{K-M}{2}; \frac{K}{2}; -\frac{\cos^2 \rho}{2\sigma^2}\right).$$

PROOF. When $p = 1$, $C_\kappa(A^*) = (\cos \rho / 2\sigma)^{2k}$ when $\kappa = (k)$ and $C_\kappa(A^*) = 0$ otherwise. Thus (2.6) becomes

$$(4.6)\qquad \frac{\prod^n w_{ii}^{K-i} J(\mathbf{u}) 2^{n-1} \Gamma(M/2)}{\pi^{(M-nK)/2} \Gamma_n(K/2)} \exp\left\{-\frac{1}{2\sigma^2}\right\} \sum_{k=0}^{\infty} \frac{(M/2)_k}{(K/2)_k k!} \left(\frac{\cos^2 \rho}{2\sigma^2}\right)^k,$$

and (4.5) follows from the Kummer relation. □

With $p = 1$, the shape density is a function solely of ρ only when $K \leq 2$. The series ${}_1F_1$ is finite when $(K - M)/2 = K(2 - N)/2$ is a negative integer, in which case either K or N must be even. When $K = n = 2$, (4.5) divided by 2 simplifies to

$$(4.7)\qquad \frac{w_{11} J(\mathbf{u})(N-2)!}{\pi^{N-2}} \exp\left\{\frac{\cos^2 \rho - 1}{2\sigma^2}\right\} {}_1F_1\left(2 - N; 1; \frac{\cos^2 \rho}{2\sigma^2}\right),$$

which is the Mardia and Dryden (1989a) shape density. Goodall and Mardia (1991a) show the equality of (4.7) and Dryden and Mardia [(1991), equation (3.3)].

5. Higher rank cases. We now first consider distributions when $p = 2$. Subsequently we comment on the cases when $p > 2$. The zonal polynomial expansion of ${}_0F_1$ is an infinite series of symmetric functions in the latent roots of the matrix argument A. When rank $A = 1$, this is a single series. When rank $A \geq 2$, the zonal polynomials can be computed using James (1968), but the summation over partitions, Σ_κ is nontrivial. We know of no simplification of $\Sigma_\kappa C_\kappa(A)/(K/2)_\kappa$ that may be used in Theorem 2. One approach is to substitute explicit formulae for the $C_\kappa(A)$ when rank $A = 2$ from James (1964). An alternative is to use Anderson and Girshick's (1944) and Anderson's (1946) expansion of the multivariate Wishart. Suppose λ_1 and λ_2 are the latent roots of A [so $\lambda_i = (\eta_i/2\sigma^2)^2$], and define $a_1 = \lambda_1 + \lambda_2$, $a_2 = \lambda_1\lambda_2$. Then comparing their formula to the noncentral Wishart density gives

$$(5.1)\quad \begin{aligned} &\frac{{}_0F_1(K/2;A)}{\Gamma(K/2)\Gamma((K-1)/2)} \\ &\quad = 2^{(K-2)/2}\sum_{w=0}^{\infty}\frac{a_2^w a_1^{-w} a_1^{(K-2)/4} 2^{-(K-2)/2}}{w!\Gamma((K-1)/2+w)} I_{(K-2)/2+2w}\left(2\sqrt{a_1}\right). \end{aligned}$$

Expanding the Bessel function $I_\cdot(\cdot)$ [Watson (1962)] gives the following:

COROLLARY 7. *When $p = 2$ the reflection size-and-shape density is*

$$(5.2)\quad \begin{aligned} &\frac{\prod_{i=1}^n t_{ii}^{K-i}}{\sigma^M 2^{(M-2n)/2}\pi^{(M-nK)/2}\Gamma_n(K/2)} \operatorname{etr}\left\{-\frac{T^tT}{2\sigma^2}-\frac{\mu^t\mu}{2\sigma^2}\right\} \\ &\quad \times \sum_{\ell=0}^{\infty}\sum_{w=0}^{\infty}\frac{a_1^\ell a_1^w \Gamma(K/2)\Gamma(K/2-1/2)}{\ell!w!\Gamma(K/2+w-1/2)\Gamma(K/2+\ell+2w)}. \end{aligned}$$

An equivalent expression was obtained by Kshirsagar [(1963), equation (4.15)]. Anderson and Girshick (1944) also give the Wishart density when rank $A = 1$. Analogous calculations confirm (4.3) with the usual (scalar) hypergeometric series expansion. Furthermore, (5.2) simplifies to (4.3) when $\lambda_2 = 0$.

To derive the rank 2 shape density we follow the proof of Theorem 2. We write $\lambda_1^* = \lambda_1\sigma^2/r^2$ and $\lambda_2^* = \lambda_2\sigma^2/r^2$ for the roots of A^*, substitute $a_1^* = \lambda_1^* + \lambda_2^*$ and $a_2^* = \lambda_1^*\lambda_2^*$ into (5.2), and integrate out r, with $k = \ell + 2w$ in (2.9).

COROLLARY 8. *When $p = 2$ the reflection shape density is*

$$(5.3)\quad \begin{aligned} &\frac{\prod w_{ii}^{K-i} J(\mathbf{u})}{\pi^{(M-nK)/2}\Gamma_n(K/2)} \operatorname{etr}\left\{-\frac{\mu^t\mu}{2\sigma^2}\right\} \\ &\quad \times \sum_{\ell=0}^{\infty}\sum_{w=0}^{\infty}\frac{a_1^{*\ell} a_2^{*w} 2^{n+\ell+2w-1}\Gamma(M/2+\ell+2w)}{\ell!w!((K-1)/2)_w (K/2)_{\ell+2w}}. \end{aligned}$$

As a check (5.3) simplifies to (4.6) when $\lambda_2^* = 0$.

There are a number of simplifications when $K = p = 2$, using (2.11) and the identity

$$\frac{r^2\xi^2\cos^2\rho}{4\sigma^4} = \left(\sqrt{\lambda_1} + \sqrt{\lambda_2}\right)^2 = a_1 + 2\sqrt{a_2}. \tag{5.4}$$

Define $\bar{\rho}$ to be the Riemannian distance between μ_X and the reflection of X. Thus

$$\frac{r^2\xi^2\cos^2\bar{\rho}}{4\sigma^4} = \left(\sqrt{\lambda_1} - \sqrt{\lambda_2}\right)^2 = a_1 - 2\sqrt{a_2}. \tag{5.5}$$

The hypergeometric term in (2.2) is written, using (2.11),

$${}_0F_1(1;\lambda_1,\lambda_2) = \tfrac{1}{2}\,{}_0F_1\left(1;a_1 + 2\sqrt{a_2}\right) + \tfrac{1}{2}\,{}_0F_1\left(1;a_1 - 2\sqrt{a_2}\right). \tag{5.6}$$

Then, with the substitution ${}_0F_1(1;x) = \Sigma_0^\infty x^k/(k!)^2$ and Legendres's duplication formula, both the expressions (2.2) and (5.2) for the reflection size-and-shape density equal

$$\frac{t_{11}}{\sigma^{2N-2}2^{N-3}\pi^{N-2}}e^{-(r^2+\xi^2)/2\sigma^2}\sum_{\ell=0}^{\infty}\sum_{w=0}^{\infty}\frac{4^w a_1^\ell a_2^\ell}{\ell!(2w)!(\ell+2w)!}. \tag{5.7}$$

Further, each scalar hypergeometric function in (5.6) can be integrated with respect to r to give the reflection shape density, containing the term

$$\begin{aligned}&\frac{1}{2}\exp\left\{\frac{\cos^2\rho}{2\sigma^2}\right\}{}_1F_1\left(2-N;1;-\frac{\cos^2\rho}{2\sigma^2}\right)\\ &\quad+\frac{1}{2}\exp\left\{\frac{\cos^2\bar{\rho}}{2\sigma^2}\right\}{}_1F_1\left(2-N;1;-\frac{\cos^2\bar{\rho}}{2\sigma^2}\right).\end{aligned} \tag{5.8}$$

Thus when $K = p = 2$, the reflection size-and-shape density and the reflection shape density is each the average of two rank 1 densities, one a function of ρ, the other a function of $\bar{\rho}$, given by (4.3) and (4.5), respectively. The corresponding size-and-shape and shape densities excluding reflection involve only the left-hand terms in (5.6) and (5.8); they are functions of ρ equal to the rank 1 densities. This illustrates homogeneity of shape spaces with $K = 2$.

When $n = K > 2$ and $p = 2$, the reflection densities are still functions of ρ and $\bar{\rho}$, and the densities excluding reflection are half of the reflection density, by (2.10).

For the case $p = 3$, we may use a triple-summation expansion of ${}_0F_1$ in terms of the roots of A, given by James (1955). The details, leading to corollaries analogous to Corollary 7 and Corollary 8, are not illuminating. Goodall and Mardia (1992) demonstrate the integration of $\operatorname{etr}(\mu^t TH/\sigma^2)$ over $H \in SO(3)$. The details are messy.

6. Configuration densities. In the rank 1 case, integrating size r out of the type ${}_0F_1$ size-and-shape density gives the type ${}_1F_1$ shape density. This follows from the inductive definition of the hypergeometric functions via the Laplace transform [Herz (1955)]. In the matrix setting, the Laplace transform involves integration over the positive definite matrices. Thus, the reflection shape density is not type ${}_1F_1$. Instead, the multivariate approach leads to distributions of type ${}_1F_1$ on equivalence classes of figures modulo affine transformations. [The next step would be to type ${}_2F_1$ distributions of canonical correlations, James (1964).] An affine transformation is specified by the pair (e: $K \times 1$, E: $K \times K$) where e is the translation and E is nonsingular. Then, we have the following:

DEFINITION. Two figures X: $N \times K$ and X': $N \times K$ have the same *configuration*, or *affine shape*, if $X' = XE + 1_N e^t$ for some (e, E).

There is a single configuration for figures comprising $N \le K + 1$ landmarks. We assume $N > K + 1$. Then, analogous to (1.2), we write

$$LX = Y = UE. \tag{6.1}$$

The matrix U: $N - 1 \times K$ contains configuration coordinates analogous to the size-and-shape matrix T. Let $Y = (Y_1^t | Y_2^t)^t$ be the partition into Y_1: $K \times K$ and Y_2: $q \times K$, where $q = N - K - 1 \ge 1$. With probability 1, Y_1 is nonsingular, and we choose $E = Y_1^{-1}$ and $U^t = (I | V^t)^t$ with $V = Y_2 Y_1^{-1}$. For completeness, we must also define configuration coordinates when Y_1 is singular. If rank $Y = m \le K$, then, if we can choose L, we replace I by a diagonal matrix with the first m diagonal elements 1 and the rest 0. But in general L is fixed once and for all, in which case U is "special" lower-triangular, defined by: For each j, $1 \le j \le K$, let $j^* = \arg\min_i u_{ij} \ne 0$; then $u_{j^*j} = 1$ and $u_{j^*k} = 0$ for $k \ne j$. That is, each row of Y is either a linear combination of the previous rows, or it is transformed to a (new) standard basis vector.

THEOREM 3. *With the model* (1.1) *and the notation above, the configuration density is*

$$\frac{\Gamma_K((N-1)/2)}{\pi^{Kq/2}|I + V^tV|^{q/2}\Gamma_K(K/2)} \operatorname{etr}\left\{-\frac{\Omega}{2\sigma^2} - \frac{\mu^t\mu}{2\sigma^2}\right\} {}_1F_1\left(-\frac{q}{2}; \frac{k}{2}; -\frac{\Omega}{2}\right), \tag{6.2}$$

where $\Omega = \mu^t U(U^tU)^{-1}U^t\mu = \mu^t Y(Y^tY)^{-1}Y^t\mu$, *for all* Y *for which rank* $Y_1 = K$. *For* q *a positive integer, the expression* ${}_1F_1$ *in* (6.2) *is a polynomial with degree* $Kq/2$ *in the latent roots of* Ω.

PROOF. Make the substitution $Y = UF^{1/2}H$ where $E = F^{1/2}H$, $H \in O(K)$, and $(F^{1/2})^2 = F > 0$. The Jacobian is $(dY) = |F|^{(q-1)/2}(dV)(dF)(H\,dH^t)$. We integrate out first H as in Theorem 1 and then F. The second integral is the matrix Laplace transform [Muirhead (1982), Theorem 7.3.4]. □

The configuration density is more familiar as a noncentral multivariate beta distribution [Olkin and Rubin (1964)], when $q > K - 1$ (that is, $N > 2K$), with the appropriate change of variables. Also, we define *rescaled* configuration coordinates by replacing the submatrix I of U by the singular values of μ, and scaling each column of V to match. For small σ (Section 9) these coordinates involve only an orthogonal transformation of Y.

7. Distribution of latent roots. The joint distribution of the latent roots of Y is of considerable interest in the statistics of size-and-shape, and of shape. The latent roots are the invariants with respect to the isometries of size-and-shape space and of shape space [Small (1983), Kendall (1984), Goodall (1991) and Le and Kendall (1993)] induced by multiplication of Y on the left by an orthogonal matrix $O(N-1)$. If reflection is included, then left multiplication by $SO(N-1)$ suffices, but in any case the latent roots ℓ_i, $i = 1, \dots, n$ are (1) ordered $\ell_1 \geq \ell_2 \geq \cdots \geq \ell_n$ and (2) nonnegative, $\ell_n \geq 0$, and the roots have the same distribution in the reflection and nonreflection cases. The derivation of the joint density of the ℓ_i^2 in the central case [Muirhead (1982), Corollary 9.4.2] extends easily to the *size-and-shape cone density*, or joint density of the ℓ_i in the noncentral case. Then, parallel to our development of Theorem 2 from Theorem 1, we obtain the *shape disk density*, or joint density of the $\ell_i^* = \ell_i / r$ ($\Sigma \ell_i^{*2} = 1$). Goodall and Mardia (1991b) give full details. For example, when $\mu = 0$, $K = 2$, and $N \geq 3$, the cone density is

$$\text{(7.1)} \qquad \frac{e^{-(\ell_1^2+\ell_2^2)/2\sigma^2}\left(\ell_1\ell_2\right)^{N-3}\left(\ell_1^2-\ell_2^2\right)}{\sigma^{2N-2}(N-3)!}.$$

This factorizes into the size distribution $r^2 \sim \sigma^2\chi_M^2$ and the invariant shape disk density

$$\text{(7.2)} \qquad \frac{2(\sin 2\theta)^{N-3}\cos 2\theta}{N-3},$$

where $\ell_1^* = \cos\theta$, $\ell_2^* = \sin\theta$, $0 \leq \theta \leq \pi/4$. When $p = 1$, some simplifications are also possible.

8. General distributions. The results in Sections 2–7 for the isotropic Gaussian model extend easily to a Gaussian model with covariance $\Sigma \otimes I_K$. With this model the errors at each landmark are isotropic, and the errors are correlated between landmarks. The key step is to write

$$\text{(8.1)} \qquad \Sigma^{-1/2}Y \sim N\left(\Sigma^{-1/2}\mu, I_{N-1} \otimes I_K\right).$$

For example, the inner product $r\xi\cos\rho = \max_H \operatorname{tr}\mu^t TH$ becomes $\max_H \operatorname{tr}\mu^t\Sigma^{-1}TH$, corresponding to a weighted Procrustes distance. Shape analysis with general Σ is discussed by Dryden and Mardia (1991) and Goodall (1991).

9. Inference. We can perform inference using the densities derived above. But the procedure will be computationally intensive as seen for $K = 2$ in Mardia and Dryden (1989b). One alternative is to use the densities to derive some useful approximations which we now describe.

9.1. *Approximations.* For inference we use a small variation result, $\sigma \to 0$. Recall that the singular values of $\mu^t T$ are written $\eta_1, \dots, \eta_p$. From Khatri and Mardia (1977)

$$(9.1)\quad {}_0F_1(K/2; A) \simeq C \exp\left(\sum \eta_i/\sigma^2\right) \Big/ \left[\prod_{j=2}^{p} \prod_{i=1}^{j-1} \sqrt{\eta_i + \eta_j} \prod_{i=1}^{p} \eta_i^{(K-p)/2} \right],$$

where C is a constant. Let the singular values of μ be $\xi\beta_i$, where $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_K \geq 0$ and $\Sigma\beta_i^2 = 1$. Since $\eta_i = \xi^2\beta_i^2 + O(\sigma)$ and the denominator includes only sums and products of the η_i, we deduce that, for small σ, the contribution of ${}_0F_1$ to the joint density is a normalization constant and $\exp(\Sigma\eta_i/\sigma^2) = \exp(\xi r \cos\rho/\sigma^2)$. But when $\tilde{\mu}$ is the Procrustes superimposition of μ on T (by rotation only), that is, $\tilde{\mu}^t T$ is symmetric (Goodall 1991), then $\xi r \cos\rho = \operatorname{tr} \tilde{\mu}^t T$ and the joint density includes the exponential terms

$$(9.2)\qquad -\operatorname{tr}(T - \tilde{\mu})^t (T - \tilde{\mu})/2\sigma^2.$$

An explicit expression for $\tilde{\mu}$ is not known (we cannot write down the SVD of $\mu^t T$). However, (9.2) is invariant to orthogonal transformations of Y and μ on both left and right. A suitable choice of isometry is to assume μ is diagonal and let $\tilde{T}$ be the Procrustes superimposition of T on μ. Then $\tilde{t}_{ji} = \tilde{t}_{ij}\beta_i/\beta_j$ for $j < i \leq K$. Note that $\tilde{T} = T$ when rank $\mu \leq 1$. The Jacobian is given by

$$(9.3)\qquad (dY) = \left[\prod_{j<i} \left(\tilde{t}_{ii}\beta_i/\beta_j - \tilde{t}_{jj} \right) + O(\sigma) \right] (d\tilde{T})(dHH^t).$$

(The exact distribution of $\tilde{T}$ is more complicated than that seen in Theorems 1 and 2.) Assuming that the β_i are distinct and nonzero, the Jacobian can be shown to be $O(1)$ for small σ. We now deduce the asymptotic distribution of $\tilde{T}$ using (9.2).

THEOREM 4. *When μ is diagonal and of rank K, and σ is small, the* Procrustes shape coordinates $\tilde{t}_{ij}$, $i \geq j$, *follow independent Gaussian distributions with means μ_{ij} and variances σ^2, $i = j$, $\sigma^2/(1 + \mu_{ii}^2/\mu_{jj}^2)$, $j < i \leq K$, and σ^2, $i > K$.*

Mardia and Dryden (1989b) show that size and shape are asymptotically independent when $K = 2$. This result is true generally. We separate out a normalized size variable, $v = (r - \xi)/\sigma$. Then (9.2) becomes $-v^2/2 - w/2 + O(\sigma)$ where $w = (2\xi^2/\sigma^2)(1 - \cos\rho)$ is the normalized Procrustes sum of squares for the difference in shape between μ and Y. Sibson (1979) shows by a perturbation argument that approximately $w \sim \chi_m^2$. But w can be partitioned: Goodall (1991) extends the perturbation argument, while we give here an analogous approximation to the exact density. Removing scale imposes a constraint on the diagonal (noncentral) elements of $\tilde{T}$.

THEOREM 5. *With the setup above, including β_i distinct and nonzero, let $r^* = \sqrt{\Sigma \tilde{t}_{ii}^2}$, $u_i = \xi(\tilde{t}_{ii}/r^* - \beta_i)/\sigma$ and $a_{ij} = (\tilde{t}_{ij}/\sigma)\sqrt{1 + \beta_i^2/\beta_j^2}$, $j < i \le K$. The size-and-shape distribution factorizes, with (a) v standard Gaussian, (b) the u_i $(K-1)$-variate standard Gaussian in the tangent space to the positive orthant of S^{K-1} at the β_i, and (c) the a_{ij} and the $\tilde{t}_{ij}/\sigma$, $i > K$, independent standard Gaussian. Since rank $\mu = K$, these distributional results are the same with and without reflection.*

9.2. *Example: the geometry of chairs.* We consider a hypothetical situation in manufacturing, but the same methodology applies in other complicated examples, for example, in morphometrics and in face recognition with given landmarks. A manufacturer of chairs takes measurements in $\mathfrak{R}^3$ at 10 positions on each chair. In the ideal mean chair, μ, Figure 1, the edges have length 1 and are at 90° to one another. Figure 1 also includes the principal axes of μ (singular values, $\xi\beta_1 = 2.449$, $\xi\beta_2 = 1.581$, $\xi\beta_3 = 1.414$), together with an affine deformed chair with $\tilde{T} = \mu$ except $\tilde{t}_{33} = 1.5\xi\beta_3$. Three sets of data are analyzed, with $\sigma = 0.02$ in all three. [Note that we have $(\tilde{t}_{33} - \xi\beta_3)/\sigma = 35$ in Figure 1.]

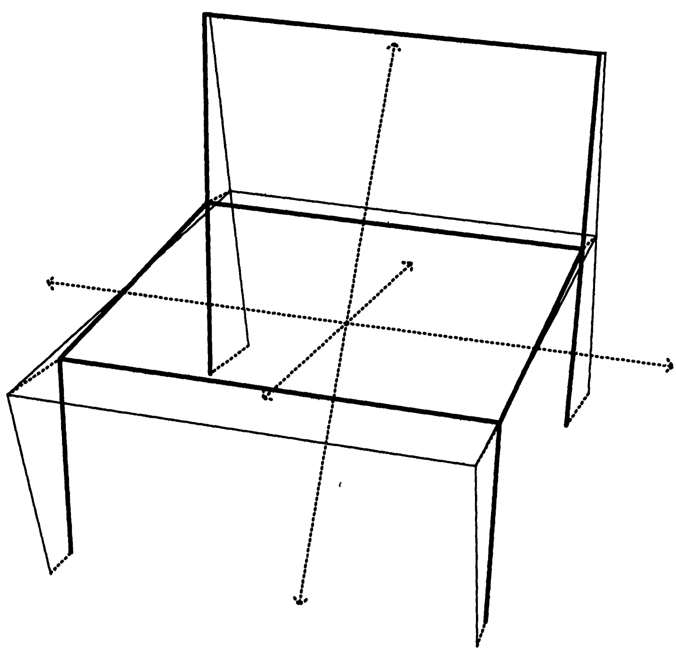

FIG. 1. *Target chair μ (bold lines) with each edge length 1 and all angles right angles. Principal axes of μ (arrows) and affine deformation (normal lines) of μ multiplying $\xi\beta_3$ by 1.5.*

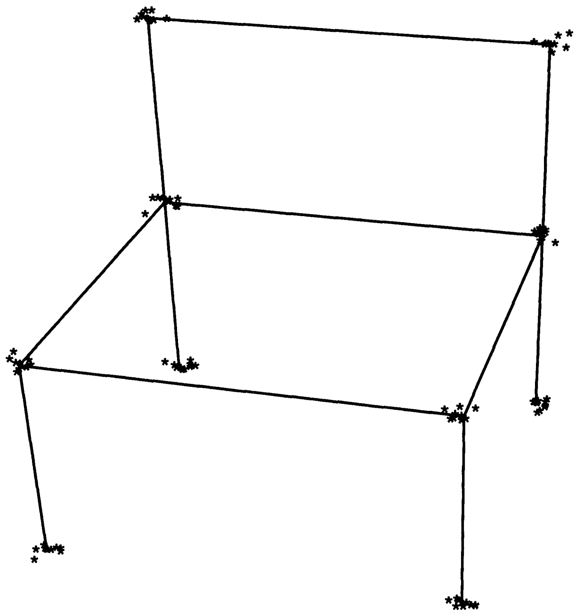

FIG. 2. *Sheared chair* μ' *with back tilted forward* 3°, *and* 10 *simulations with mean* μ' *and* $\sigma = 0.02$.

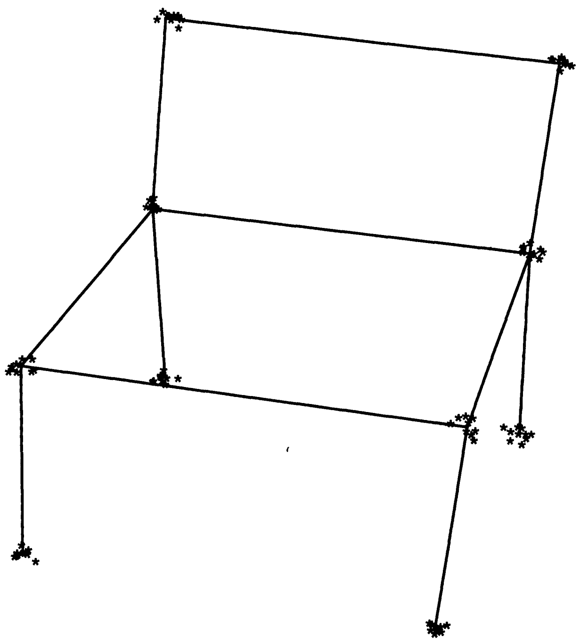

FIG. 3. *Deformed chair* μ'' *with seat rotated* 3°, *and* 10 *simulations with mean* μ'' *and* $\sigma = 0.02$.

1. $n = 100$ and mean μ.
2. $n = 10$ and mean μ' obtained by a shear of μ parallel to the horizontal plane rotating the back forwards 3° (Figure 2).
3. $n = 10$ and mean μ'' obtained by rotating only the seat of μ by 3° in the horizontal plane (Figure 3).

The objectives are in 1 to estimate σ and assess the approximations in Theorem 4, and in 2 and 3 to test for and characterize the bias specifically in shape (not size-and-shape) assuming μ' and μ'' are unknown.

SET 1. The analysis of size and shape is invariant to separate rigid body motions of the individual figures. Thus to estimate σ we use generalized Procrustes superimposition of the 100 figures, yielding a Procrustes (residual) sum of squares of 0.92483 on $(100 - 1) \times (30 - 3 - 3)$ df. Hence the mle of σ is $\hat{\sigma} = 0.02015$, which is very precise, and in computing the standardized

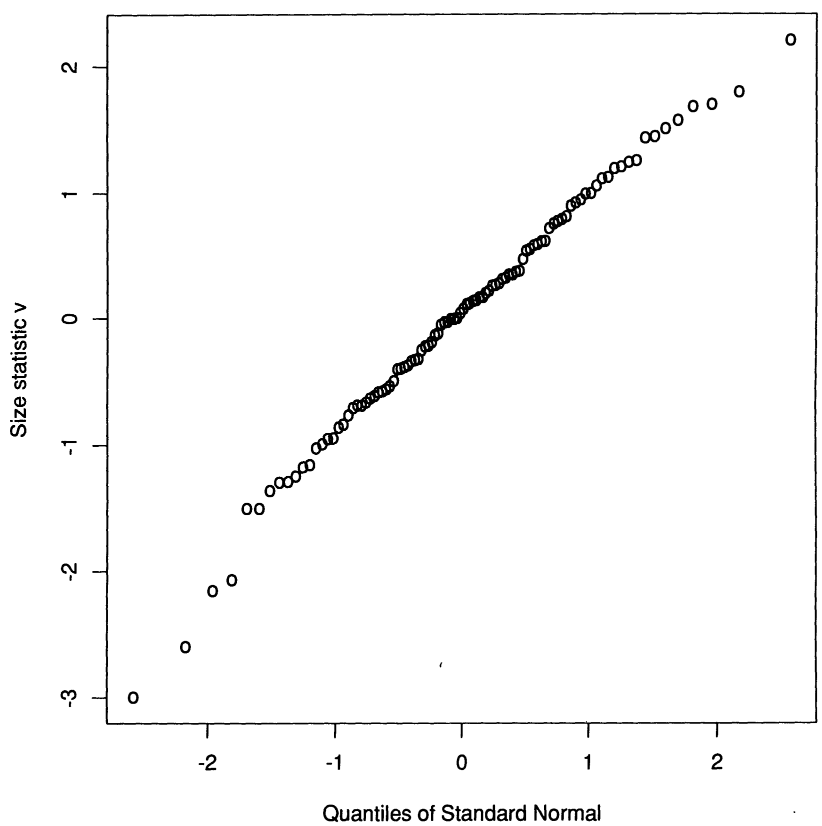

FIG. 4. *Normal probability plot of scale variable v from* 100 *simulations with mean μ and* $\sigma = 0.02$.

variates we substitute the correct σ. The normal probability plot of v, Figure 4, helps validate the Gaussian approximation.

SET 2. When $K = 3$, the affine group has dimension 12 and the Euclidean group dimension 6. The six additional affine dimensions correspond to the variables $\tilde{t}_{ij}$, $j \le i \le 3$, that is, v, the u_i and the a_{ij}. Table 1 shows the standardized size-and-shape coordinates and μ and μ', the mean coordinates $\hat{\mu}'$ of the 10 figures, and any P-values less than 10% in using $\hat{\mu}'$ to test that the mean is μ. The shear is seen only in u_1, u_3 and a_{31}, and the size change v, with a secondary "squashing" of the chair ($u_2 > 0$).

TABLE 1
Analysis of deformed chairs

	Target	**Sheared chair**			**Skewed chair**		
	μ	μ'	$\hat{\mu}'$	***P*-value**	μ''	$\hat{\mu}''$	***P*-value**
v	0.0	−0.9	−0.6	5.7	0.0	0.6	6.8
u_1	0.0	−1.1	−0.9	0.3	0.0	0.0	
u_2	0.0	0.4	0.0		0.0	−0.4	
u_3	0.0	1.5	1.5	0.0	0.0	0.4	
Σu_i^2	0.0	3.6	5.3	0.0	0.0	1.8	
a_{21}	0.0	0.0	0.1		−0.3	−0.5	
a_{31}	0.0	−2.6	−2.7	0.0	0.0	0.1	
a_{32}	0.0	0.0	−0.4		0.3	0.0	
Σa_{ij}^2	0.0	7.0	10.5	0.0	0.1	2.1	
$\tilde{t}_{41}$	0.0	0.0	0.1		0.0	−0.2	
$\tilde{t}_{42}$	0.0	0.0	0.1		0.4	0.1	
$\tilde{t}_{43}$	0.0	0.0	0.3		−0.3	−0.1	
$\hat{t}_{51}$	0.0	0.0	−0.1		0.1	0.4	
$\hat{t}_{52}$	0.0	0.0	0.5	8.4	0.0	0.0	
$\hat{t}_{53}$	0.0	0.0	0.4		−0.4	1.5	0.0
$\hat{t}_{61}$	0.0	0.0	0.2		−0.3	0.0	
$\tilde{t}_{62}$	0.0	0.0	−0.1		0.1	1.1	0.0
$\tilde{t}_{63}$	0.0	0.0	−0.2		1.2	−0.6	7.5
$\tilde{t}_{71}$	0.0	0.0	−0.2		1.4	0.7	2.8
$\tilde{t}_{72}$	0.0	0.0	0.3		0.0	1.4	0.0
$\tilde{t}_{73}$	0.0	0.0	0.2		−0.4	0.0	
$\tilde{t}_{81}$	0.0	0.0	−0.2		−0.1	0.2	
$\tilde{t}_{82}$	0.0	0.0	0.2		1.1	0.3	
$\tilde{t}_{83}$	0.0	0.0	−0.2		−0.9	0.1	
$\tilde{t}_{91}$	0.0	0.0	0.2		0.4	−0.2	
$\tilde{t}_{92}$	0.0	0.0	0.1		0.0	−0.1	
$\tilde{t}_{93}$	0.0	0.0	0.1		−1.2	−1.3	0.0
$\Sigma \tilde{t}_{ij}^2$	0.0	0.0	16.3		7.8	20.8	7.3

SET 3. For these data there is no affine component, and the rotation is seen primarily in the coordinates $\tilde{t}_{53}$, $\tilde{t}_{62}$, $\tilde{t}_{63}$, $\tilde{t}_{72}$, and $\tilde{t}_{93}$. (Table 1). [Note that because the u_i and a_{ij} are small, the coordinates $\tilde{t}_{ij}$ $(i > 3)$ are almost the rescaled configuration coordinates of Section 6.] Investigating the effect of changing the rotation angle, θ say, in μ'', the $\tilde{t}_{ij}/\theta$, $i > 3$, a_{21}/θ, and a_{32}/θ are almost constant, and a_{31} and the u_i vary quadratically with θ (with $u_2 \approx u_3$). Similar constraints are found when μ' is varied. In general the power of our tests is increased when, through a coordinate transformation of $\tilde{T}$, the difference in size-and-shape or shape between $\mu(\theta)$ and μ appears in a minimal number of coordinates. There is some flexibility because the isometry is only partly specified by μ.

Acknowledgments. The authors wish to thank Ian Dryden, Yasunori Fujikoshi and two anonymous referees for their comments. The comments of a referee on the proof of Theorem 3 are especially acknowledged.

REFERENCES

ANDERSON, T. W. (1946). The non-central Wishart distribution and certain problems of multivariate statistics. *Ann. Math. Statist.* **17** 409–431.

ANDERSON, T. W. and GIRSHICK, M. A. (1944). Some extensions of the Wishart distribution. *Ann. Math. Statist.* **15** 345–357.

BOOKSTEIN, F. L. (1986). Size and shape spaces for landmark data in two dimensions (with discussion). *Statist. Sci.* **1** 181–242.

CARNE, T. K. (1990). The geometry of shape spaces. *Proc. London Math. Soc.* **61** 407–432.

DRYDEN, I. L. and MARDIA, K. V. (1991). General shape distributions in a plane. *Adv. in Appl. Probab.* **23** 259–276.

FARRELL, R. H. (1985). *Multivariate Calculations: Use of the Continuous Groups*. Springer, New York.

GOODALL, C. R. (1991). Procrustes methods in the statistical analysis of shape (with discussion). *J. Roy. Statist. Soc. Ser. B* **53** 285–339.

GOODALL, C. R. and MARDIA, K. V. (1991a). A geometrical derivation of the shape density. *Adv. in Appl. Probab.* **23** 496–514.

GOODALL, C. R. and MARDIA, K. V. (1991b). Multivariate aspects of shape theory. Technical Report 91-20, Center for Multivariate Analysis, Pennsylvania State Univ.

GOODALL, C. R. and MARDIA, K. V. (1992). The noncentral Bartlett decompositions and shape densities. *J. Multivariate Anal.* **40** 94–108.

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

JAMES, A. T. (1955). A generating function for averages over the orthogonal group. *Proc. Royal Soc. London Ser. A* **229** 367–375.

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

JAMES, A. T. (1968). Calculation and zonal polynomial coefficients by use of the Laplace–Beltrami operator. *Ann. Math. Statist.* **39** 1711–1718.

KENDALL, D. G. (1984). Shape manifolds, Procrustean metrics, and complex projective spaces. *Bull. London Math. Soc.* **16** 81–121.

KENDALL, D. G. (1989). A survey of the statistical theory of shape. *Statist. Sci.* **4** 87–120.

KENDALL, W. S. (1988). Symbolic computation and the diffusion of shapes of triads. *Adv. in Appl. Probab.* **20** 775–797.

KHATRI, C. G. and MARDIA, K. V. (1977). The von Mises–Fisher matrix distribution in orientation statistics. *J. Roy. Statist. Soc. Ser. B* **39** 95–106.

KSHIRSAGAR, A. M. (1963). Effect of non-centrality on the Bartlett decompositions of a Wishart matrix. *Ann. Inst. Statist. Math.* **14** 217–228.

LE, H.-L. (1990). A stochastic calculus approach to the shape distribution induced by a complex normal model. *Proc. Cambridge Philos. Soc.* **109** 221–228.

LE, H.-L. and KENDALL, D. G. (1993). The Riemannian structure of Euclidean spaces: A novel environment for statistics. *Ann. Statist.* **21**. To appear.

MARDIA, K. V. (1989a). Shape analysis of triangles through directional techniques. *J. Roy. Statist. Soc. Ser. B* **51** 449–458.

MARDIA, K. V. (1989b). Discussion of "A survey of the statistical theory of shape" by D. G. Kendall. *Statist. Sci.* **4** 108–111.

MARDIA, K. V. and DRYDEN, I. L. (1989a). Shape distributions for landmark data. *Adv. in Appl. Probab.* **21** 742–755.

MARDIA, K. V. and DRYDEN, I. L. (1989b). The statistical analysis of shape data. *Biometrika* **76** 271–281.

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

OLKIN, I. and RUBIN, H. (1964). Multivariate beta distributions and independence properties of the Wishart distributions. *Ann. Math. Statist.* **35** 261–269.

SIBSON, R. (1979). Studies in the robustness of multidimensional scaling: perturbational analysis of classical scaling. *J. Roy. Statist. Soc. Ser. B* **41** 217–229.

SMALL, C. G. (1981). Distributions of shape and maximal invariant statistics. Ph.D. dissertation, Univ. Cambridge.

SMALL, C. G. (1983). Characterization of distributions from maximal invariant statistics. *Z. Wahrsch. Verw. Gebiete* **63** 517–527.

WATSON, G. N. (1962). *A Treatise on the Theory of Bessel Functions*, 2nd ed. Cambridge Univ. Press.

DEPARTMENT OF STATISTICS
CLASSROOM BUILDING
PENNSYLVANIA STATE UNIVERSITY
UNIVERSITY PARK, PENNSYLVANIA 16802

DEPARTMENT OF STATISTICS
UNIVERSITY OF LEEDS
LEEDS LS2 9JT
UNITED KINGDOM