STATISTICS ON RIEMANNIAN MANIFOLDS

WITH APPLICATIONS TO THE PLANER SHAPE SPACE

- Abhishek Bhattacharya

Project Advisor: Dr.Rabi Bhattacharya

Abstract

This article presents certain recent methodologies and some new results for the statistical analysis of distributions of shapes on manifolds. An important example considered in some detail here is the 2-D shape space of k-ads

The statistical analysis of shape distributions based on random samples is important in many areas such as morphometrics (discrimination and classification of biological shapes), medical diagnostics (detection of change or deformation of shapes in some organs due to some disease, for example) and machine vision (e.g., digital recording and analysis based on planar views of 3-D objects, when the position from which the object was viewed or pictured is unknown).

Frechet Mean on Metric Spaces

- Let $(\mathbf{M}, \mathbf{d})$ be a metric space and $\mathbf{Q}$ a probability measure on M. The **Frechet function** of Q is

$$F(p) = \int_M d^2(p, x) Q(dx), \quad p \in M.$$

- The **Frechet Mean** set of Q is the set of all p for which $F(p)$ is the minimum.

- Suppose every closed and bounded subset of M is compact. If the Frechet function is finite for some p, then the Frechet mean set is nonempty and compact.

Sample Frechet Mean

- Let $X_1, X_2, \ldots, X_n$ be an independent and identically distributed (iid) random sample with common distribution Q. The sample **Frechet Function** is

$$F_n(p) = \frac{1}{n} \sum_{i=1}^{n} d^2(X_i, p)$$

- The set of all minimizers of F_n is called the **sample Frechet mean** set.

Consistency of the Sample Frechet Mean

- **Note** A sequence of estimator θ_n defined on a probability space $(\Omega, P, \mathbb{B})$ is said to be a strongly consistent estimator of a parameter θ, if $\theta_n(\omega) \longrightarrow \theta$ as $n \to \infty$ for every ω outside of a P-null set.

- If every closed and bounded subset of M is compact and the Frechet mean of Q exists (as a unique minimizer of F), then every measurable selection from the Frechet sample mean set is a strongly consistent estimator of the Frechet mean of Q.

Asymptotic distribution of Frechet Sample mean

Suppose the following assumptions hold:

A1 Q has support in a single coordinate patch, (U, ϕ), $\phi : U \longrightarrow \mathbb{R}^d$ smooth. Call $Y_j = \phi(X_j)$, $j = 1, \ldots, n$.
A2 Frechet Mean μ_F of Q is unique.
A3 $\forall x$, $y \rightarrow h(x, y) = (d^\phi)^2(x, y) = d^2(\phi^{-1}x, \phi^{-1}y)$ is twice continuously differentiable in a neighborhood of $\phi(\mu_F) = \mu$.
A4 $E(D_r h(Y, \mu))^2 < \infty \ \forall r$.
A5 $E\{\sup_{|u-v|\leq\epsilon} |D_s D_r h(Y, v) - D_s D_r h(Y, u)|\} \rightarrow 0$ as $\epsilon \rightarrow 0 \ \forall \ r, s$.
A6 $\Lambda = ((\ E D_s D_r h(Y, \mu)\ ))$ is nonsingular.
A7 $\Sigma = \text{Cov } Dh(Y_1, \mu)$ is nonsingular.

Let $\mu_{F,n} =$ Frechet Sample mean, $\mu_n = \phi(\mu_{F,n})$. Then under the assumptions A1-A7,

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

Frechet Variation

- The **Frechet Variation**, V of Q is the minimum value attained by the Frechet function on M.

$$V = \int_M d^2(\mu, x) Q(dx),$$

$\mu \in$ The Frechet mean set of Q.

- The minimum value attained by the sample Frechet function, F_n is called the **sample Frechet Variation**, V_n.

$$V_n = \frac{1}{n} \sum_{i=1}^{n} d^2(X_i, \mu_n),$$

$\mu_n \in$ Sample Frechet mean set.

Consistency of the Sample Frechet Variation

- If every closed and bounded subset of M is compact, and F is finite on M, then the sample variation V_n is a strongly consistent estimator of the population variation, V.

- The sample variation is consistent even when the Frechet mean does not exist.

Asymptotic distribution of the Sample Frechet Variation

- Under assumptions A1-A7,

$$\sqrt{n}(V_n - V) \xrightarrow{\mathcal{L}} N(0, Vd^2(X_1, \mu_F))$$

- We need the Frechet mean to exist for the sample variation to be asymptotically normal.

Statistics on Riemannian manifolds

- Let $(\mathbf{M}, \mathbf{g})$ be a d-dimensional connected complete Riemannian manifold, g being the riemannian metric on M.

- We define "Extrinsic" and "Intrinsic" means and variations on M depending on the distance chosen on M.

Extrinsic Mean and Variation

- Let $\phi : M \rightarrow \mathbb{R}^k$ be an isometric map of M onto $\tilde{M} = \phi(M) \subset \mathbb{R}^k$.

- Define $d(x, y) = \|\phi(x) - \phi(y)\|$, where $\|.\|$ denotes Euclidean norm.

- Let Q be a probability measure on M with finite Frechet function. The Frechet mean (set) of Q is the **Extrinsic Mean (set)** of Q. The Frechet Variation is called the **Extrinsic Variation**.

- If $X_i (i \geq 1)$ are iid observations from Q, and $Q_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}$ is the corresponding empirical distribution, then the Frechet mean (set) of Q_n (wrt d) is called the **Extrinsic sample mean(set)**. The sample Frechet variation is called the **sample extrinsic variation**.

Existence and Consistency of Extrinsic Mean

- Let $\tilde{Q}, \tilde{Q}_n$ be the images of Q, Q_n respectively on $\mathbb{R}^k$: $\tilde{Q} = Q \circ \phi^{-1}$, $\tilde{Q}_n = Q_n \circ \phi^{-1}$.

- If $\mu = \int_{R^k} u \, \tilde{Q}(du)$, then the extrinsic mean set of Q is given by $\phi^{-1}(P_{\tilde{M}}\mu)$, where $P_{\tilde{M}}\mu$ is the set of points in $\tilde{M}$ whose distance from μ is the smallest among all points in $\tilde{M}$.

- If this set is a singleton, μ is called a **non focal point** of $\mathbb{R}^k$ (w.r.t. $\tilde{M}$); otherwise it is a **focal point**.

- Let $\bar{Y} = \frac{1}{n}\sum_{j=1}^{n} Y_j$ be the (sample) mean of $Y_j = \phi(X_j)$. The extrinsic sample mean set is $\phi^{-1}(P(\bar{Y}))$.

- If μ is a nonfocal point of $\mathbb{R}^k$, then any measurable selection from the extrinsic sample mean set is a strongly consistent estimator of the extrinsic mean $\mu_E = \phi^{-1}(P_{\tilde{M}}\mu)$.

Examples

Example1: Unit sphere S^{k-1}

- Consider the inclusion map $i : S^{k-1} \to \mathbb{R}^k$, $i(x) = x$.

- The extrinsic mean (set) of a probability measure Q on S^{k-1} is then the point(set) $P_{S^{k-1}}\tilde{\mu}$ on S^{k-1} closest to $\tilde{\mu} = \int_{\mathbb{R}^k} x\tilde{Q}(dx)$, where $\tilde{Q}$ is Q regarded as a probability measure on $\mathbb{R}^k$.

- $\tilde{\mu}$ is non-focal iff $\tilde{\mu} \neq 0$. Then $P_{S^{k-1}}\tilde{\mu} = \frac{\tilde{\mu}}{\|\tilde{\mu}\|}$, else $P_{S^{k-1}}(0) = S^{k-1}$.

- The extrinsic variation of Q is $2(1 - \|\tilde{\mu}\|)$.

Example2: Axial Space $\mathbb{R}P^{k-1}$

- Consists of all lines through the origin in $\mathbb{R}^k$.

- May be regarded as the quotient space of S^{k-1} under the equivalence relation $u \sim v$ iff $u = -v$. The elements of $\mathbb{R}P^{k-1}$ may be represented as $[u] = \{-u, u\}$ $(u \in S^{k-1})$.

- An embedding of $\mathbb{R}P^{k-1}$ is via the **Veronese-Whitney embedding** ϕ into the space of all $k \times k$ matrices identified with $\mathbb{R}^{k^2}$, $\phi([u]) = uu' = ((u_i u_j))_{1 \leq i,j \leq k}$, $u = (u_1, .., u_k)' \in S^{k-1}$.

- Define the extrinsic distance d on $\mathbb{R}P^{k-1}$ as $d^2([u],[v]) = \|uu' - vv'\|^2 = Trace(uu' - vv')^2$.

- Let Q be a probability measure on $\mathbb{R}P^{k-1}$, and $\tilde{\mu}$ be the mean of $\tilde{Q} = Q \circ \phi^{-1}$ considered as a probability measure on $\mathbb{R}^{k^2}$.

- Since $\tilde{\mu}$ is a mixture of elements of the form uu', $\tilde{\mu} \in S^+(k, \mathbb{R})$: the space of all symmetric nonnegative definite $k \times k$ matrices.

- $\tilde{\mu}$ is **nonfocal** iff its largest eigenvalue is **simple**, i.e., if the eigenspace corresponding to the largest eigenvalue is one dimensional.

- Then the extrinsic mean of Q is $[\mu]$, where μ is an eigen vector corresponding to the largest eigen value.

- The Extrinsic variation of Q is $2(1 - \lambda_k)$ where λ_k is the largest eigen value of $\tilde{\mu}$.

Example 3: Planer Shape Space of k-ads

- Consider a set of k points on the plane, not all points being the same. Assume $k > 2$ and refer to such a set as a k-ad (or a set of k landmarks).

- Denote a k-ad by a complex k-vector, $\mathbf{z} = (z_1, z_2, \ldots, z_k)$, $z_j = x_j + iy_j, 1 \leq j \leq k$.

- By the shape of a k-ad $\mathbf{z}$, we mean the equivalence class, or orbit of $\mathbf{z}$ under translation, scaling and rotation.

- Hence one may represent the shape of the k-ad as the complex line passing through $\mathbf{z} - \langle z \rangle$ where $\langle z \rangle = \frac{1}{k}\sum_{j=1}^{k} z_j$. Thus the space of k-ads is the set of all complex lines on the (complex(k-1)-dimensional) hyperplane, $H^{k-1} = \{w \in \mathbb{C}^k \setminus \{0\}\colon \sum_1^k w_j = 0\}$.

- So the shape space Σ_2^k has the structure of the complex projective space $\mathbb{C}P^{k-2}$.

- It is convenient to represent the element of Σ_2^k corresponding to a k-ad $\mathbf{z}$ by the curve $\pi(z) = [z] = \{e^{i\theta}\frac{(z-\langle z \rangle)}{\|z-\langle z \rangle\|}\colon 0 \leq \theta < 2\pi\}$ on the unit sphere in H^{k-1}.

- Can be embedded in $\mathbb{C}^{k^2}$ via the **Veronese-Whitney embedding** given by $\phi([z]) = uu^*$, where $u = \frac{(z - \langle \mathbf{z} \rangle)}{\|z - \langle \mathbf{z} \rangle\|}$ is called the preshape of the shape of z.

- Define the extrinsic distance d on Σ_2^k by that induced from this embedding, namely, $d^2([z],[w]) = \|uu^* - vv^*\|^2$ where u and v are the preshapes of z and w respectively.

- Here for arbitrary $k \times k$ complex matrices A, B; $\|A - B\|^2 = \sum_{j,j'} |a_{jj'} - b_{jj'}\|^2 = Trace(A - B)(A - B)^*$.

- Let Q be a probability measure on the shape space Σ_2^k, and let μ_0 denote the mean vector of $Q_0 \doteq Q \circ \phi^{-1}$, regarded as a probability measure on $\mathbb{C}^{k^2}$ (or, $\mathbb{R}^{2k^2}$). μ_0 belongs to the convex hull of $\phi(\Sigma_2^k)$ and in particular, is an element of H^{k-1}.

- The Extrinsic mean μ, say, of Q is unique iff the eigenspace for the largest eigenvalue of μ_0 is (complex) one dimensional, and then $\mu = [w]$, $w(\neq 0) \in$ the eigenspace of the largest eigenvalue of μ_0.

- The Extrinsic Variation of Q has the expression $V = 2(1 - \lambda_k)$ where λ_k is the largest eigen value of μ_0.

- The distance d on Σ_2^k may be expressed as $2(1-|u^*v|^2)$ and is called the **Full Procrustes distance**.

- Let $[z]$ and $[w]$ be two shapes and let u and v be their preshapes. Then the **Procrustes coordinates** of v onto u is defined as $v^P = (\widehat{a}+i\widehat{b})\mathbf{1}_k+\widehat{\beta}e^{i\widehat{\theta}}v$ where (β, θ, a, b) are chosen to minimized $D^2(u,v) = \|u - \beta e^{i\theta}v - (a + ib)\mathbf{1}_k\|^2$.

- Then $v^P = (v^*u)v$ and $D^2(u,v) = (1 - |v^*u|^2) = \frac{1}{2}d^2([z],[w])$.

- As a numerical example, consider 8 locations on a gorilla skull projected on a plane. There are 30 female and 29 male samples.

- Figures 1 and 2 are the plot of the Procrustes coordinates of the female and male samples onto their extrinsic sample means respectively.

Procrustes coordinates onto Extrinsic Mean, * denotes coordinates of the mean shape

Figure 1: Female landmarks

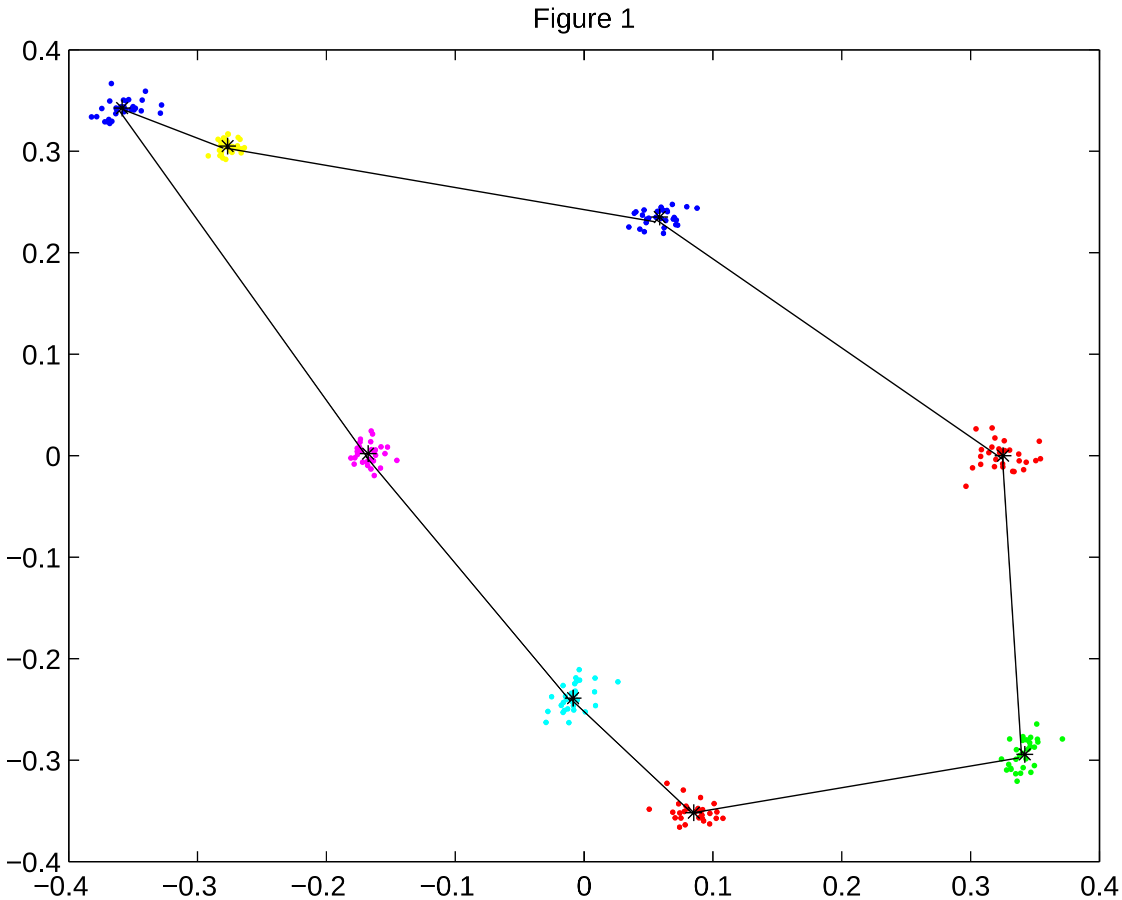
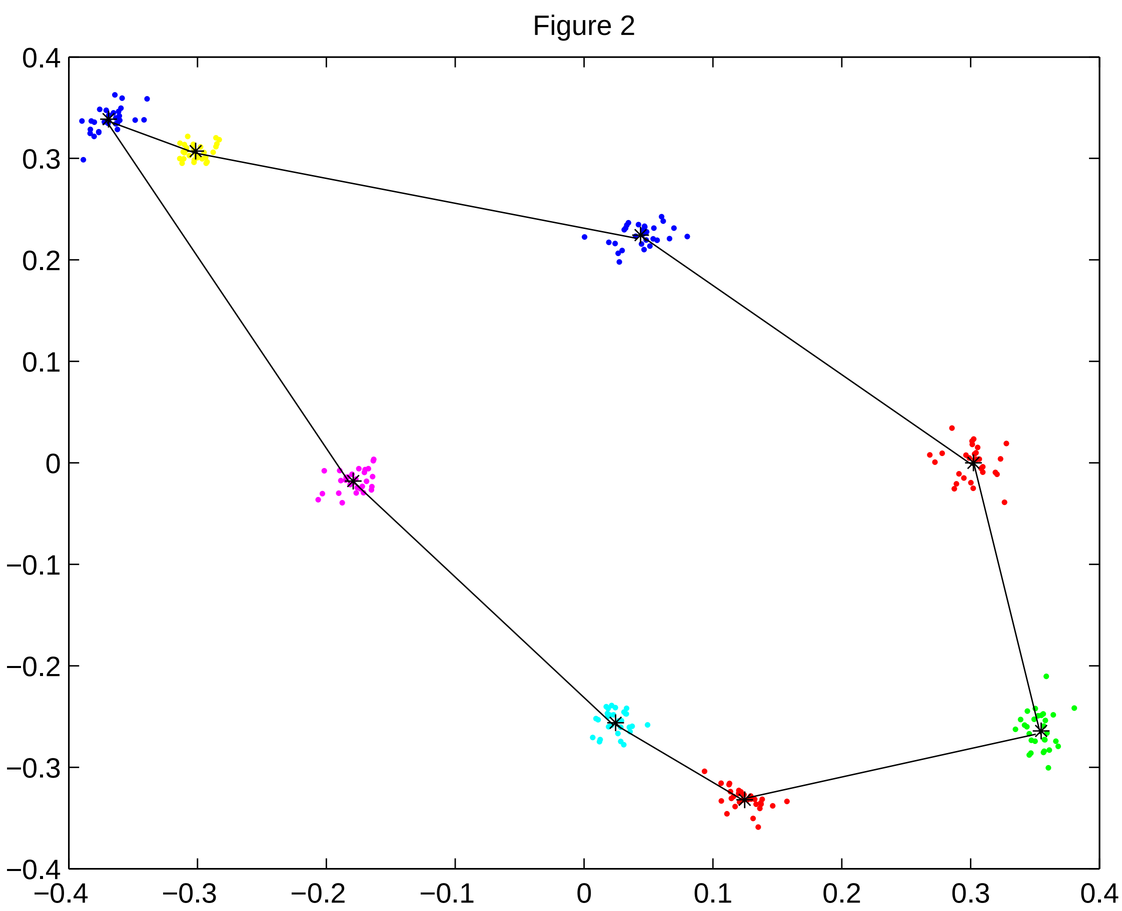


Figure 2: Male landmarks



Asymptotic Distribution of the sample extrinsic mean

- Let ϕ is an embedding of M into $\mathbb{R}^k$. Suppose the mean μ of the image $\tilde{Q} = Q \circ \phi^{-1}$ is a non-focal point of $\mathbb{R}^k$, so that the projection $P(\mu)$ of μ on $\phi(M)$ is unique, and the extrinsic mean of Q is $\mu_E = \phi^{-1}P(\mu)$.

- Let $\bar{Y} = \frac{1}{n}\sum_{j=1}^{n} Y_j$ denote the sample mean of $Y_j = \phi(X_j)$, where $X_1, \ldots, X_n$ is a random sample from Q. The extrinsic sample mean set is $\phi^{-1}(P(\bar{Y}))$.

- In a neighborhood of a nonfocal point such as μ, $P(.)$ is smooth. Write

$$\begin{aligned}\sqrt{n}[P(\bar{Y}) - P(\mu)] &= \sqrt{n}(d_\mu P)(\bar{Y} - \mu) + o_P(1) \\ &= (d_\mu P)(\sqrt{n}(\bar{Y} - \mu)) + o_P(1)\end{aligned}$$

where $d_\mu P$ is the differential (map) of the projection $P(.)$, which takes vectors in the tangent space of $\mathbb{R}^k$ at μ to tangent vectors of $\phi(M)$ at $P(\mu)$.

- Let $f_1, f_2, \ldots, f_d$ be an orthonormal basis of $T_{P(\mu)}\phi(M)$ and $e_1, e_2, \ldots, e_k$ be an orthonormal basis (frame) of $\mathbb{R}^k$. One has

$$\begin{aligned}
\sqrt{n}(\bar{Y} - \mu) &= \sum_{j=1}^{k} \sqrt{n}(\bar{Y} - \mu)_j e_j, \\
d_\mu P(\sqrt{n}(\bar{Y} - \mu)) &= \sum_{j=1}^{k} \sqrt{n}(\bar{Y} - \mu)_j d_\mu P(e_j) \\
&= \sum_{j=1}^{k} \sqrt{n}(\bar{Y} - \mu)_j \sum_{r=1}^{d} a_{jr}(\mu) f_r \\
&= \sum_{r=1}^{d} [\sum_{j=1}^{k} a_{jr}(\mu)\sqrt{n}(\bar{Y} - \mu)_j] f_r
\end{aligned}$$

where $a_{jr}(\mu) = \langle d_\mu P(e_j), f_r \rangle$.

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

  $\Sigma = A'VA \;\; [A = A(\mu) = ((a_{jr}(\mu)))_{1 \leq j \leq k, 1 \leq r \leq d}]$

  V being the $k \times k$ covariance matrix of $\tilde{Q}$ (wrt the basis $\{e_j : 1 \leq j \leq k\}$).

- In matrix notation,

$$\sqrt{n}A'(\bar{Y} - \mu) \xrightarrow{\mathcal{L}} N(0, \Sigma) \text{ as } n \to \infty$$

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

$$n(\bar{Y} - \mu)' A \Sigma^{-1} A' (\bar{Y} - \mu) \longrightarrow \mathcal{X}_d^2 \text{ as } n \to \infty$$

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

$$\{P(\mu) : n(\bar{Y} - \mu)' A \Sigma^{-1} A' (\bar{Y} - \mu) \leq \mathcal{X}_d^2 (1 - \alpha)\}$$

Asymptotic distribution of the mean shape

- Let us apply the above result to find the asymptotic distribution of the sample extrinsic mean shape of a sample of size n from the planer shape space.

- $M = \Sigma_2^k$ can be embedded into $S(k, \mathbb{C})$, the space of all $k \times k$ complex self adjoint matrices, via the Veroneese-Whitney embedding, ϕ.

- Consider $S(k, \mathbb{C})$ as a linear subspace of $\mathbb{C}^{k^2}$ (over $\mathbb{R}$) and as such a regular submanifold of $\mathbb{C}^{k^2}$ embedded by the inclusion map, and inheriting the metric $\langle A, B \rangle = Re\ Trace(A\bar{B}')$.

- The dimension of $S(k,\mathbb{C})$ is k^2. An orthonormal basis for $S(k,\mathbb{C})$ is given by $\{v_b^a : 1 \le a \le b \le k\}$ and $\{w_b^a : 1 \le a < b \le k\}$:

$$\begin{aligned} v_b^a &= \frac{1}{\sqrt{2}}(e_a e_b^t + e_b e_a^t), \;\; a < b \\ &= e_a e_a^t, \;\; a = b \\ w_b^a &= \frac{i}{\sqrt{2}}(e_a e_b^t - e_b e_a^t), \;\; a < b. \end{aligned}$$

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

- For $U \in O(k)$ $(UU^* = U^*U = I)$, $\{U v_b^a U^* : 1 \le a \le b \le k\}$, $\{U w_b^a U^* : 1 \le a < b \le k\}$ is also an orthogonal frame for $S(k,\mathbb{C})$.

- Assume that the mean μ of $\tilde{Q} = Q \circ \phi^{-1}$ has its largest eigen value simple. Then

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

  Here we view $d_\mu P : S(k,\mathbb{C}) \to T_{P(\mu)}\phi(\Sigma_2^k)$.

- Choose $U \in O(k)$ such that $U^*\mu U = D \equiv Diag(\lambda_1, \ldots, \lambda_k)$, $\lambda_1 \leq \ldots \leq \lambda_{k-1} < \lambda_k$ being the eigenvalues of μ.

- Choose the basis $\{Uv_b^aU^*, Uw_b^aU^*\}$ for $S(k, \mathbb{C})$. Then $d_\mu P(Uv_b^aU^*) =$

$$\begin{cases} 0 & \text{if } 1 \leq a \leq b < k,\ a = b = k \\ (\lambda_k - \lambda_a)^{-1} U v_k^a U^* & \text{if } 1 \leq a < k,\ b = k. \end{cases}$$

and $d_\mu P(Uw_b^aU^*) =$

$$\begin{cases} 0 & \text{if } 1 \leq a < b < k \\ (\lambda_k - \lambda_a)^{-1} U w_k^a U^* & \text{if } 1 \leq a < k,\ b = k. \end{cases}$$

- Write $\sqrt{n}(\bar{Y}-\mu) =$

$$\sum\sum_{1\leq a\leq b\leq k} < \sqrt{n}(\bar{Y}-\mu), Uv_b^a U^* > Uv_b^a U^*$$
$$+ \sum\sum_{1\leq a<b\leq k} < \sqrt{n}(\bar{Y}-\mu), Uw_b^a U^* > Uw_b^a U^*$$

- Then $d_\mu P(\sqrt{n}(\bar{Y}-\mu)) =$

$$\sum_{a=2}^{k-1} < \sqrt{n}(\bar{Y}-\mu), Uv_k^a U^* > (\lambda_k-\lambda_a)^{-1} Uv_k^a U^*$$
$$+ \sum_{a=2}^{k-1} < \sqrt{n}(\bar{Y}-\mu), Uw_k^a U^* > (\lambda_k-\lambda_a)^{-1} Uw_k^a U^*$$

- So $\sqrt{n}(P(\bar{Y}) - P(\mu))$ has an asymptotic Gaussian distribution on a subspace of $S(k, \mathbb{C})$ with asymptotic coordinates $T_n(\mu) =$

  $(\langle\sqrt{n}(\bar{Y}-\mu), Uv_k^a U^*\rangle_{a=2}^{k-1}, \langle\sqrt{n}(\bar{Y}-\mu), Uw_k^a U^*\rangle_{a=2}^{k-1})$

  wrt the basis vector

  $\{(\lambda_k - \lambda_a)^{-1} Uv_k^a U^*, (\lambda_k - \lambda_a)^{-1} Uw_k^a U^*\}_{a=2}^{k-1}$

- $T_n(\mu)'\Sigma(\mu)^{-1}T_n(\mu) \longrightarrow \mathcal{X}^2_{2k-4}$ as $n \to \infty$.

- This result can be used to construct asymptotic confidence set of $P(\mu)$ as in general M.

A two sample testing problem

- Let Q_1 and Q_2 be two probability measures on Σ_2^k, and let μ_1 and μ_2 denote the mean vectors of $Q_1 \circ \phi^{-1}$ and $Q_2 \circ \phi^{-1}$ respectively. Suppose $[x_1], \ldots, [x_n]$ and $[y_1], \ldots, [y_m]$ are iid random samples from Q_1 and Q_2 respectively. Let $X_i = \phi([x_i])$, $Y_i = \phi([y_i])$.

- We are to test if the extrinsic means of Q_1 and Q_2 are equal, i.e.

$$H_0 : P\mu_1 = P\mu_2$$

- Assume that both μ_1 and μ_2 have simple largest eigen values. Then under H_0, the corresponding eigen vectors differ by a rotation.

- Choose $\mu \in S(k, \mathbb{C})$ with same projection as μ_1 and μ_2.

- Suppose $\mu = U\Lambda U^*$, where $\Lambda = Diag(\lambda_1 \leq \lambda_2 \leq \ldots < \lambda_k)$ are its eigen values and $U = [U_1, U_2, \ldots, U_k]$ are the corresponding eigen vectors. Under H_0, $P\mu_1 = P\mu_2 = U_k U_k^*$.

$$
\begin{aligned}
d_\mu P(\bar{X} - \mu) &= \sum_{a=2}^{k-1} (\lambda_k - \lambda_a)^{-1} (U_a^* \bar{X} U_k) U_a U_k^* \\
&\quad + \sum_{a=2}^{k-1} (\lambda_k - \lambda_a)^{-1} (U_k^* \bar{X} U_a) U_k U_a^* \\
d_\mu P(\bar{Y} - \mu) &= \sum_{a=2}^{k-1} (\lambda_k - \lambda_a)^{-1} (U_a^* \bar{Y} U_k) U_a U_k^* \\
&\quad + \sum_{a=2}^{k-1} (\lambda_k - \lambda_a)^{-1} (U_k^* \bar{Y} U_a) U_k U_a^*
\end{aligned}
$$

- Let $T_n(\mu) = (Re(U_a^* \bar{X} U_k),\ Im(U_a^* \bar{X} U_k))_{a=2}^{k-1}$ and $S_m(\mu) = (Re(U_a^* \bar{Y} U_k),\ Im(U_a^* \bar{Y} U_k))_{a=2}^{k-1}$.

- Under H_0, $T_n(\mu)$ and $S_m(\mu)$ have mean zero, and as $n, m \to \infty$, $\sqrt{n}T_n(\mu) \xrightarrow{\mathcal{L}} N(0, \Sigma_1(\mu))$ and $\sqrt{m}S_m(\mu) \xrightarrow{\mathcal{L}} N(0, \Sigma_2(\mu))$.

- Suppose $\frac{n}{m+n} \to p$, $\frac{m}{m+n} \to q$, for some $p, q > 0$; $p + q = 1$. Then $(n+m)(T_n(\mu) - S_m(\mu))'(\frac{1}{p}\Sigma_1(\mu) + \frac{1}{q}\Sigma_2(\mu))^{-1}(T_n(\mu) - S_m(\mu)) \xrightarrow{\mathcal{L}} \mathcal{X}^2_{2k-4}$.

- We can choose μ to be any positive linear combination of μ_1 and μ_2. We may take $\mu = p\mu_1 + q\mu_2$.

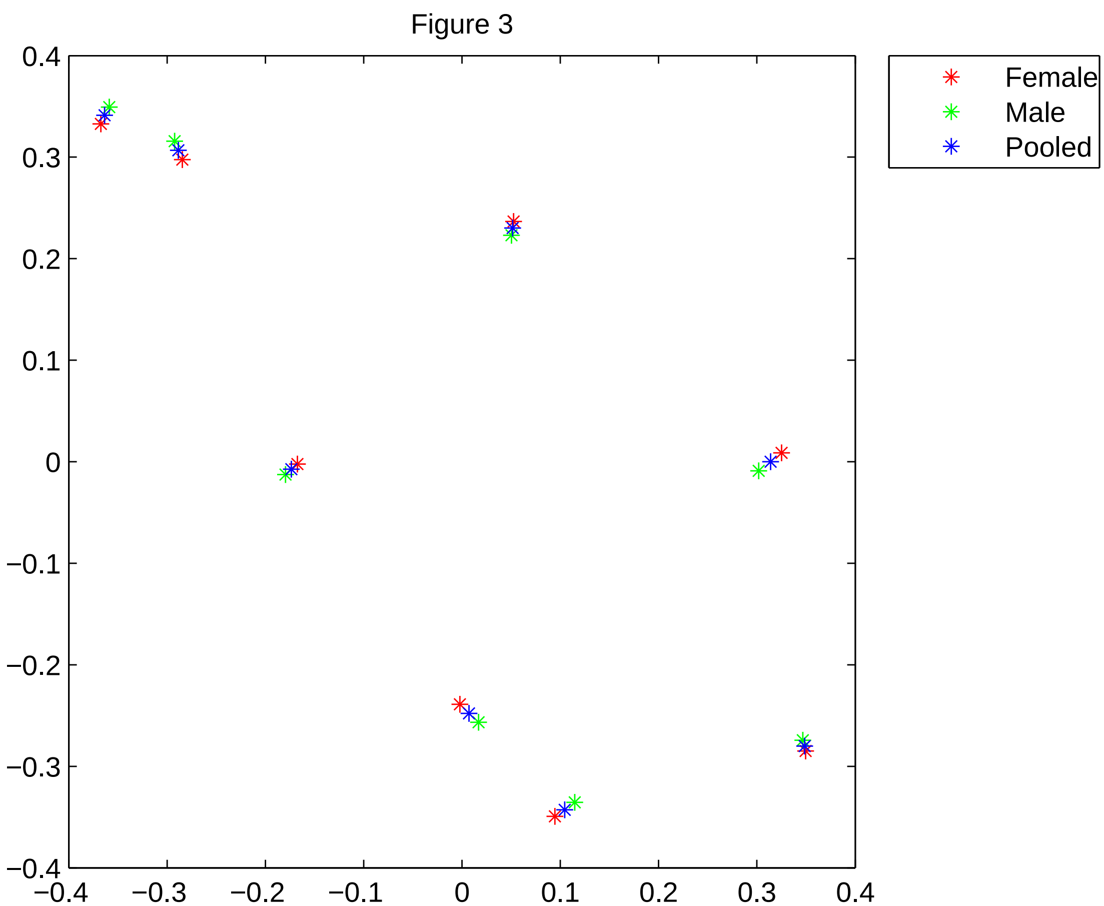
- In practice μ_1 and μ_2 are unknown, so is μ. Then we may estimate μ by the pooled sample mean, $\widehat{\mu} = \frac{n\bar{X}+m\bar{Y}}{m+n}$; $\Sigma_1(\mu)$ and $\Sigma_2(\mu)$ by their sample estimates $\widehat{\Sigma}_1(\widehat{\mu})$ and $\widehat{\Sigma}_2(\widehat{\mu})$.

- Then the two sample test statistic becomes $(\bar{X}(\widehat{\mu})-\bar{Y}(\widehat{\mu}))'(\frac{1}{n}\widehat{\Sigma}_1(\widehat{\mu})+\frac{1}{m}\widehat{\Sigma}_2(\widehat{\mu}))^{-1}(\bar{X}(\widehat{\mu})-\bar{Y}(\widehat{\mu}))$

- For the skull data, suppose we are to test if the male and female populations have the same mean shape.

- Figure 3 is the plot of the full Procrustes coordinates for the (sample) Extrinsic mean shapes of female and male skulls onto the Extrinsic mean for the pooled sample.

Figure 3: Procrustes Coordinates of sample means



- Value of the two sample test statistic is 392.68.

- P-value for the test using chi square approximation is 0.

- So we reject H_0 and conclude that the mean shapes are different.

Asymptotic distribution of Extrinsic Variation

- Suppose V is the Extrinsic variation of Q, and V_n its sample analogue. Then

$$\sqrt{n}(V_n - V) \xrightarrow{\mathcal{L}} N(0, \sigma^2),$$
$$\sigma^2 = \int_M (d^2(x, \mu_E) - V)^2 Q(dx)$$

- An asymptotic level α confidence interval for V is given by:

$$(V_n - \frac{s}{\sqrt{n}} Z_{1-\frac{\alpha}{2}}, V_n + \frac{s}{\sqrt{n}} Z_{1-\frac{\alpha}{2}})$$

where $s^2 = \frac{1}{n}\sum_{j=1}^{n}(d^2(X_j, \mu_{nE}) - V_n)^2$ is the sample estimate of σ^2 and $Z_{1-\frac{\alpha}{2}}$ is the upper $(1 - \frac{\alpha}{2})$ quantile for standard normal distribution.

- For the gorilla skull data, 95% C.I. for the variations of females and males are:

- Females: $(0.0031, 0.0046)$

- Males: $(0.0034, 0.0056)$

Testing equality of Extrinsic Variations

- To test if two populations have same spread V, use the statistic

$$T_{nm} = \frac{(V_{1n} - V_{2m})}{\sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{m}}}$$

, where V_{1n} and V_{2m} are the sample spreads;

$$s_1^2 = \frac{1}{n} \sum_{j=1}^{n} (d^2([x_j], [\mu_{nE}]) - V_{1n})^2$$

$$s_2^2 = \frac{1}{m} \sum_{j=1}^{m} (d^2([y_j], [\mu_{mE}]) - V_{2m})^2$$

- Under H_0, $T_{nm} \xrightarrow{\mathcal{L}} N(0, 1)$.

- For the shape space, $T_{nm} = 2\frac{(\lambda_{km}-\lambda_{kn})}{\sqrt{\frac{s_1^2}{n}+\frac{s_2^2}{m}}}$ where λ_{kn} and λ_{km} are the largest eigen values of $\bar{X}$ and $\bar{Y}$ respectively.

- For the skull data, $T_{nm} = -0.923$.

- P-value of the test using normal approximation is 0.356.

- So we conclude that the two populations have same average spread around their respective mean shapes.

Intrinsic Mean and Variation

- Let the distance, $d = d_g$, the geodesic distance under g.

- Let Q be a probability distribution with finite Frechet function. The Frechet mean set of Q is called its **Intrinsic Mean** set. The Frechet Variation of Q under d_g is called its **Intrinsic Variation**.

- Let $X_1, X_2, \ldots, X_n$ be iid with common distribution Q. The sample Frechet mean (set) is called the **Sample Intrinsic mean** (set) and the sample Frechet Variation is called the **Sample Intrinsic variation**.

Existence (Uniqueness) of Intrinsic mean

Suppose all sectional curvatures on M are bounded above by some $C \geq 0$. Suppose the support of Q is contained in a convex ball of radius r, $B(p, r)$ (wrt d_g) where

$$r = \begin{cases} \infty & \text{if } C = 0 \\ \frac{\pi}{4\sqrt{C}} & \text{if } C > 0 \end{cases}$$

Then the Frechet function, F of Q is strictly convex on $B(p, r)$, has a unique global minima, which is attained in $B(p, r)$ and is the unique local minima of F on $B(p, 2r)$. Hence the intrinsic mean of Q exists (as a unique minimizer of F) and lies in $B(p, r)$.

Asymptotic distribution of sample Intrinsic Mean

- Suppose the support of Q is contained in a convex geodesic ball $B(p,r)$ with center p and radius r which is disjoint from the cutlocus of p, $C(p)$.

- Let $\phi = Exp_p^{-1} : B(p,r) \longrightarrow T_pM(\approx \mathbb{R}^d)$.

- Define $h(x,y) = d_g^2(\phi^{-1}x, \phi^{-1}y);\ x, y \in \mathbb{R}^d$.

- Let $((D_r h))_{r=1}^d$ and $((D_r D_s h))_{r,s=1}^d$ be the matrix of first and second order derivatives of $y \mapsto h(x,y)$.

- Let $Y_j = \phi(X_j); j = 1, \ldots, n$; $X_1, \ldots, X_n$ being iid observations from Q. Let $\mu = \phi(\mu_I)$, μ_I being the intrinsic mean of Q. Let $\mu_n = \phi(\mu_{n,I})$, $\mu_{n,I}$ being a measurable selection from the sample mean set of X_j's. Define $\Lambda = E((D_r D_s h(Y_1, \mu)))_{r,s=1}^d$; $\Sigma = Cov((D_r h(Y_1, \mu)))_{r=1}^d$.

- Then Λ and Σ are positive definite and

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

Expressions for Λ and Σ

- Suppose Q has an intrinsic mean μ_I and satisfies the following assumption:
  A. For any geodesic γ, $\gamma(0) = \mu_I$; there exists $s(\mu_I) > 0$ such that the cut-locus of $\gamma\,|_{[0,s(\mu_I)]}$, $C(\gamma\,|_{[0,s(\mu_I)]})$ has Q measure 0.

- This is satisfied if the support of Q is contained in a convex open ball of radius r.

- Let $Y_j = exp_{\mu_I}^{-1} X_j = (Y_j^1, \ldots, Y_j^d)$ be the normal coordinates of the sample around μ_I. Then we have the following expressions:

  1. $D_r h(Y_j, \mu) = -2Y_j^r$
  2. $E(D_r h(Y_1, \mu)) = 0$
  3. $\Sigma_{rs} = 4Cov(Y_1^r, Y_1^s)$

If M has constant sectional curvature C, then

4. $\Lambda_{rs} = 2E((\left(\frac{1 - f|Y_1|}{|Y_1|^2}\right) Y_1^r Y_1^s + (f|Y_1|)\delta_{rs}))$,

$|Y_1| = \sqrt{(Y_1^1)^2 + (Y_1^2)^2 + \ldots (Y_1^d)^2}$

where

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

Asymptotic Confidence set for μ_I

- Consider the statistic $T_n = d_g^2(\mu_{nI}, \mu_I)$.

- If $\phi = Exp_{\mu_I}^{-1}$, $T_n = \|\mu_n\|^2$.

- So $nT_n \xrightarrow{\mathcal{L}} \sum_{i=1}^d \lambda_i Z_i^2$ where $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_d$ are the eigen values of $\Lambda^{-1}\Sigma\Lambda^{-1}$ and $Z_1, \ldots, Z_d$ iid $N(0,1)$.

- So an asymptotic level $(1-\alpha)$ confidence set for μ_I is given by:

$$\{\mu_I : nd_g^2(\mu_{nI}, \mu_I) \leq \hat{c}_{1-\alpha}\}$$

  where $\hat{c}_{1-\alpha}$ is the estimated upper $(1-\alpha)$ quantile of the distribution of $\sum_{i=1}^{d} \hat{\lambda}_i Z_i^2$, $\hat{\lambda}_i$ being the eigen values estimated from the sample $X_1, \ldots, X_n$ and $(Z_1, Z_2, \ldots)$ is a sample of iid $N(0,1)$ independent of the sample $X_1, \ldots, X_n$.

- The corresponding bootstrapped confidence region is given by

$$\{\mu_I : nd_g^2(\mu_{nI}, \mu_I) \leq c^*_{1-\alpha}\}$$

  where $c^*_{(1-\alpha)}$ is the upper $(1-\alpha)$-quantile of the bootstrapped values of the statistic nT_n.

Examples

Example 1: Unit sphere S^d

- At each $p \in S^d$, and $v_1, v_2 \in T_pS^d = \{v \in \mathbb{R}^{d+1} : v.p = 0\}$, $g(v_1, v_2) = v_1.v_2$, where . is the eucledian dot product.

- The geodesics are the big circles,

$$\gamma_{p,v}(t) = (\cos t)p + (\sin t)\frac{v}{\|v\|}, -\pi < t \leq \pi$$

- The exponential map, $Exp : T_p \to S^d$ is

$$Exp_p(0) = p,$$
$$Exp_p(v) = \cos(\|v\|)p + \sin(\|v\|)\frac{v}{\|v\|} \quad (v \in T_p)$$

- The cutlocus of p is $C(p) = \{-p\}$.

- The inverse of the Exponential map on $S^d \setminus \{-p\}$ into T_p is

  $$Exp_p^{-1}(q) = \frac{arccos(p.q)}{\sqrt{1-(p.q)^2}}[q-(p.q)p] \quad (q \neq p, -p),$$
  $$Exp_p^{-1}(p) = 0$$

- The geodesic distance d_g is

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

- This space has constant sectional curvature 1.

- So if Q is a probability measure on S^d, Q has an intrinsic mean if its support is contained in a geodesic ball of radius at most $\pi/4$. Then the sample Intrinsic mean (i.e., any measurable selection from the sample intrinsic mean set) based on a random sample from Q is consistent.

- In case Q has a mean μ_I, pick an orthonormal basis for $T_{\mu_I}S^d$: $\{v_1, \ldots, v_d\}$. For $x \in S^{k-1}$, $|x.\mu_I| < 1$, its normal coordinates around μ_I is

$$\phi(x) = exp_{\mu_I}^{-1}(x) = \frac{arccos(x.\mu_I)}{\sqrt{1-(x.\mu_I)^2}}[x - (x.\mu_I)\mu_I]$$

- Let $\mathbf{y} = (y^1, \ldots, y^d) = y^r v_r$ denote the normal coordinates of x. Then

$$y^r = \frac{arccos(x.\mu_I)}{\sqrt{1-(x.\mu_I)^2}} x.v_r \quad r = 1, 2, \ldots, d.$$
$$|y| = arccos(x.\mu_I) = d_g(x, \mu_I)$$

- Q satisfies assumption (**A**), if all one dimensional curves have measure 0. This is true in particular if Q is absolutely continuous with respect to the volume measure.

Example 2: Σ_2^k

- Consider the complex projective space $\mathbb{C}P^d$: the space of all complex lines through the origin in $\mathbb{C}^{d+1}$.

- Consider the map

$$\begin{aligned} \pi : \mathbb{C}S^d &\to \mathbb{C}P^d \\ z &\mapsto \pi(z) = [z]; \\ z &\in \mathbb{C}^{d+1},\ \|z\| = \sum_{j=1}^{d+1} |z_j|^2 = 1 \end{aligned}$$

  This π is a Riemannian submersion.

- The tangent space $T_z\mathbb{C}S^d$ at $\mathbf{z}$ is the set of all vectors, $\mathbf{v}$ in $\mathbb{C}^{d+1}$ orthogonal to $\mathbf{z}$. Here for $\mathbf{v}, \mathbf{w} \in \mathbb{C}^{d+1}$, $< \mathbf{v}, \mathbf{w} > = \mathrm{Re}(\mathbf{v}'\bar{\mathbf{w}})$.

- For any $[z] \in \mathbb{C}P^d$, the tangent space $T_{[z]}\mathbb{C}P^d$ at $[z]$ is isomorphic with a subspace called the horizontal subspace of $T_z\mathbb{C}S^d$,

$$H_z = \{v \in \mathbb{C}^{d+1} \ : z^{'}\bar{v} = 0\}$$

- $exp_{[z]} = \pi \circ exp_z \circ d\pi_z^{-1}$, and if $z'\bar{w} \neq 0$,

$$d\pi_z^{-1}(exp_{[z]}^{-1}([w]) = \frac{r}{sinr}\{-zcosr + e^{i\theta}w\}$$

where $r = d_g([z],[w]) = \ arccos(|z^{'}\bar{w}|) \in \ [0, \frac{\pi}{2})$

and $e^{i\theta} = \frac{z^{'}\bar{w}}{|z^{'}\bar{w}|}$

- Σ_2^k may be identified with the set of all complex lines in H^{k-1}, so one may express the geodesics, geodesic distances, the exponential map and its inverse in Σ_2^k by simply taking $d = k-1$ above and replacing the k-ads $\mathbf{z}, \mathbf{w}$ by their preshapes.

- Σ_2^k has all sectional curvatures bounded between 1 and 4.

- The cut-locus of $[p]$ is

$$C([p]) = \{[x] : d_g([x],[p]) = \frac{\pi}{2}\}$$
$$= \{[x] : p^{'}\bar{x} = 0\}$$

- Q has an intrinsic mean, if its support is contained in a geodesic ball of radius at most $\frac{\pi}{8}$.

- In case Q has intrinsic mean $[\mu]$, consider the coordinate patch $\phi = d\pi_\mu^{-1}(exp_{[\mu]}^{-1}[z])$ around $[\mu]$, that is

$$\phi : \Sigma_2^k \setminus C([\mu]) \to H_\mu$$
$$\phi([z]) = \frac{r(z)}{sinr(z)}[-\mu cosr(z) + e^{i\theta(z)}z]$$

where $z, \mu \in \mathbb{C}S^{k-1} \cap H^{k-1}$, $r(z) = d_g([z], [\mu])$ $= arccos(|\mu^{'}\bar{z}|) \in [0, \frac{\pi}{2})$ and $e^{i\theta(z)} = \frac{\mu^{'}\bar{z}}{|\mu^{'}\bar{z}|}$.

- Q satisfies assumption (**A**) if it is absolutely continuous wrt the volume measure on Σ_2^k.

- If $\{[X_1], [X_2], \ldots, [X_j]\}$ is an iid sample from Q, X_j's being the pre-shapes and $Y_j = \phi(X_j)$, then

$$\begin{aligned} Dh(Y_1, \mu) &= -2Y_1 \\ &= \frac{-2r_1}{sinr_1}[-cosr_1\mu + e^{i\theta_1}X_1] \\ \text{where } r_1 &= arccos(|\mu^{'}\bar{X}_1|) \\ \text{and } e^{i\theta_1} &= \frac{\mu^{'}\bar{X}_1}{cosr_1}. \\ E(Dh(Y_1, \mu)) &= 0 \end{aligned}$$

- To find μ, we need to find the zeros of the function: $\mu \mapsto E(Dh(Y_1, \mu))$. This is equivalent to finding the fixed points of the map

$$\begin{aligned} f : \pi^{-1}\Sigma_2^k &\to \mathbb{C}S^{k-1} \\ \mu &\mapsto exp_\mu(-E(Dh(Y_1, \mu))) \end{aligned}$$

- Here exp_μ is the exponential map on $\mathbb{C}S^{k-1}$,

$$exp_\mu : \ T_\mu \mathbb{C}S^{k-1} \to \mathbb{C}S^{k-1}$$
$$exp_\mu(v) = (cos|v|)\mu + \frac{sin|v|}{|v|} v.$$

- So $f(\mu) = (cos|v|)\mu + \frac{sin|v|}{|v|} v$ where $v = 2E \frac{r_1}{sin r_1}(-cos r_1 \mu + e^{i\theta_1} X_1)$.

- If the support of Q is contained in a geodesic ball of radius at most $\frac{3\pi}{40}$, then f has a unique fixed point μ, and then $[\mu]$ is the intrinsic mean of Q.

- To find the sample intrinsic mean, replace Q by the empirical distribution, Q_n. That is, start with some $\theta_0 \in \mathbb{C}S^{k-1} \cap H^{k-1}$ and compute θ_m iteratively:

$$\begin{aligned}\theta_{m+1} &= f(\theta_m); \;\; m = 0, 1, 2, \dots \\ &= (cos|v_m|)\theta_m + \frac{sin|v_m|}{|v_m|} v_m\end{aligned}$$

where $v_m = 2 \sum_{j=1}^{n} \frac{1}{n} \frac{r_j}{sin r_j} (-cos r_j \theta_m + e^{i\theta_j} X_j)$

$$\begin{aligned} r_j &= arccos(|\theta_m^{'} \bar{X}_j|) \\ e^{i\theta_j} &= \frac{\theta_m^{'} \bar{X}_j}{cos r_j}\end{aligned}$$

- If all the sample points are in a geodesic ball of radius at most $\frac{3\pi}{40}$, then the above algorithm converges to μ_n, $[\mu_n]$ being the sample intrinsic mean, whetever θ_0 we start with in that ball. We may take $[\theta_0]$ to be the sample extrinsic mean.

- For the skull data, the female sample is contained in a geodesic ball of radius 0.0703 while the male sample is contained in a geodesic ball of radius 0.0855 around their respective sample extrinsic means. Since both radii are $<< 3\pi/40$, their sample intrinsic means exist and the above algorithm converges to that. The geodesic distance between the extrinsic and intrinsic means come to the order of 10^{-6}:

$$d_g(\mu_{nE}, \mu_{nI}) = 5.5395e - 07$$
$$d_g(\mu_{mE}, \mu_{mI}) = 1.9609e - 06$$

- 95% asymptotic confidence regions of the population intrinsic means are

Females: $\{[\mu_1] : nd_g^2(\mu_{nI}, \mu_1) \leq 0.0003247\}$
Males: $\{[\mu_2] : md_g^2(\mu_{mI}, \mu_2) \leq 0.0004691\}$

- To test if the male and female populations have same intrinsic mean shapes, we accept H_0: $[\mu_1] = [\mu_2]$ if the regions overlap that is if

$$d_g(\mu_{nI}, \mu_{mI}) < \sqrt{\frac{0.0003247}{n}} + \sqrt{\frac{0.0004691}{m}}$$

- For this sample, $d_g(\mu_{nI}, \mu_{mI}) = 0.0587$ while $\sqrt{\frac{0.0003247}{n}} + \sqrt{\frac{0.0004691}{m}} = 0.0073$. So we reject H_0.

Conclusion

There are many outstanding statistical problems in shape analysis which remain unresolved. One of them is a proper analysis of 3-D shape spaces and of distributions on them. Another is the study of time series for evolution of the distribution of shapes, in discrete and in continuous time. For 2-D shape spaces and other Riemannian manifolds, a mathematical problem of some significance is to find broad conditions for the uniqueness of the intrinsic mean.

Acknowledgements

I am very thankful to my advisor, Dr. Rabi Bhattacharya for guiding me through out my research. I would also like to thank other professors for their thoughtful and constructive suggestions.

References

[1] BHATTACHARYA, R. and PATRANGENARU, V. (2003). Large Sample Methods of Intrinsic and Extrinsic Sample Means on Manifolds-I. *Ann. Statist.*31 1-29.

[2] BHATTACHARYA, R. and PATRANGENARU, V. (2005). Large Sample Methods of Intrinsic and Extrinsic Sample Means on Manifolds-II. *Ann. Statist.*33 1225-1259.

[3] KARCHAR, H. (1977). Riemannian Center of Mass & Mollifier Smoothing. *Comm. on Pure & Applied Math.*XXX 509-541.

[4] LE, HUILING (2001). Locating Frechet Means with Application to Shape Spaces. *Adv. Appl. Prob.*33 324-338.

[5] PENNEC, XAVIER (1999). Probabilities and Statistics on Riemannian Manifolds: Basic Tools for Geometric Measurements. *NSIP'99*.

[6] DRYDEN, I.L. and MARDIA, K.V. (1998). Statistical Shape Analysis. *Wiley N.Y.*

[7] KENDALL, D.G.; BARDEN, D.; CARNE, T.K. and LE, H. (1999). Shape & Shape Theory. *Wiley N.Y.*

[8] LEE, J.M. (1997). Riemannian Manifolds an introduction to Curvature. *Springer*.