

Statistics on Manifolds with Applications to Shape Spaces

Rabi Bhattacharya and Abhishek Bhattacharya

ABSTRACT. This article provides an exposition of recent developments on the analysis of *landmark based shapes* in which a k -ad, i.e., a set of k points or landmarks on an object or a scene, are observed in 2D or 3D, for purposes of identification, discrimination, or diagnostics. Depending on the way the data are collected or recorded, the appropriate shape of an object is the maximal invariant specified by the space of orbits under a group G of transformations. All these spaces are manifolds, often with natural Riemannian structures. The statistical analysis based on Riemannian structures is said to be *intrinsic*. In other cases, proper distances are sought via an *equivariant embedding* of the manifold M in a vector space E , and the corresponding statistical analysis is called *extrinsic*.

1. Introduction

Statistical analysis of a probability measure Q on a differentiable manifold M has diverse applications in directional and axial statistics, morphometrics, medical diagnostics and machine vision. In this article, we are mostly concerned with the analysis of landmark based data, in which each observation consists of $k > m$ points in m -dimension, representing k locations on an object, called a k -ad. The choice of landmarks is generally made with expert help in the particular field of application. The objects of study can be anything for which two k -ads are equivalent modulo a group of transformations appropriate for the particular problem depending on the method of recording of the observations. For example, one may look at k -ads modulo size and Euclidean rigid body motions of translation and rotation. The analysis of shapes under this invariance was pioneered by Kendall (1977, 1984) and Bookstein (1978). Bookstein's approach is primarily registration-based requiring two or three landmarks to be brought into a standard position by translation, rotation and scaling of the k -ad. For these shapes, we would prefer Kendall's more invariant view of a shape identified with the orbit under rotation (in m -dimension) of the k -ad centered at the origin and scaled to have unit size. The resulting shape space is denoted Σ_k^m . A fairly comprehensive account of parametric inference on these manifolds, with many references to the literature, may be found in Dryden and Mardia (1998). The nonparametric methodology pursued here, along with the geometric and other mathematical issues that accompany it, stems from the earlier

Key words and phrases. shape space of k -ads, Fréchet mean, extrinsic and intrinsic means, nonparametric analysis.

Research supported in part by NSF Grant DMS 0406143.

work of Bhattacharya and Patrangenaru(2002, 2003, 2005).

Recently there has been much emphasis on the statistical analysis of other notions of shapes of k -ads, namely, *affine shapes* invariant under affine transformations, and *projective shapes* invariant under projective transformations. Reconstruction of a scene from two (or more) aerial photographs taken from a plane is one of the research problems in affine shape analysis. Potential applications of projective shape analysis include face recognition and robotics-for robots to visually recognize a scene. (Mardia and Patrangenaru (2005), Bandulasiri et al.(2007)).

Examples of analysis with real data suggest that appropriate nonparametric methods are more powerful than their parametric counterparts in the literature, for distributions that occur in applications (Bhattacharya and Bhattacharya (2008a)).

There is a large literature on registration via landmarks in functional data analysis (see, e.g., Bigot (2006), Xia and Liu (2004), Ramsay and Silverman (2005)), in which proper alignments of curves are necessary for purposes of statistical analysis. However this subject is not closely related to the topics considered in the present article.

The article is organized as follows. Section 2 provides a brief expository description of the geometries of the manifolds that arise in shape analysis. Section 3 introduces the basic notion of the *Fréchet mean* as the unique minimizer of the *Fréchet function* $F(p)$, which is used here to nonparametrically discriminate different distributions. Section 4 outlines the asymptotic theory for *extrinsic mean*, namely, the unique minimizer of the Fréchet function $F(p) = \int_M \rho^2(p, x)Q(dx)$ where ρ is the distance inherited by the manifold M from an equivariant embedding J . In Section 5, we describe the corresponding asymptotic theory for *intrinsic means* on Riemannian manifolds, where ρ is the geodesic distance. In Section 6, we apply the theory of extrinsic and intrinsic analysis to some manifolds including the shape spaces of interest. Finally, Section 7 illustrates the theory with three applications to real data.

2. Geometry of Shape Manifolds

Many differentiable manifolds M naturally occur as submanifolds, or surfaces or hypersurfaces, of an Euclidean space. One example of this is the sphere $S^d = \{p \in \mathbb{R}^{d+1} : \|p\| = 1\}$. The shape spaces of interest here are not of this type. They are generally quotients of a Riemannian manifold N under the action of a transformation group. A number of them are quotient spaces of $N = S^d$ under the action of a compact group G , i.e., the elements of the space are orbits in S^d traced out by the application of G . Among important examples of this kind are axial spaces and Kendall's shape spaces. In some cases the *action of the group is free*, i.e., $gp = p$ only holds for the identity element $g = e$. Then the elements of the orbit $O_p = \{gp : g \in G\}$ are in one-one correspondence with elements of G , and one can identify the orbit with the group. The orbit inherits the differential structure of the Lie group G . The tangent space $T_p N$ at a point p may then be decomposed into a *vertical subspace* of dimension that of the group G along the orbit space to which p belongs, and a *horizontal* one which is orthogonal to it. The projection

π , $\pi(p) = O_p$ is a *Riemannian submersion* of N onto the quotient space N/G . In other words, $\langle d\pi(v), d\pi(w) \rangle_{\pi(p)} = \langle v, w \rangle_p$ for horizontal vectors $v, w \in T_p N$, where $d\pi : T_p N \rightarrow T_{\pi(p)} N/G$ denotes the differential, or Jacobian, of the projection π . With this metric tensor, N/G has the natural structure of a Riemannian manifold. The intrinsic analysis proposed for these spaces is based on this Riemannian structure (See Section 5).

Often it is simpler both mathematically and computationally to carry out an extrinsic analysis, by embedding M in some Euclidean space $E^k \approx \mathbb{R}^k$, with the distance induced from that of E^k . This is also pursued when an appropriate Riemannian structure on M is not in sight. Among the possible embeddings, one seeks out *equivariant embeddings* which preserve many of the geometric features of M .

DEFINITION 2.1. For a Lie group H acting on a manifold M , an embedding $J : M \rightarrow \mathbb{R}^k$ is *H-equivariant* if there exists a group homomorphism $\phi : H \rightarrow GL(k, \mathbb{R})$ such that

$$(2.1) \quad J(hp) = \phi(h)J(p) \quad \forall p \in M, \forall h \in H.$$

Here $GL(k, \mathbb{R})$ is the *general linear group* of all $k \times k$ non-singular matrices.

[Note: Henceforth, BP (...) stands for Bhattacharya and Patrangenaru (...) and BB (...) stands for Bhattacharya and Bhattacharya (...).]

2.1. The Real Projective Space $\mathbb{R}P^d$. This is the axial space comprising axes or lines through the origin in \mathbb{R}^{d+1} . Thus elements of $\mathbb{R}P^d$ may be represented as equivalence classes

$$(2.2) \quad [x] = [x^1 : x^2 : \dots : x^{d+1}] = \{\lambda x : \lambda \neq 0\}, \quad x \in \mathbb{R}^{d+1} \setminus \{0\}.$$

One may also identify $\mathbb{R}P^d$ with S^d/G , with G comprising the identity map and the antipodal map $p \mapsto -p$. Its structure as a d -dimensional manifold (with quotient topology) and its Riemannian structure both derive from this identification. Among applications are observations on galaxies, on axes of crystals, or on the line of a geological fissure (Watson (1983), Mardia and Jupp (1999), Fisher et al. (1987), Beran and Fisher (1998), Kendall (1989)).

2.2. Kendall's (Direct Similarity) Shape Spaces Σ_m^k . Kendall's shape spaces are quotient spaces S^d/G , under the action of the *special orthogonal group* $G = SO(m)$ of $m \times m$ orthogonal matrices with determinant +1. For the important case $m = 2$, consider the space of all planar k -ads (z_1, z_2, \dots, z_k) ($z_j = (x_j, y_j)$), $k > 2$, excluding those with k identical points. The set of all centered and normed k -ads, say $u = (u_1, u_2, \dots, u_k)$ comprise a unit sphere in a $(2k-2)$ -dimensional vector space and is, therefore, a $(2k-3)$ -dimensional sphere S^{2k-3} , called the *preshape sphere*. The group $G = SO(2)$ acts on the sphere by rotating each landmark by the same angle. The orbit under G of a point u in the preshape sphere can thus be seen to be a circle S^1 , so that Kendall's planer shape space Σ_2^k can be viewed as the quotient space $S^{2k-3}/G \sim S^{2k-3}/S^1$, a $(2k-4)$ -dimensional compact manifold. An algebraically simpler representation of Σ_2^k is given by the complex projective space $\mathbb{C}P^{k-2}$, described in Section 6.4. For many applications in archaeology, astronomy, morphometrics, medical diagnosis, etc., see Bookstein (1986, 1997), Kendall (1989), Dryden and Mardia (1998), BP (2003, 2005), BB (2008a) and Small (1996).

2.3. Reflection (Similarity) Shape Spaces $R\Sigma_m^k$. Consider now the *reflection shape* of a k -ad as defined in Section 2.2, but with $SO(m)$ replaced by the larger *orthogonal group* $O(m)$ of all $m \times m$ orthogonal matrices (with determinants either $+1$ or -1). The reflection shape space $R\Sigma_m^k$ is the space of orbits of the elements u of the preshape sphere whose columns span \mathbb{R}^m .

2.4. Affine Shape Spaces $A\Sigma_m^k$. The *affine shape* of a k -ad in \mathbb{R}^m may be defined as the orbit of this k -ad under the group of all *affine transformations* $x \mapsto F(x) = Ax + b$, where A is an arbitrary $m \times m$ non-singular matrix and b is an arbitrary point in \mathbb{R}^m . Note that two k -ads $x = (x_1, \dots, x_k)$ and $y = (y_1, \dots, y_k)$, ($x_j, y_j \in \mathbb{R}^m$ for all j) have the same affine shape if and only if the centered k -ads $u = (u_1, u_2, \dots, u_k) = (x_1 - \bar{x}, \dots, x_k - \bar{x})$ and $v = (v_1, v_2, \dots, v_k) = (y_1 - \bar{y}, \dots, y_k - \bar{y})$ are related by a transformation $Au \doteq (Au_1, \dots, Au_k) = v$. The centered k -ads lie in a linear subspace of \mathbb{R}^m of dimension $m(k-1)$. Assume $k > m + 1$. The affine shape space is then defined as the quotient space $H(m, k)/GL(m, \mathbb{R})$, where $H(m, k)$ consists of all centered k -ads whose landmarks span \mathbb{R}^m , and $GL(m, \mathbb{R})$ is the general linear group on \mathbb{R}^m (of all $m \times m$ nonsingular matrices) which has the relative topology (and distance) of \mathbb{R}^{m^2} and is a manifold of dimension m^2 . It follows that $A\Sigma_m^k$ is a manifold of dimension $m(k-1) - m^2$. For $u, v \in H(m, k)$, since $Au = v$ iff $u'A' = v'$, and as A varies $u'A'$ generates the linear subspace L of $H(m, k)$ spanned by the m rows of u . The affine shape of u , (or of x), is identified with this subspace. Thus $A\Sigma_m^k$ may be identified with the set of all m dimensional subspaces of \mathbb{R}^{k-1} , namely, the *Grassmannian* $G_m(k-1)$ -a result of Sparr (1995) (Also see Boothby (1986), pp. 63-64, 362-363). Affine shape spaces arise in certain problems of bioinformatics, cartography, machine vision and pattern recognition (Berthilsson and Heyden (1999), Berthilsson and Astrom (1999), Sepiashvili et al. (2003), Sparr (1992, 1996)).

2.5. Projective Shape Spaces $P\Sigma_m^k$. For purposes of machine vision, if images are taken from a great distance, such as a scene on the ground photographed from an airplane, affine shape analysis is appropriate. Otherwise, *projective shape* is a more appropriate choice. If one thinks of images or photographs obtained through a central projection (a pinhole camera is an example of this), a ray is received as a point on the image plane (e.g., the film of the camera). Since axes in 3D comprise the projective space $\mathbb{R}P^2$, k -ads in this view are valued in $\mathbb{R}P^2$. Note that for a 3D k -ad to represent a k -ad in $\mathbb{R}P^2$, the corresponding axes must all be distinct. To have invariance with regard to camera angles, one may first look at the original noncollinear (centered) 3D k -ad u and achieve affine invariance by its affine shape (i.e., by the equivalence class Au , $A \in GL(3, \mathbb{R})$), and finally take the corresponding equivalence class of axes in $\mathbb{R}P^2$ to define the projective shape of the k -ad as the equivalence class, or orbit, with respect to projective transformations on $\mathbb{R}P^2$. A projective shape (of a k -ad) is singular if the k axes lie on a vector plane ($\mathbb{R}P^1$). For $k > 4$, the space of all non-singular shapes is the 2D projective shape space, denoted $P_0\Sigma_2^k$.

In general, a projective (general linear) transformation α on $\mathbb{R}P^m$ is defined in terms of an $(m+1) \times (m+1)$ nonsingular matrix $A \in GL(m+1, \mathbb{R})$ by

$$(2.3) \quad \alpha([x]) = \alpha([x^1 : \dots : x^{m+1}]) = [A(x^1, \dots, x^{m+1})'],$$

where $x = (x^1, \dots, x^{m+1}) \in \mathbb{R}^{m+1} \setminus \{0\}$. The group of all projective transformations on $\mathbb{R}P^m$ is denoted by $PGL(m)$. Now consider a k -ad (y_1, \dots, y_k) in $\mathbb{R}P^m$, say $y_j = [x_j]$ ($j = 1, \dots, k$), $k > m + 2$. The projective shape of this k -ad is its orbit under $PGL(m)$, i.e., $\{(\alpha y_1, \dots, \alpha y_k) : \alpha \in PGL(m)\}$. To exclude singular shapes, define a k -ad $(y_1, \dots, y_k) = ([x_1], \dots, [x_k])$ to be in *general position* if the linear span of $\{y_1, \dots, y_k\}$ is $\mathbb{R}P^m$, i.e., if the linear span of the set of k representative points $\{x_1, \dots, x_k\}$ in \mathbb{R}^{m+1} is \mathbb{R}^{m+1} . The space of shapes of all k -ads in general position is the projective shape space $P_0\Sigma_m^k$. Define a projective frame in $\mathbb{R}P^m$ to be an ordered system of $m + 2$ points in general position. Let $I = i_1 < \dots < i_{m+2}$ be an ordered subset of $\{1, \dots, k\}$. A manifold structure on $P_I\Sigma_m^k$, the open dense subset of $P_0\Sigma_m^k$, of k -ads for which $(y_{i_1}, \dots, y_{i_{m+2}})$ is a projective frame in $\mathbb{R}P^m$, was derived in Mardia and Patrangenaru (2005) as follows. The standard frame is defined to be $([e_1], \dots, [e_{m+1}], [e_1 + e_2 + \dots + e_{m+1}])$, where $e_j \in \mathbb{R}^{m+1}$ has 1 in the j -th coordinate and zeros elsewhere. Given two projective frames (p_1, \dots, p_{m+2}) and (q_1, \dots, q_{m+2}) , there exists a unique $\alpha \in PGL(m)$ such that $\alpha(p_j) = q_j$ ($j = 1, \dots, m+2$). By ordering the points in a k -ad such that the first $m + 2$ points are in general position, one may bring this ordered set, say, (p_1, \dots, p_{m+2}) , to the standard form by a unique $\alpha \in PGL(m)$. Then the ordered set of remaining $k - m - 2$ points is transformed to a point in $(\mathbb{R}P^m)^{k-m-2}$. This provides a diffeomorphism between $P_I\Sigma_m^k$ and the product of $k - m - 2$ copies of the real projective space $\mathbb{R}P^m$.

We will return to these manifolds again in Section 6. Now we turn to nonparametric inference on general manifolds.

3. Fréchet Means on Metric Spaces

Let (M, ρ) be a metric space, ρ being the distance, and let $f \geq 0$ be a given continuous increasing function on $[0, \infty)$. For a given probability measure Q on (the Borel sigmafield of) M , define the *Fréchet function* of Q as

$$(3.1) \quad F(p) = \int_M f(\rho(p, x))Q(dx), \quad p \in M.$$

DEFINITION 3.1. Suppose $F(p) < \infty$ for some $p \in M$. Then the set of all p for which $F(p)$ is the minimum value of F on M is called the *Fréchet Mean set* of Q , denoted by C_Q . If this set is a singleton, say $\{\mu_F\}$, then μ_F is called the *Fréchet Mean* of Q . If X_1, X_2, \dots, X_n are independent and identically distributed (iid) M -valued random variables defined on some probability space (Ω, \mathcal{F}, P) with common distribution Q , and $Q_n \doteq \frac{1}{n} \sum_{j=1}^n \delta_{X_j}$ is the corresponding empirical distribution, then the Fréchet mean set of Q_n is called the *sample Fréchet mean set*, denoted by C_{Q_n} . If this set is a singleton, say $\{\mu_{F_n}\}$, then μ_{F_n} is called the *sample Fréchet mean*.

Proposition 3.1 proves the consistency of the sample Fréchet mean as an estimator of the Fréchet mean of Q .

PROPOSITION 3.1. *Let M be a compact metric space. Consider the Fréchet function F of a probability measure given by (3.1). Given any $\epsilon > 0$, there exists an integer-valued random variable $N = N(\omega, \epsilon)$ and a P -null set $A(\omega, \epsilon)$ such that*

$$(3.2) \quad C_{Q_n} \subset C_Q^\epsilon \equiv \{p \in M : \rho(p, C_Q) < \epsilon\}, \quad \forall n \geq N$$

outside of $A(\omega, \epsilon)$. In particular, if $C_Q = \{\mu_F\}$, then every measurable selection, μ_{F_n} from C_{Q_n} is a strongly consistent estimator of μ_F .

PROOF. For simplicity of notation, we write $C = C_Q$, $C_n = C_{Q_n}$, $\mu = \mu_F$ and $\mu_n = \mu_{F_n}$. Choose $\epsilon > 0$ arbitrarily. If $C^\epsilon = M$, then (3.2) holds with $N = 1$. If $D = M \setminus C^\epsilon$ is nonempty, write

$$(3.3) \quad \begin{aligned} l &= \min\{F(p) : p \in M\} = F(q) \quad \forall q \in C, \\ l + \delta(\epsilon) &= \min\{F(p) : p \in D\}, \quad \delta(\epsilon) > 0. \end{aligned}$$

It is enough to show that

$$(3.4) \quad \max\{|F_n(p) - F(p)| : p \in M\} \longrightarrow 0 \text{ a.s., as } n \rightarrow \infty.$$

For if (3.4) holds, then there exists $N \geq 1$ such that, outside a P-null set $A(\omega, \epsilon)$,

$$(3.5) \quad \begin{aligned} \min\{F_n(p) : p \in C\} &\leq l + \frac{\delta(\epsilon)}{3}, \\ \min\{F_n(p) : p \in D\} &\geq l + \frac{\delta(\epsilon)}{2}, \quad \forall n \geq N. \end{aligned}$$

Clearly (3.5) implies (3.2).

To prove (3.4), choose and fix $\epsilon' > 0$, however small. Note that $\forall p, p', x \in M$,

$$|\rho(p, x) - \rho(p', x)| \leq \rho(p, p').$$

Hence

$$(3.6) \quad \begin{aligned} |F(p) - F(p')| &\leq \max\{|f(\rho(p, x)) - f(\rho(p', x))| : x \in M\} \\ &\leq \max\{|f(u) - f(u')| : |u - u'| \leq \rho(p, p')\}, \\ |F_n(p) - F_n(p')| &\leq \max\{|f(u) - f(u')| : |u - u'| \leq \rho(p, p')\}. \end{aligned}$$

Since f is uniformly continuous on $[0, R]$ where R is the diameter of M , so are F and F_n on M , and there exists $\delta(\epsilon') > 0$ such that

$$(3.7) \quad |F(p) - F(p')| \leq \frac{\epsilon'}{4}, \quad |F_n(p) - F_n(p')| \leq \frac{\epsilon'}{4}$$

if $\rho(p, p') < \delta(\epsilon')$. Let $\{q_1, \dots, q_k\}$ be a $\delta(\epsilon')$ -net of M , i.e., $\forall p \in M$ there exists $q(p) \in \{q_1, \dots, q_k\}$ such that $\rho(p, q(p)) < \delta(\epsilon')$. By the strong law of large numbers, there exists an integer-valued random variable $N(\omega, \epsilon')$ such that outside of a P-null set $A(\omega, \epsilon')$, one has

$$(3.8) \quad |F_n(q_i) - F(q_i)| \leq \frac{\epsilon'}{4} \quad \forall i = 1, 2, \dots, k; \text{ if } n \geq N(\omega, \epsilon').$$

From (3.7) and (3.8) we get

$$\begin{aligned} |F(p) - F_n(p)| &\leq |F(p) - F(q(p))| + |F(q(p)) - F_n(q(p))| + |F_n(q(p)) - F_n(p)| \\ &\leq \frac{3\epsilon'}{4} < \epsilon', \quad \forall p \in M, \end{aligned}$$

if $n \geq N(\omega, \epsilon')$ outside of $A(\omega, \epsilon')$. This proves (3.4). \square

REMARK 3.1. Under an additional assumption guaranteeing the existence of a minimizer of F , Proposition 3.1 can be extended to all metric spaces whose closed and bounded subsets are all compact. We will consider such an extension elsewhere,

thereby generalizing Theorem 2.3 in BP (2003). For statistical analysis on shape spaces which are compact manifolds, Proposition 3.1 suffices.

REMARK 3.2. One can show that the reverse of (3.2) that is “ $C_Q \subset C_{Q_n}^\epsilon \forall n \geq N(\omega, \epsilon)$ ” does not hold in general. See for example Remark 2.6 in BP (2003).

REMARK 3.3. In view of Proposition 3.1, if the Fréchet mean μ_F of Q exists as a unique minimizer of F , then every measurable selection of a sequence $\mu_{F_n} \in C_{Q_n}$ ($n \geq 1$) converges to μ_F with probability one. In the rest of the paper it therefore suffices to define the sample Fréchet mean as a measurable selection from C_{Q_n} ($n \geq 1$).

Next we consider the asymptotic distribution of μ_{F_n} . For Theorem 3.2, we assume M to be a differentiable manifold of dimension d . Let ρ be a distance metrizing the topology of M . The proof of the theorem is similar to that of Theorem 2.1 in BP (2005). Denote by D_r the partial derivative w.r.t. the r^{th} coordinate. ($r = 1, \dots, d$).

THEOREM 3.2. *Suppose the following assumptions hold:*

A1 Q has support in a single coordinate patch, (U, ϕ) . [$\phi : U \rightarrow \mathbb{R}^d$ smooth.] Let $Y_j = \phi(X_j)$, $j = 1, \dots, n$.

A2 Fréchet mean μ_F of Q is unique.

A3 $\forall x, y \rightarrow h(x, y) = (\rho^\phi)^2(x, y) = \rho^2(\phi^{-1}x, \phi^{-1}y)$ is twice continuously differentiable in a neighborhood of $\phi(\mu_F) = \mu$.

A4 $E\{D_r h(Y, \mu)\}^2 < \infty \forall r$.

A5 $E\left\{ \sup_{|u-v| \leq \epsilon} |D_s D_r h(Y, v) - D_s D_r h(Y, u)| \right\} \rightarrow 0$ as $\epsilon \rightarrow 0 \forall r, s$.

A6 $\Lambda = ((E\{D_s D_r h(Y, \mu)\})$ is nonsingular.

A7 $\Sigma = \text{Cov}[\text{grad } h(Y_1, \mu)]$ is nonsingular.

Let $\mu_{F,n}$ be a measurable selection from the sample Fréchet mean set. Then under the assumptions **A1-A7**,

$$(3.9) \quad \sqrt{n}(\mu_n - \mu) \xrightarrow{\mathcal{L}} N(0, \Lambda^{-1} \Sigma (\Lambda')^{-1}).$$

4. Extrinsic Means on Manifolds

From now on, we assume that M is a Riemannian manifold of dimension d . Let G be a Lie group acting on M and let $J : M \rightarrow \mathbb{E}^N$ be a H -equivariant embedding of M into some euclidean space \mathbb{E}^N of dimension N . For all our applications, H is compact. Then J induces the metric

$$(4.1) \quad \rho(x, y) = \|J(x) - J(y)\|$$

on M , where $\|\cdot\|$ denotes Euclidean norm ($\|u\|^2 = \sum_{i=1}^N u_i^2 \forall u = (u_1, u_2, \dots, u_N)$). This is called the *extrinsic distance* on M .

For the Fréchet function F in (3.1), let $f(r) = r^2$ on $[0, \infty)$. This choice of the Fréchet function makes the Fréchet mean computable in a number of important examples using Proposition 4.1. Assume $J(M) = \tilde{M}$ is a closed subset of \mathbb{E}^N . Then for every $u \in \mathbb{E}^N$ there exists a compact set of points in \tilde{M} whose distance from u is the smallest among all points in \tilde{M} . We denote this set by

$$(4.2) \quad P_{\tilde{M}} u = \{x \in \tilde{M} : \|x - u\| \leq \|y - u\| \forall y \in \tilde{M}\}.$$

If this set is a singleton, u is said to be a *nonfocal point* of \mathbb{E}^N (w.r.t. \tilde{M}), otherwise it is said to be a *focal point* of \mathbb{E}^N .

DEFINITION 4.1. Let $(M, \rho), J$ be as above. Let Q be a probability measure on M such that the Fréchet function

$$(4.3) \quad F(x) = \int \rho^2(x, y)Q(dy)$$

is finite. The Fréchet mean (set) of Q is called the *extrinsic mean (set)* of Q . If $X_i, i = 1, \dots, n$ are iid observations from Q and $Q_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$, then the Fréchet mean(set) of Q_n is called the *extrinsic sample mean(set)*.

Let \tilde{Q} and \tilde{Q}_n be the images of Q and Q_n respectively in \mathbb{E}^N : $\tilde{Q} = Q \circ J^{-1}$, $\tilde{Q}_n = Q_n \circ J^{-1}$.

PROPOSITION 4.1. (a) If $\tilde{\mu} = \int_{\mathbb{E}^N} u \tilde{Q}(du)$ is the mean of \tilde{Q} , then the extrinsic mean set of Q is given by $J^{-1}(P_{\tilde{M}} \tilde{\mu})$. (b) If $\tilde{\mu}$ is a nonfocal point of \mathbb{E}^N then the extrinsic mean of Q exists (as a unique minimizer of F).

PROOF. See Proposition 3.1, BP (2003). \square

COROLLARY 4.2. If $\tilde{\mu} = \int_{\mathbb{E}^N} u \tilde{Q}(du)$ is a nonfocal point of \mathbb{E}^N then the extrinsic sample mean μ_n (any measurable selection from the extrinsic sample mean set) is a strongly consistent estimator of the extrinsic mean μ of Q .

PROOF. Follows from Proposition 3.1 for compact M . For the more general case, see BP (2003). \square

4.1. Asymptotic Distribution of the Extrinsic Sample Mean. Although one can apply Theorem 3.2 here, we prefer a different, and more widely applicable approach, which does not require that the support of Q be contained in a coordinate patch. Let $\bar{Y} = \frac{1}{n} \sum_{j=1}^n Y_j$ be the (sample) mean of $Y_j = P(X_j)$. In a neighborhood of a nonfocal point such as $\tilde{\mu}$, $P(\cdot)$ is smooth. Hence it can be shown that

$$(4.4) \quad \sqrt{n}[P(\bar{Y}) - P(\tilde{\mu})] = \sqrt{n}(d_{\tilde{\mu}}P)(\bar{Y} - \tilde{\mu}) + o_P(1)$$

where $d_{\tilde{\mu}}P$ is the differential (map) of the projection $P(\cdot)$, which takes vectors in the tangent space of \mathbb{E}^N at $\tilde{\mu}$ to tangent vectors of \tilde{M} at $P(\tilde{\mu})$. Let f_1, f_2, \dots, f_d be an orthonormal basis of $T_{P(\tilde{\mu})}J(M)$ and e_1, e_2, \dots, e_N be an orthonormal basis (frame) for $T\mathbb{E}^N \approx \mathbb{E}^N$. One has

$$(4.5) \quad \begin{aligned} \sqrt{n}(\bar{Y} - \tilde{\mu}) &= \sum_{j=1}^N \langle \sqrt{n}(\bar{Y} - \tilde{\mu}), e_j \rangle e_j, \\ d_{\tilde{\mu}}P(\sqrt{n}(\bar{Y} - \tilde{\mu})) &= \sum_{j=1}^N \langle \sqrt{n}(\bar{Y} - \tilde{\mu}), e_j \rangle d_{\tilde{\mu}}P(e_j) \\ &= \sum_{j=1}^N \langle \sqrt{n}(\bar{Y} - \tilde{\mu}), e_j \rangle \sum_{r=1}^d \langle d_{\tilde{\mu}}P(e_j), f_r \rangle f_r \\ &= \sum_{r=1}^d \left[\sum_{j=1}^N \langle d_{\tilde{\mu}}P(e_j), f_r \rangle \langle \sqrt{n}(\bar{Y} - \tilde{\mu}), e_j \rangle \right] f_r. \end{aligned}$$

Hence $\sqrt{n}[P(\bar{Y}) - P(\tilde{\mu})]$ has an asymptotic Gaussian distribution on the tangent space of $J(M)$ at $P(\tilde{\mu})$, with mean vector zero and a dispersion matrix (w.r.t. the basis vector $\{f_r : 1 \leq r \leq d\}$)

$$\Sigma = A'VA$$

where

$$A \equiv A(\tilde{\mu}) = ((\langle d_{\tilde{\mu}}P(e_j), f_r \rangle))_{1 \leq j \leq N, 1 \leq r \leq d}$$

and V is the $N \times N$ covariance matrix of $\tilde{Q} = Q \circ J^{-1}$ (w.r.t. the basis $\{e_j : 1 \leq j \leq N\}$). In matrix notation,

$$(4.6) \quad \sqrt{n}\bar{T} \xrightarrow{\mathcal{L}} N(0, \Sigma) \text{ as } n \rightarrow \infty,$$

where

$$T_j(\tilde{\mu}) = A'[\langle (Y_j - \tilde{\mu}), e_1 \rangle \dots \langle (Y_j - \tilde{\mu}), e_N \rangle]', \quad j = 1, \dots, n$$

and

$$\bar{T} \equiv \bar{T}(\tilde{\mu}) = \frac{1}{n} \sum_{j=1}^n T_j(\tilde{\mu}).$$

This implies, writing \mathcal{X}_d^2 for the chisquare distribution with d degrees of freedom,

$$(4.7) \quad n\bar{T}'\Sigma^{-1}\bar{T} \xrightarrow{\mathcal{L}} \mathcal{X}_d^2, \text{ as } n \rightarrow \infty.$$

A confidence region for $P(\tilde{\mu})$ with asymptotic confidence level $1 - \alpha$ is then given by

$$(4.8) \quad \{P(\tilde{\mu}) : n\bar{T}'\hat{\Sigma}^{-1}\bar{T} \leq \mathcal{X}_d^2(1 - \alpha)\}$$

where $\hat{\Sigma} \equiv \hat{\Sigma}(\tilde{\mu})$ is the sample covariance matrix of $\{T_j(\tilde{\mu})\}_{j=1}^n$. The corresponding bootstrapped confidence region is given by

$$(4.9) \quad \{P(\tilde{\mu}) : n\bar{T}'\hat{\Sigma}^{-1}\bar{T} \leq c_{(1-\alpha)}^*\}$$

where $c_{(1-\alpha)}^*$ is the upper $(1 - \alpha)$ -quantile of the bootstrapped values U^* , $U^* = n\bar{T}^{*'}\hat{\Sigma}^{*-1}\bar{T}^*$ and \bar{T}^* , $\hat{\Sigma}^*$ being the sample mean and covariance respectively of the bootstrap sample $\{T_j^*(\bar{Y})\}_{j=1}^n$.

5. Intrinsic Means on Manifolds

Let (M, g) be a complete connected Riemannian manifold with metric tensor g . Then the natural choice for the distance metric ρ in Section 3 is the geodesic distance d_g on M . Unless otherwise stated, we consider the function $f(r) = r^2$ in (3.1) throughout this section and later sections. However one may take more general f . For example one may consider $f(r) = r^a$, for suitable $a \geq 1$.

Let Q be a probability distribution on M with finite Fréchet function

$$(5.1) \quad F(p) = \int_M d_g^2(p, m)Q(dm).$$

Let X_1, \dots, X_n be an iid sample from Q .

DEFINITION 5.1. The Fréchet mean set of Q under $\rho = d_g$ is called the *intrinsic mean set* of Q . The Fréchet mean set of the empirical distribution Q_n is called the *sample intrinsic mean set*.

Before proceeding further, let us define a few technical terms related to Riemannian manifolds which we will use extensively in this section. For details on Riemannian Manifolds, see DoCarmo (1992), Gallot et al. (1990) or Lee (1997).

- (1) *Geodesic*: These are curves γ on the manifold with zero acceleration. They are locally length minimizing curves. For example, consider great circles on the sphere or straight lines in \mathbb{R}^d .
- (2) *Exponential map*: For $p \in M$, $v \in T_p M$, we define $\exp_p v = \gamma(1)$, where γ is a geodesic with $\gamma(0) = p$ and $\dot{\gamma}(0) = v$.
- (3) *Cut locus*: For a point $p \in M$, define the cut locus $C(p)$ of p as the set of points of the form $\gamma(t_0)$, where γ is a unit speed geodesic starting at p and t_0 is the supremum of all $t > 0$ such that γ is distance minimizing from p to $\gamma(t)$. For example, $C(p) = \{-p\}$ on the sphere.
- (4) *Sectional Curvature*: Recall the notion of Gaussian curvature of two dimensional surfaces. On a Riemannian manifold M , choose a pair of linearly independent vectors $u, v \in T_p M$. A two dimensional submanifold of M is swept out by the set of all geodesics starting at p and with initial velocities lying in the two-dimensional section π spanned by u, v . The Gaussian curvature of this submanifold is called the sectional curvature at p of the section π .
- (5) *Injectivity Radius*: Define the injectivity radius of M as

$$\text{inj}(M) = \inf\{d_g(p, C(p)) : p \in M\}.$$

For example the sphere of radius 1 has injectivity radius equal to π .

Also let $r_* = \min\{\text{inj}(M), \frac{\pi}{\sqrt{\bar{C}}}\}$, where \bar{C} is the least upper bound of sectional curvatures of M if this upper bound is positive, and $\bar{C} = 0$ otherwise. The exponential map at p is injective on $\{v \in T_p(M) : |v| < r_*\}$. By $B(p, r)$ we will denote an open ball with center $p \in M$ and radius r , and $\bar{B}(p, r)$ will denote its closure.

In case Q has a unique intrinsic mean μ_I , it follows from Proposition 3.1 and Remark 3.1 that the sample intrinsic mean μ_{nI} (a measurable selection from the sample intrinsic mean set) is a consistent estimator of μ_I . Broad conditions for the existence of a unique intrinsic mean are not known. From results due to Karchar (1997) and Le (2001), it follows that if the support of Q is in a geodesic ball of radius $\frac{r_*}{4}$, i.e. $\text{supp}(Q) \subseteq B(p, \frac{r_*}{4})$, then Q has a unique intrinsic mean. This result has been substantially extended by Kendall (1990) which shows that if $\text{supp}(Q) \subseteq B(p, \frac{r_*}{2})$, then there is a unique local minimum of the Fréchet function F in that ball. Then we redefine the (local) intrinsic mean of Q as that unique minimizer in the ball. In that case one can show that the (local) sample intrinsic mean is a consistent estimator of the intrinsic mean of Q . This is stated in Proposition 5.1.

PROPOSITION 5.1. *Let Q have support in $B(p, \frac{r_*}{2})$ for some $p \in M$. Then (a) Q has a unique (local) intrinsic mean μ_I in $B(p, \frac{r_*}{2})$ and (b) the sample intrinsic mean μ_{nI} in $B(p, \frac{r_*}{2})$ is a strongly consistent estimator of μ_I .*

PROOF. (a) Follows from Kendall (1990).

(b) Since $\text{supp}(Q)$ is compact, $\text{supp}(Q) \subseteq \bar{B}(p, r)$ for some $r < \frac{r_*}{2}$. From Lemma 1, Le (2001), it follows that $\mu_I \in B(p, r)$ and μ_I is the unique intrinsic mean of Q restricted to $\bar{B}(p, r)$. Now take the compact metric space in Proposition 3.1 to be $\bar{B}(p, r)$ and the result follows. \square

For the asymptotic distribution of the sample intrinsic mean, we may use Theorem 3.2. For that we need to verify assumptions A1-A7. Theorem 5.2 gives sufficient

conditions for that. In the statement of the theorem, the usual partial order $A \geq B$ between $d \times d$ symmetric matrices A, B , means that $A - B$ is nonnegative definite.

THEOREM 5.2. *Assume $\text{supp}(Q) \subseteq B(p, \frac{r_*}{2})$. Let $\phi = \exp_{\mu_I}^{-1} : B(p, \frac{r_*}{2}) \rightarrow T_{\mu_I}M (\approx \mathbb{R}^d)$. Then the map $y \mapsto h(x, y) = d_g^2(\phi^{-1}x, \phi^{-1}y)$ is twice continuously differentiable in a neighborhood of 0 and in terms of normal coordinates with respect to a chosen orthonormal basis for $T_{\mu_I}M$,*

$$(5.2) \quad D_r h(x, 0) = -2x^r, \quad 1 \leq r \leq d,$$

$$(5.3) \quad [D_r D_s h(x, 0)] \geq \left[2 \left\{ \left(\frac{1 - f(|x|)}{|x|^2} \right) x^r x^s + f(|x|) \delta_{rs} \right\} \right]_{1 \leq r, s \leq d}.$$

Here $x = (x^1, \dots, x^d)'$, $|x| = \sqrt{(x^1)^2 + (x^2)^2 + \dots + (x^d)^2}$ and

$$(5.4) \quad f(y) = \begin{cases} 1 & \text{if } \bar{C} = 0 \\ \sqrt{\bar{C}} y \frac{\cos(\sqrt{\bar{C}} y)}{\sin(\sqrt{\bar{C}} y)} & \text{if } \bar{C} > 0 \\ \sqrt{-\bar{C}} y \frac{\cosh(\sqrt{-\bar{C}} y)}{\sinh(\sqrt{-\bar{C}} y)} & \text{if } \bar{C} < 0 \end{cases}$$

There is equality in (5.3) when M has constant sectional curvature \bar{C} , and in this case Λ has the expression:

$$(5.5) \quad \Lambda_{rs} = 2E \left\{ \left(\frac{1 - f(|\tilde{X}_1|)}{|\tilde{X}_1|^2} \right) \tilde{X}_1^r \tilde{X}_1^s + f(|\tilde{X}_1|) \delta_{rs} \right\}, \quad 1 \leq r, s \leq d.$$

Λ is positive definite if $\text{supp}(Q) \in B(\mu_I, \frac{r_*}{2})$.

PROOF. See Theorem 2.2., BB (2008b). □

From Theorem 5.2 it follows that $\Sigma = 4\text{Cov}(Y_1)$ where $Y_j = \phi(X_j)$, $j = 1, \dots, n$ are the normal coordinates of the sample X_1, \dots, X_n from Q . It is nonsingular if $Q \circ \phi^{-1}$ has support in no smaller dimensional subspace of \mathbb{R}^d . That holds if for example Q has a density with respect to the volume measure on M .

6. Applications

In this section we apply the results of the earlier sections to some important manifolds. We start with the unit sphere S^d in \mathbb{R}^{d+1} .

6.1. S^d . Consider the space of all directions in \mathbb{R}^{d+1} which can be identified with the unit sphere

$$S^d = \{x \in \mathbb{R}^{d+1} : \|x\| = 1\}.$$

Statistics on S^2 , often called *directional statistics*, have been among the earliest and most widely used statistics on manifolds. (See, e.g., Watson (1983), Fisher et al. (1996), Mardia and Jupp (1999)). Among important applications, we cite paleomagnetism, where one may detect and/or study the shifting of magnetic poles on earth over geological times. Another application is the estimation of the direction of a signal.

6.1.1. Extrinsic Mean on S^d . The inclusion map $i : S^d \rightarrow \mathbb{R}^{d+1}$, $i(x) = x$ provides a natural embedding for S^d into \mathbb{R}^{d+1} . The extrinsic mean set of a probability distribution Q on S^d is then the set $P_{S^d} \tilde{\mu}$ on S^d closest to $\tilde{\mu} = \int_{\mathbb{R}^{d+1}} x \tilde{Q}(dx)$, where \tilde{Q} is Q regarded as a probability measure on \mathbb{R}^{d+1} . Note that $\tilde{\mu}$ is non-focal iff $\tilde{\mu} \neq 0$ and then Q has a unique extrinsic mean $\mu = \frac{\tilde{\mu}}{\|\tilde{\mu}\|}$.

6.1.2. **Intrinsic Mean on S^d .** At each $p \in S^d$, endow the tangent space $T_p S^d = \{v \in \mathbb{R}^{d+1} : v \cdot p = 0\}$ with the metric tensor $g_p : T_p \times T_p \rightarrow \mathbb{R}$ as the restriction of the scalar product at p of the tangent space of $\mathbb{R}^{d+1} : g_p(v_1, v_2) = v_1 \cdot v_2$. The geodesics are the big circles,

$$(6.1) \quad \gamma_{p,v}(t) = (\cos t|v|)p + (\sin t|v|)\frac{v}{|v|}.$$

The exponential map, $exp_p : T_p S^d \rightarrow S^d$ is

$$(6.2) \quad exp_p(v) = \cos(|v|)p + \sin(|v|)\frac{v}{|v|},$$

and the geodesic distance is

$$(6.3) \quad d_g(p, q) = \arccos(p \cdot q) \in [0, \pi].$$

This space has constant sectional curvature 1 and injectivity radius π . Hence if Q has support in an open ball of radius $\frac{\pi}{2}$, then it has a unique intrinsic mean in that ball.

6.2. $\mathbb{R}P^d$. Consider the real projective space $\mathbb{R}P^d$ of all lines through the origin in \mathbb{R}^{d+1} . The elements of $\mathbb{R}P^d$ may be represented as $[u] = \{-u, u\}$ ($u \in S^d$).

6.2.1. **Extrinsic Mean on $\mathbb{R}P^d$.** $\mathbb{R}P^d$ can be embedded into the space of $k \times k$ real symmetric matrices $S(k, \mathbb{R})$, $k = d + 1$ via the *Veronese-Whitney embedding* $J : \mathbb{R}P^d \rightarrow S(k, \mathbb{R})$ which is given by

$$(6.4) \quad J([u]) = uu' = ((u_i u_j))_{1 \leq i, j \leq k} \quad (u = (u_1, \dots, u_k)' \in S^d).$$

As a linear subspace of \mathbb{R}^{k^2} , $S(k, \mathbb{R})$ has the Euclidean distance

$$(6.5) \quad \|A - B\|^2 \equiv \sum_{1 \leq i, j \leq k} (a_{ij} - b_{ij})^2 = \text{Trace}(A - B)(A - B)'$$

This endows $\mathbb{R}P^d$ with the extrinsic distance ρ given by

$$(6.6) \quad \rho^2([u], [v]) = \|uu' - vv'\|^2 = 2(1 - (u'v)^2).$$

Let Q be a probability distribution on $\mathbb{R}P^d$ and let $\tilde{\mu}$ be the mean of $\tilde{Q} = Q \circ J^{-1}$ considered as a probability measure on $S(k, \mathbb{R})$. Then $\tilde{\mu} \in S^+(k, \mathbb{R})$ -the space of $k \times k$ real symmetric nonnegative definite matrices, and the projection of $\tilde{\mu}$ into $J(\mathbb{R}P^d)$ is given by the set of all uu' where u is a unit eigenvector of $\tilde{\mu}$ corresponding to the largest eigenvalue. Hence the projection is unique, i.e. $\tilde{\mu}$ is nonfocal iff its largest eigenvalue is simple, i.e., if the eigenspace corresponding to the largest eigenvalue is one dimensional. In that case the extrinsic mean of Q is $[u]$, u being a unit eigenvector in the eigenspace of the largest eigenvalue..

6.2.2. **Intrinsic Mean on $\mathbb{R}P^d$.** $\mathbb{R}P^d$ is a complete Riemannian manifold with geodesic distance

$$(6.7) \quad d_g([p], [q]) = \arccos(|p \cdot q|) \in [0, \frac{\pi}{2}].$$

It has constant sectional curvature 4 and injectivity radius $\frac{\pi}{2}$. Hence if the support of Q is contained in an open geodesic ball of radius $\frac{\pi}{4}$, it has a unique intrinsic mean in that ball.

6.3. Σ_m^k . Consider a set of k points in \mathbb{R}^m , not all points being the same. Such a set is called a k -ad or a configuration of k landmarks. We will denote a k -ad by the $m \times k$ matrix, $x = [x_1 \dots x_k]$ where $x_i, i = 1, \dots, k$ are the k landmarks from the object of interest. Assume $k > m$. The *direct similarity shape* of the k -ad is what remains after we remove the effects of translation, rotation and scaling. To remove translation, we subtract the mean $\bar{x} = \frac{1}{k} \sum_{i=1}^k x_i$ from each landmark to get the centered k -ad $w = [x_1 - \bar{x} \dots x_k - \bar{x}]$. We remove the effect of scaling by dividing w by its euclidean norm to get

$$(6.8) \quad u = \left[\frac{x_1 - \bar{x}}{\|w\|} \dots \frac{x_k - \bar{x}}{\|w\|} \right] = [u_1 u_2 \dots u_k].$$

This u is called the *preshape* of the k -ad x and it lies in the unit sphere S_m^k in the hyperplane

$$(6.9) \quad H_m^k = \{u \in \mathbb{R}^{km} : \sum_{j=1}^k u_j = 0\}.$$

Thus the preshape space S_m^k may be identified with the sphere S^{km-m-1} . Then the shape of the k -ad x is the orbit of z under left multiplication by $m \times m$ rotation matrices. In other words $\Sigma_m^k = S^{km-m-1}/SO(m)$. The cases of importance are $m = 2, 3$. Next we turn to the case $m = 2$.

6.4. Σ_2^k . As pointed out in Sections 2.2 and 6.3, $\Sigma_2^k = S^{2k-3}/SO(2)$. For a simpler representation, we denote a k -ad in the plane by a set of k complex numbers. The preshape of this complex k -vector x is $z = \frac{x - \bar{x}}{\|x - \bar{x}\|}$, $x = (x_1, \dots, x_k) \in \mathbb{C}^k$, $\bar{x} = \frac{1}{k} \sum_{i=1}^k x_i$. z lies in the complex sphere

$$(6.10) \quad S_2^k = \{z \in \mathbb{C}^k : \sum_{j=1}^k |z_j|^2 = 1, \sum_{j=1}^k z_j = 0\}$$

which may be identified with the real sphere of dimension $2k - 3$. Then the shape of x can be represented as the orbit

$$(6.11) \quad \sigma(x) = \sigma(z) = \{e^{i\theta} z : -\pi < \theta \leq \pi\}$$

and

$$(6.12) \quad \Sigma_2^k = \{\sigma(z) : z \in S_2^k\}.$$

Thus Σ_2^k has the structure of the complex projective space $\mathbb{C}P^{k-2}$ of all complex lines through the origin in \mathbb{C}^{k-1} , an important and well studied manifold in differential geometry (See Gallot et al. (1993), pp. 63-65, 97-100, BB (2008b)).

6.4.1. Extrinsic Mean on Σ_2^k . Σ_2^k can be embedded into $S(k, \mathbb{C})$ -the space of $k \times k$ complex Hermitian matrices, via the *Veronese-Whitney embedding*

$$(6.13) \quad J : \Sigma_2^k \rightarrow S(k, \mathbb{C}), J(\sigma(z)) = zz^*.$$

J is equivariant under the action of $SU(k)$ -the group of $k \times k$ complex matrices Γ such that $\Gamma^* \Gamma = I$, $\det(\Gamma) = 1$. To see this, let $\Gamma \in SU(k)$. Then Γ defines a diffeomorphism,

$$(6.14) \quad \Gamma : \Sigma_2^k \rightarrow \Sigma_2^k, \Gamma(\sigma(z)) = \sigma(\Gamma(z)).$$

The map ϕ_Γ on $S(k, \mathbb{C})$ defined by

$$(6.15) \quad \phi_\Gamma(A) = \Gamma A \Gamma^*$$

preserves distances and has the property

$$(6.16) \quad (\phi_\Gamma)^{-1} = \phi_{\Gamma^{-1}}, \quad \phi_{\Gamma_1\Gamma_2} = \phi_{\Gamma_1} \circ \phi_{\Gamma_2}.$$

That is (6.15) defines a group homomorphism from $SU(k)$ into a group of isometries of $S(k, \mathbb{C})$. Finally note that $J(\Gamma(\sigma(z))) = \phi_\Gamma(J(\sigma(z)))$. Informally, the symmetries $SU(k)$ of Σ_2^k are preserved by the embedding J .

$S(k, \mathbb{C})$ is a (real) vector space of dimension k^2 . It has the Euclidean distance,

$$(6.17) \quad \|A - B\|^2 = \sum_{i,j} |a_{ij} - b_{ij}|^2 = \text{Trace}(A - B)^2.$$

Thus the extrinsic distance ρ on Σ_2^k induced from the Veronese-Whitney embedding is given by

$$(6.18) \quad \rho^2(\sigma(x), \sigma(y)) = \|uu^* - vv^*\|^2 = 2(1 - |u^*v|^2),$$

where x and y are two k -ads, u and v are their preshapes respectively.

Let Q be a probability distribution on Σ_2^k and let $\tilde{\mu}$ be the mean of $\tilde{Q} = Q \circ J^{-1}$, regarded as a probability measure on \mathbb{C}^{k^2} . Then $\tilde{\mu} \in S_+(k, \mathbb{C})$: the space of $k \times k$ complex positive semidefinite matrices. Its projection into $J(\Sigma_2^k)$ is given by $P(\tilde{\mu}) = \{uu^*\}$ where u is a unit eigenvector of $\tilde{\mu}$ corresponding to its largest eigenvalue. The projection is unique, i.e. $\tilde{\mu}$ is nonfocal, and Q has a unique extrinsic mean μ_E , iff the eigenspace for the largest eigenvalue of $\tilde{\mu}$ is (complex) one dimensional, and then $\mu_E = \sigma(u)$, $u (\neq 0) \in$ eigenspace of the largest eigenvalue of $\tilde{\mu}$. Let X_1, \dots, X_n be an iid sample from Q . If $\tilde{\mu}$ is nonfocal, the sample extrinsic mean μ_{nE} is a consistent estimator of μ_E and $J(\mu_{nE})$ has an asymptotic Gaussian distribution on the tangent space $T_{P(\tilde{\mu})}J(\Sigma_2^k)$ (see Section 4),

$$(6.19) \quad \sqrt{n}(J(\mu_{nE}) - J(\mu_E)) = \sqrt{n}d_{\tilde{\mu}}P(\bar{X} - \tilde{\mu}) + o_P(1) \xrightarrow{\mathcal{L}} N(0, \Sigma).$$

Here $\bar{X}_j = J(X_j)$, $j = 1, \dots, n$. In (6.19), $d_{\tilde{\mu}}P(\bar{X} - \tilde{\mu})$ has coordinates

$$(6.20) \quad \bar{T}(\tilde{\mu}) = (\sqrt{2}\text{Re}(U_a^* \bar{X} U_k), \sqrt{2}\text{Im}(U_a^* \bar{X} U_k))_{a=2}^{k-1}$$

with respect to the basis

$$(6.21) \quad \{(\lambda_k - \lambda_a)^{-1} U v_k^a U^*, (\lambda_k - \lambda_a)^{-1} U w_k^a U^*\}_{a=2}^{k-1}$$

for $T_{P(\tilde{\mu})}J(\Sigma_2^k)$ (see Section 3.3, BB (2008a)). Here $U = [U_1 \dots U_k] \in SO(k)$ is such that $U^* \tilde{\mu} U = D \equiv \text{Diag}(\lambda_1, \dots, \lambda_k)$, $\lambda_1 \leq \dots \leq \lambda_{k-1} < \lambda_k$ being the eigenvalues of $\tilde{\mu}$. $\{v_b^a : 1 \leq a \leq b \leq k\}$ and $\{w_b^a : 1 \leq a < b \leq k\}$ is the canonical orthonormal basis frame for $S(k, \mathbb{C})$, defined as

$$(6.22) \quad v_b^a = \begin{cases} \frac{1}{\sqrt{2}}(e_a e_b^t + e_b e_a^t), & a < b \\ e_a e_a^t, & a = b \end{cases}$$

$$w_b^a = \frac{i}{\sqrt{2}}(e_a e_b^t - e_b e_a^t), \quad a < b$$

where $\{e_a : 1 \leq a \leq k\}$ is the standard canonical basis for \mathbb{R}^k .

Given two independent samples X_1, \dots, X_n iid Q_1 and Y_1, \dots, Y_m iid Q_2 on Σ_2^k , we may like to test if $Q_1 = Q_2$ by comparing their extrinsic mean shapes. Let μ_{iE} denote the extrinsic mean of Q_i and let μ_i be the mean of $Q_i \circ J^{-1}$ $i = 1, 2$.

Then $\mu_{iE} = J^{-1}P(\mu_i)$, and we wish to test $H_0 : P(\mu_1) = P(\mu_2)$. Let $\tilde{X}_j = J(X_j)$, $j = 1, \dots, n$ and $\tilde{Y}_j = J(Y_j)$, $j = 1, \dots, m$. Let T_j, S_j denote the asymptotic coordinates for \tilde{X}_j, \tilde{Y}_j respectively in $T_{P(\hat{\mu})}J(\Sigma_2^k)$ as defined in (6.20). Here $\hat{\mu} = \frac{n\bar{X} + m\bar{Y}}{m+n}$ is the pooled sample mean. We use the two sample test statistic

$$(6.23) \quad T_{nm} = (\bar{T} - \bar{S})' \left(\frac{1}{n} \hat{\Sigma}_1 + \frac{1}{m} \hat{\Sigma}_2 \right)^{-1} (\bar{T} - \bar{S}).$$

Here $\hat{\Sigma}_1, \hat{\Sigma}_2$ denote the sample covariances of T_j, S_j respectively. Under H_0 , $T_{nm} \xrightarrow{\mathcal{L}} \chi_{2k-4}^2$ (see Section 3.4, BB (2008a)). Hence given level α , we reject H_0 if $T_{nm} > \chi_{2k-4}^2(1 - \alpha)$.

6.4.2. **Intrinsic Mean on Σ_2^k .** Identified with $\mathbb{C}P^{k-2}$, Σ_2^k is a complete connected Riemannian manifold. It has all sectional curvatures bounded between 1 and 4 and injectivity radius of $\frac{\pi}{2}$ (see Gallot et al. (1990), pp. 97-100, 134). Hence if $\text{supp}(Q) \in B(p, \frac{\pi}{4})$, $p \in \Sigma_2^k$, it has a unique intrinsic mean μ_I in the ball.

Let X_1, \dots, X_n be iid Q and let μ_{nI} denote the sample intrinsic mean. Under the hypothesis of Theorem 5.2,

$$(6.24) \quad \sqrt{n}(\phi(\mu_{nI}) - \phi(\mu_I)) \xrightarrow{\mathcal{L}} N(0, \Lambda^{-1}\Sigma\Lambda^{-1}).$$

However Theorem 5.2 does not provide an analytic computation of Λ , since Σ_2^k does not have constant sectional curvature. Proposition 6.1 below gives the precise expression for Λ . It also relaxes the support condition required for Λ to be positive definite.

PROPOSITION 6.1. *With respect to normal coordinates, $\phi : B(p, \frac{\pi}{4}) \rightarrow \mathbb{C}^{k-2} (\approx \mathbb{R}^{2k-4})$, Λ as defined in Theorem 3.2 has the following expression:*

$$(6.25) \quad \Lambda = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{12}' & \Lambda_{22} \end{bmatrix}$$

where for $1 \leq r, s \leq k-2$,

$$\begin{aligned} (\Lambda_{11})_{rs} &= 2E \left[d_1 \cot(d_1) \delta_{rs} - \frac{\{1 - d_1 \cot(d_1)\}}{d_1^2} (\text{Re}\tilde{X}_{1,r})(\text{Re}\tilde{X}_{1,s}) \right. \\ &\quad \left. + \frac{\tan(d_1)}{d_1} (\text{Im}\tilde{X}_{1,r})(\text{Im}\tilde{X}_{1,s}) \right], \\ (\Lambda_{22})_{rs} &= 2E \left[d_1 \cot(d_1) \delta_{rs} - \frac{\{1 - d_1 \cot(d_1)\}}{d_1^2} (\text{Im}\tilde{X}_{1,r})(\text{Im}\tilde{X}_{1,s}) \right. \\ &\quad \left. + \frac{\tan(d_1)}{d_1} (\text{Re}\tilde{X}_{1,r})(\text{Re}\tilde{X}_{1,s}) \right], \end{aligned}$$

$$(\Lambda_{12})_{rs} = -2E \left[\frac{\{1 - d_1 \cot(d_1)\}}{d_1^2} (\text{Re}\tilde{X}_{1,r})(\text{Im}\tilde{X}_{1,s}) + \frac{\tan(d_1)}{d_1} (\text{Im}\tilde{X}_{1,r})(\text{Re}\tilde{X}_{1,s}) \right]$$

where $d_1 = d_g(X_1, \mu_I)$ and $\tilde{X}_j \equiv (\tilde{X}_{j,1}, \dots, \tilde{X}_{j,k-2}) = \phi(X_j)$, $j = 1, \dots, n$. Λ is positive definite if $\text{supp}(Q) \in B(\mu_I, 0.37\pi)$.

PROOF. See Theorem 3.1, BB (2008b). □

Note that with respect to a chosen orthonormal basis $\{v_1, \dots, v_{k-2}\}$ for $T_{\mu_I}\Sigma_2^k$, ϕ has the expression

$$\phi(m) = (\tilde{m}_1, \dots, \tilde{m}_{k-2})'$$

where

$$(6.26) \quad \tilde{m}_j = \frac{r}{\sin r} e^{i\theta} \bar{v}_j' z, \quad r = d_g(m, \mu_I) = \arccos(|z_0' \bar{z}|), \quad e^{i\theta} = \frac{z_0' \bar{z}}{|z_0' \bar{z}|}.$$

Here z, z_0 are the preshapes of m, μ_I respectively (see Section 3, BB (2008b)).

Given two independent samples X_1, \dots, X_n iid Q_1 and Y_1, \dots, Y_m iid Q_2 , one may test if Q_1 and Q_2 have the same intrinsic mean μ_I . The test statistic used is

$$(6.27) \quad T_{nm} = (n+m)(\hat{\phi}(\mu_{nI}) - \hat{\phi}(\mu_{mI}))' \hat{\Sigma}^{-1} (\hat{\phi}(\mu_{nI}) - \hat{\phi}(\mu_{mI})).$$

Here μ_{nI} and μ_{mI} are the sample intrinsic means for the X and Y samples respectively and $\hat{\mu}$ is the pooled sample intrinsic mean. Then $\hat{\phi} = \exp_{\hat{\mu}}^{-1}$ gives normal coordinates on the tangent space at $\hat{\mu}$, and $\hat{\Sigma} = (m+n) \left(\frac{1}{n} \hat{\Lambda}_1^{-1} \hat{\Sigma}_1 \hat{\Lambda}_1^{-1} + \frac{1}{m} \hat{\Lambda}_2^{-1} \hat{\Sigma}_2 \hat{\Lambda}_2^{-1} \right)$, where (Λ_1, Σ_1) and (Λ_2, Σ_2) are the parameters in the asymptotic distribution of $\sqrt{n}(\phi(\mu_{nI}) - \phi(\mu_I))$ and $\sqrt{m}(\phi(\mu_{mI}) - \phi(\mu_I))$ respectively, as defined in Theorem 3.2., and $(\hat{\Lambda}_1, \hat{\Sigma}_1)$ and $(\hat{\Lambda}_2, \hat{\Sigma}_2)$ are consistent sample estimates. Assuming H_0 to be true, $T_{nm} \xrightarrow{\mathcal{L}} \mathcal{X}_{2k-4}^2$ (see Section 4.1, BB (2008a)). Hence we reject H_0 at asymptotic level $1 - \alpha$ if $T_{nm} > \mathcal{X}_{2k-4}^2(1 - \alpha)$.

6.5. $R\Sigma_m^k$. For $m > 2$, the direct similarity shape space Σ_m^k fails to be a manifold. That is because the action of $SO(m)$ is not in general free (see, e.g., Kendall et al. (1999) and Small (1996)). To avoid that one may consider the shape of only those k -ads whose preshapes have rank at least $m - 1$. This subset is a manifold but not complete (in its geodesic distance). Alternatively one may also remove the effect of reflection and redefine shape of a k -ad x as

$$(6.28) \quad \sigma(x) = \sigma(z) = \{Az : A \in O(m)\}$$

where z is the preshape. Then $R\Sigma_m^k$ is the space of all such shapes where rank of z is m . In other words

$$(6.29) \quad R\Sigma_m^k = \{\sigma(z) : z \in S_m^k, \text{rank}(z) = m\}.$$

This is a manifold. It has been shown that the map

$$(6.30) \quad J : R\Sigma_m^k \rightarrow S(k, \mathbb{R}), \quad J(\sigma(z)) = z'z$$

is an embedding of the reflection shape space into $S(k, \mathbb{R})$ (see Bandulasiri and Patrangenaru (2005), Bandulasiri et al. (2007), and Dryden et al. (2007)) and is H -equivariant where $H = O(k)$ acts on the right: $A\sigma(z) \doteq \sigma(zA')$, $A \in O(k)$.

Let Q be a probability distribution on $R\Sigma_m^k$ and let $\tilde{\mu}$ be the mean of $Q \circ J^{-1}$ regarded as a probability measure on $S(k, \mathbb{R})$. Then $\tilde{\mu}$ is positive semi-definite with rank atleast m . Let $\tilde{\mu} = UDU'$ be the singular value decomposition of $\tilde{\mu}$, where $D = \text{Diag}(\lambda_1, \dots, \lambda_k)$ consists of ordered eigen values $\lambda_1 \geq \dots \geq \lambda_m \geq \dots \geq \lambda_k \geq 0$ of $\tilde{\mu}$, and $U = [U_1 \dots U_k]$ is a matrix in $SO(k)$ whose columns are the corresponding orthonormal eigen vectors. Then we may define the mean reflection shape set of Q as the set

$$(6.31) \quad \{\mu \in R\Sigma_m^k : J(\mu) = \frac{\sum_{j=1}^m \lambda_j U_j U_j'}{\sum_{j=1}^m \lambda_j}\}$$

The set in (6.31) is a singleton, and hence Q has a unique mean reflection shape μ iff $\lambda_m > \lambda_{m+1}$. Then $\mu = \sigma(u)$ where

$$(6.32) \quad u = \left[\sqrt{\frac{\lambda_1}{\sum_{j=1}^m \lambda_j}} U_1 \dots \sqrt{\frac{\lambda_m}{\sum_{j=1}^m \lambda_j}} U_m \right]'$$

6.6. $A\Sigma_m^k$. Let z be a centered k -ad in $H(m, k)$, and let $\sigma(z)$ denote its affine shape, as defined in Section 2.4. Consider the map

$$(6.33) \quad J : A\Sigma_m^k \rightarrow S(k, \mathbb{R}), \quad J(\sigma(z)) \equiv P = FF'$$

where $F = [f_1 f_2 \dots f_m]$ is an orthonormal basis for the row space of z . This is an embedding of $A\Sigma_m^k$ into $S(k, \mathbb{R})$ with the image

$$(6.34) \quad J(A\Sigma_m^k) = \{A \in S(k, \mathbb{R}) : A^2 = A, \text{Trace}(A) = m, A\mathbf{1} = 0\}.$$

It is equivariant under the action of $O(k)$ (see Dimitric (1996)).

PROPOSITION 6.2. *Let Q be a probability distribution on $A\Sigma_m^k$ and let $\tilde{\mu}$ be the mean of $Q \circ J^{-1}$ in $S(k, \mathbb{R})$. The projection of $\tilde{\mu}$ into $J(A\Sigma_m^k)$ is given by*

$$(6.35) \quad P(\tilde{\mu}) = \left\{ \sum_{j=1}^m U_j U_j' \right\}$$

where $U = [U_1 \dots U_k] \in SO(k)$ is such that $U' \tilde{\mu} U = D \equiv \text{Diag}(\lambda_1, \dots, \lambda_k)$, $\lambda_1 \geq \dots \geq \lambda_m \geq \dots \geq \lambda_k$. $\tilde{\mu}$ is nonfocal and Q has a unique extrinsic mean μ_E iff $\lambda_m > \lambda_{m+1}$. Then $\mu_E = \sigma(F')$ where $F = [U_1 \dots U_m]$.

PROOF. See Sughatadasa (2006). □

6.7. $P_0\Sigma_m^k$. Consider the diffeomorphism between $P_I\Sigma_m^k$ and $(\mathbb{R}P^m)^{k-m-2}$ as defined in Section 2.5. Using that one can embed $P_I\Sigma_m^k$ into $S(m+1, \mathbb{R})^{k-m-2}$ via the Veronese Whitney embedding of Section 6.2 and perform extrinsic analysis in a dense open subset of $P_0\Sigma_m^k$.

7. Examples

7.1. Example 1: Gorilla Skulls. To test the difference in the shapes of skulls of male and female gorillas, eight landmarks were chosen on the midline plane of the skulls of 29 male and 30 female gorillas. The data can be found in Dryden and Mardia (1998), pp. 317-318. Thus we have two iid samples in Σ_2^k , $k = 8$. The sample extrinsic mean shapes for the female and male samples are denoted by $\hat{\mu}_{1E}$ and $\hat{\mu}_{2E}$ where

$$\begin{aligned} \hat{\mu}_{1E} &= \sigma[-0.3586 + 0.3425i, 0.3421 - 0.2943i, 0.0851 - 0.3519i, -0.0085 - 0.2388i, \\ &\quad -0.1675 + 0.0021i, -0.2766 + 0.3050i, 0.0587 + 0.2353i, 0.3253], \\ \hat{\mu}_{2E} &= \sigma[-0.3692 + 0.3386i, 0.3548 - 0.2641i, 0.1246 - 0.3320i, 0.0245 - 0.2562i, \\ &\quad -0.1792 - 0.0179i, -0.3016 + 0.3072i, 0.0438 + 0.2245i, 0.3022] \end{aligned}$$

The corresponding intrinsic mean shapes are denoted by $\hat{\mu}_{1I}$ and $\hat{\mu}_{2I}$. They are very close to the extrinsic means ($d_g(\hat{\mu}_{1E}, \hat{\mu}_{1I}) = 5.5395 \times 10^{-7}$, $d_g(\hat{\mu}_{2E}, \hat{\mu}_{2I}) = 1.9609 \times 10^{-6}$). Figure 1 shows the preshapes of the sample k -ads along with that of the extrinsic mean. The sample preshapes have been rotated appropriately so as to minimize the Euclidean distance from the mean preshape. Figure 2 shows the preshapes of the extrinsic means for the two samples along with that of the

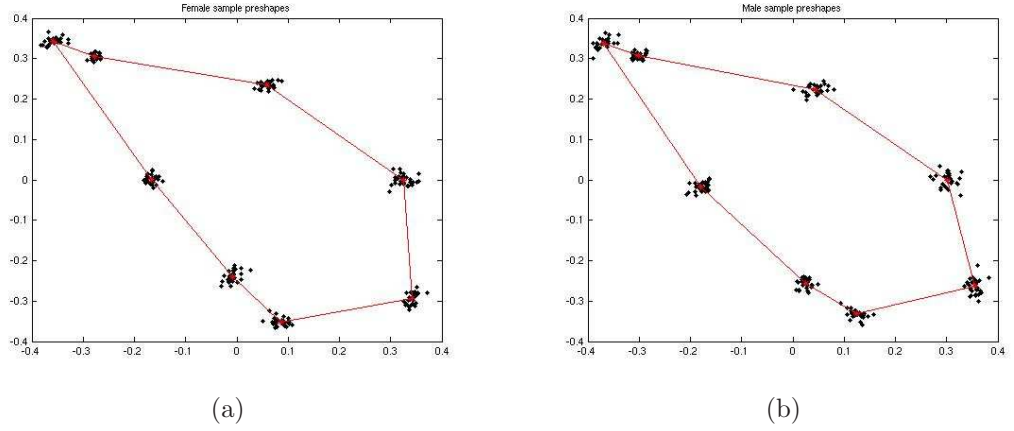


FIGURE 1. 1a and 1b show 8 landmarks from skulls of 30 female and 29 male gorillas, respectively, along with the mean shapes. * correspond to the mean shapes' landmarks.

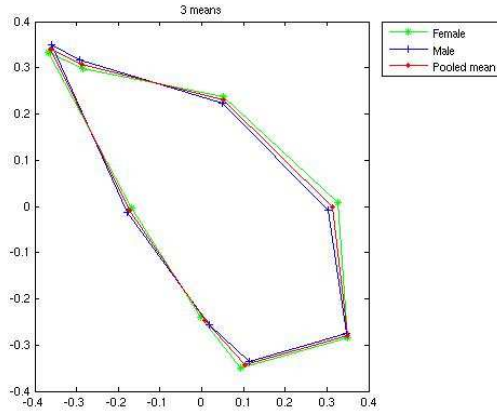


FIGURE 2. The sample extrinsic means for the 2 groups along with the pooled sample mean, corresponding to Figure 1.

pooled sample extrinsic mean. In BB (2008a), nonparametric two sample tests are performed to compare the mean shapes. The statistics (6.23) and (6.27) yield the following values:

$$\begin{aligned} \text{Extrinsic: } T_{nm} &= 392.6, \text{ p-value} = P(\chi_{12}^2 > 392.6) < 10^{-16}. \\ \text{Intrinsic: } T_{nm} &= 391.63, \text{ p-value} = P(\chi_{12}^2 > 391.63) < 10^{-16}. \end{aligned}$$

A parametric F-test (Dryden and Mardia (1998), pp. 154) yields $F = 26.47$, p-value = $P(F_{12,46} > 26.47) = 0.0001$. A parametric (Normal) model for Bookstein coordinates leads to the Hotelling's T^2 test (Dryden and Mardia (1998), pp. 170-172) yields the p-value 0.0001.

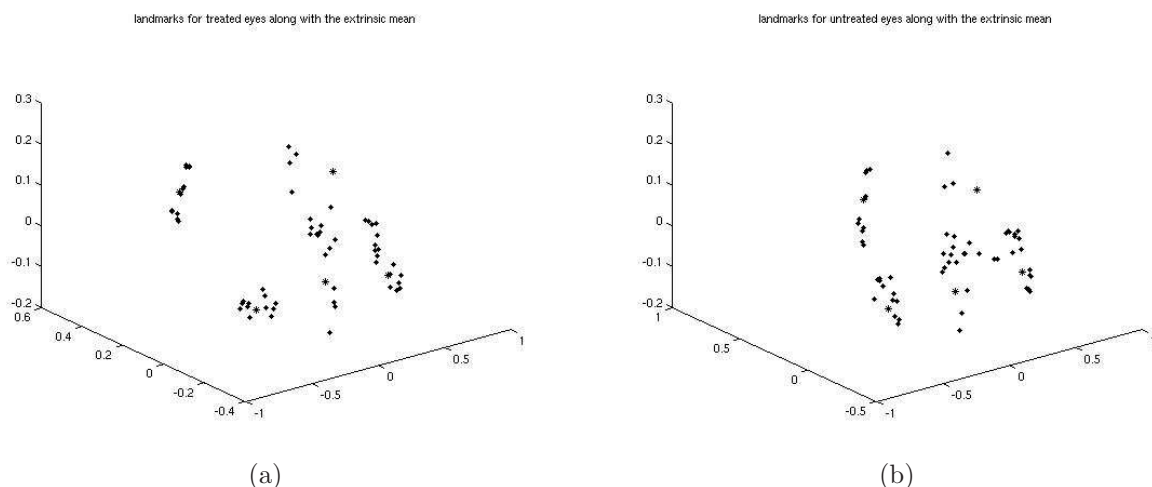


FIGURE 3. 3a and 3b show 5 landmarks from treated and untreated eyes of 12 monkeys, respectively, along with the mean shapes. * correspond to the mean shapes' landmarks.

7.2. Example 2: Schizophrenic Children. In this example from Bookstein (1991), 13 landmarks are recorded on a midsagittal two-dimensional slice from a Magnetic Resonance brain scan of each of 14 schizophrenic children and 14 normal children. In BB (2008a), nonparametric two sample tests are performed to compare the extrinsic and intrinsic mean shapes of the two samples. The values of the two-sample test statistics (6.23), (6.27), along with the p-values are as follows.

Extrinsic: $T_{nm} = 95.5476$, p-value = $P(\chi_{22}^2 > 95.5476) = 3.8 \times 10^{-11}$.

Intrinsic: $T_{nm} = 95.4587$, p-value = $P(\chi_{22}^2 > 95.4587) = 3.97 \times 10^{-11}$.

The value of the likelihood ratio test statistic, using the so-called *offset normal shape distribution* (Dryden and Mardia (1998), pp. 145-146) is $-2 \log \Lambda = 43.124$, p-value = $P(\chi_{22}^2 > 43.124) = 0.005$. The corresponding values of Goodall's F-statistic and Bookstein's Monte Carlo test (Dryden and Mardia (1998), pp. 145-146) are $F_{22,572} = 1.89$, p-value = $P(F_{22,572} > 1.89) = 0.01$. The p-value for Bookstein's test = 0.04.

7.3. Example 3: Glaucoma detection. To detect any shape change due to Glaucoma, 3D images of the Optic Nerve Head (ONH) of both eyes of 12 rhesus monkeys were collected. One of the eyes was treated while the other was left untreated. 5 landmarks were recorded on each eye and their reflection shape was considered in $R\Sigma_3^k$, $k = 5$. For details on landmark registration, see Derado et al. (2004). The landmark coordinates can be found in BP (2005). Figure 3 shows the preshapes of the sample k -ads along with that of the mean shapes. The sample points have been rotated and (or) reflected so as to minimize their Euclidean distance from the mean preshapes. Figure 4 shows the preshapes of the mean shapes for the two eyes along with that of the pooled sample mean shape. In Bandulasari et al. (2007), 4 landmarks are selected and the sample mean shapes of the two eyes are compared. Five local coordinates are used in the neighborhood of the mean to compute Bonferroni type Bootstrap Confidence Intervals for the difference between

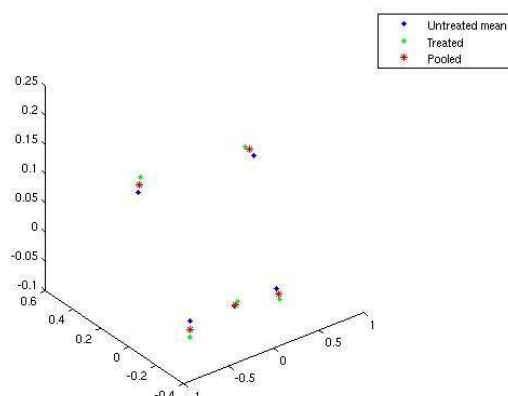


FIGURE 4. The sample means for the 2 eyes along with the pooled sample mean, corresponding to Figure 3.

the local reflection similarity shape coordinates of the paired glaucomatous versus control eye (see Section 6.1, Bandulasari et al. (2007) for details). It is found that the means are different at 1% level of significance.

Acknowledgement

The authors thank the referee for his suggestions.

References

- [1] BANDULASIRI, A., BHATTACHARYA, R. N. AND PATRANGENARU, V. (2007). Nonparametric Inference on Shape Manifolds with Applications in Medical Imaging. *To appear*.
- [2] BANDULASIRI A. AND PATRANGENARU, V. (2005). Algorithms for Nonparametric Inference on Shape Manifolds, *Proc. of JSM 2005, Minneapolis, MN* 1617-1622.
- [3] BERTHILSSON, R. AND HEYDEN A. (1999). Recognition of Planar Objects using the Density of Affine Shape. *Computer Vision and Image Understanding* **76** 2 135-145.
- [4] BERTHILSSON, RIKARD AND ASTROM, K. (1999). Extension of affine shape. *J. Math. Imaging Vision* **11**, no. 2 119-136.
- [5] BHATTACHARYA, A. AND BHATTACHARYA, R. (2008a). Nonparametric Statistics on Manifolds with Applications to Shape Spaces. *Pushing the Limits of Contemporary Statistics: Contributions in honor of J.K. Ghosh. IMS Lecture Series*. In Press.
- [6] BHATTACHARYA, A. AND BHATTACHARYA, R. (2008b). Statistics on Riemannian Manifolds: Asymptotic Distribution and Curvature. *Proc. Amer. Math. Soc.* In Press.
- [7] BHATTACHARYA, R. N. AND PATRANGENARU, V. (2002). Nonparametric estimation of location and dispersion on Riemannian manifolds. *J. Statist. Plann. Inference* **108** 23-35.
- [8] BHATTACHARYA, R. N. AND PATRANGENARU, V. (2003). Large sample theory of intrinsic and extrinsic sample means on manifolds-I. *Ann. Statist.* **31** 1-29.
- [9] BHATTACHARYA, R. AND PATRANGENARU, V. (2005). Large sample theory of intrinsic and extrinsic sample means on manifolds-II. *Ann. Statist.* **33** 1225-1259.
- [10] BIGOT, JÉRÉMIE (2006). Landmark-based registration of curves via the continuous wavelet transform. *J. Comput. Graph. Statist.* **15** 542-564.
- [11] BOOKSTEIN, F. (1978). *The Measurement of Biological Shape and Shape Change*. Lecture Notes in Biomathematics **24**. Springer, Berlin.
- [12] BOOKSTEIN, F. L. (1986). Size and shape spaces of landmark data (with discussion). *Statistical Science* **1** 181-242.

- [13] BOOKSTEIN, F.L. (1991). *Morphometric Tools for Landmark data: Geometry and Biology*. Cambridge Univ. Press.
- [14] BOOTHBY, W. (1986). *An Introduction to Differentiable Manifolds and Riemannian Geometry*. 2d ed. Academic Press, Orlando.
- [15] DERADO, G., MARDIA, K.V., PATRANGENARU, V. AND THOMPSON, H.W. (2004). A Shape Based Glaucoma Index for Tomographic Images. *J. Appl. Statist.* **31** 1241-1248.
- [16] DIMITRIĆ, IVKO (1996). A note on equivariant embeddings of Grassmannians. *Publ. Inst. Math. (Beograd) (N.S.)* **59**(73) 131-137.
- [17] DO CARMO, M. P. (1992). *Riemannian Geometry*. Birkhauser, Boston. English translation by F. Flaherty.
- [18] DRYDEN, I. L., LE, H. AND WOOD, A. (2007). The MDS model for shape. *To appear*.
- [19] DRYDEN, I. L. AND MARDIA, K. V. (1998). *Statistical Shape Analysis*. Wiley N.Y.
- [20] FISHER, N.I., HALL, P. , JING, B.-YI AND WOOD, A.T.A. (1996). Improved pivotal methods for constructing confidence regions with directional data. *J. Amer. Statist. Assoc.* **91** 1062-1070.
- [21] GALLOT, S., HULIN, D. AND LAFONTAINE, J. (1990). *Riemannian Geometry, 2nd ed.* Springer.
- [22] HENDRICKS, H. AND LANDSMAN, Z. (1998). Mean location and sample mean location on manifolds: Asymptotics, tests, confidence regions. *J. Multivariate Anal.* **67** 227-243.
- [23] KARCHAR, H. (1977). Riemannian center of mass & mollifier smoothing. *Comm. on Pure & Applied Math.* **30** 509-541.
- [24] KENDALL, D.G. (1977). The diffusion of shape. *Adv. Appl. Probab.* **9** 428-430.
- [25] KENDALL, D. G. (1984). Shape manifolds, Procrustean metrics, and complex projective spaces. *Bull. London Math. Soc.* **16** 81-121.
- [26] KENDALL, DAVID G. (1989) A survey of the statistical theory of shape. *Statist. Sci.* **4** 871-920.
- [27] KENDALL, D. G., BARDEN, D., CARNE, T. K. AND LE, H. (1999). *Shape & Shape Theory*. Wiley N.Y.
- [28] KENDALL, W. S. (1990). Probability, convexity, and harmonic maps with small image-I. Uniqueness and the fine existence. *Proc. London Math. Soc.* **61** 371-406.
- [29] KENT, J. T. (1992). New directions in shape analysis. *In The Art of Statistical Science: A Tribute to G. S. Watson* (K. V. Mardia, ed.) 115-128. Wiley, New York.
- [30] LE, H. (2001). Locating frechet means with application to shape spaces. *Adv. Appl. Prob.* **33** 324-338.
- [31] LEE, J.M. (1997). *Riemannian Manifolds: An Introduction to Curvature*. Springer, New York.
- [32] MARDIA, K.V. AND JUPP, P. E. (1999). *Statistics of Directional Data*. Wiley, New York.
- [33] MARDIA, K.V. AND PATRANGENARU, V. (2005). Directions and projective shapes. *Ann. Statist.* **33** 1666-1699.
- [34] RAMSAY, J.O. AND SILVERMAN, B.W. (2005). *Functional Data Analysis*. 2nd ed. Springer, New York.
- [35] SEPIASHVILI, D., MOURA, J.M.F. AND HA, V.H.S. (2003). Affine-Permutation Symmetry: Invariance and Shape Space (2003). *Proceedings of the 2003 Workshop on Statistical Signal Processing*. 293-296.
- [36] SMALL, C.G. (1996). *The Statistical Theory of Shape*. Springer, New York.
- [37] SPARR, G. (1992). Depth-computations from polyhedral images. *Proc. 2nd European Conf. on Computer Vision(ECCV-2)* (G. Sandini, editor). 378-386.
- [38] SUGATHADASA, M. S. (2006). *Affine and Projective Shape Analysis with Applications*. Ph.D. dissertation, Texas Tech University, Lubbock, TX.
- [39] WATSON, G. S. (1983). *Statistics on spheres*. University of Arkansas Lecture Notes in the Mathematical Sciences, 6. Wiley, New York.
- [40] XIA, MINGHUI AND LIU, BEDE (2004). Image registration by “super-curves”. *IEEE Trans. Image Process.* **13** 720-732.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ARIZONA, TUCSON, ARIZONA 85721

E-mail address: `rabi@math.arizona.edu`

URL: `http://math.arizona.edu/~rabi/`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ARIZONA, TUCSON, ARIZONA 85721

E-mail address: `abhishek@math.arizona.edu`

URL: `http://math.arizona.edu/~abhishek/`