STATISTICS ON THE PLANER SHAPE SPACE

Abhishek Bhattacharya

Project Advisor: Dr.Rabi Bhattacharya

Overview

- The notion of means and variances of probability distributions on metric spaces are defined.

- Properties of mean and variation on Riemannian manifolds are presented.

- The configuration shape space has been defined, and its properties are presented.

- The results are applied to the shape space to carry out estimation and statistical inference on shapes.

Frechet Mean and Variation on Metric Spaces

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

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

- 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. [Bhattacharya, R. and Patrangenaru, Vic. 2003]

Frechet Variation

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

- If the Frechet function is finite for some p, then the variation is finite and is attained by all p in the Frechet mean set.

Sample Frechet Mean and Variation

- 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

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

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

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

Consistency of the Sample Frechet Mean and Variation

- **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. [Bhattacharya, R. and Patrangenaru, Vic. 2003]

- If every closed and bounded subset of M is compact, the sample variation is a strongly consistent estimator of the Frechet Variation of Q. [Bhattacharya, R. and Bhattacharya, A. 2006]

Mean on Riemannian manifolds

Intrinsic Means

- Let $(\mathbf{M}, \mathbf{g})$ be a d-dimensional connected complete Riemannian manifold, g being the riemannian metric. Let $\rho = d_g$, the geodesic distance under g.

- Let Q be a probability distribution on M. 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**.

Few Definitions

1. Geodesic: $\ddot{\gamma} = 0$. Locally length minimizing curves.
   eg. Great Circles on S^n, Straight lines in $\mathbb{R}^n$.

2. Exponential map: $p \in M$, $V \in T_pM$; $exp_pV = \gamma(1)$, where γ is a geodesic with $\gamma(0) = p$ and $\dot{\gamma}(0) = V$.

3. Cut locus C(p): Let γ be a geodesic with $\gamma(0) = p$. Let t_0 be the supremum of all t for which γ is length minimizing on $[0, t]$. Then $\gamma(t_0)$ is the cut point of p along γ. $C(p)$ is the set of all cut points of p along all geodesics.
   eg. $C(p) = \{-p\}$ on S^n.

3. Sectional Curvature: For a curve γ, its sectional curvature is $\pm|\ddot{\gamma}(t)|$, $+$ if the curve is pointing towards N and $-$ if it is pointing away from N, where N is a chosen normal field along γ.

   For a 2d manifold, chose a basis (X, Y) for T_pM. Then sectional curvature at p is $\frac{Rm(X,Y,Y,X)}{|X|^2|Y|^2 - <X,Y>^2}$, where Rm is the 'Riemann Curvature Tensor'.

   For a d dimensional manifold, consider the $2-D$ submanifold swept out by geodesics with initial velocities lying in a $2-D$ subspace, Π of T_pM. That is the sectional curvature at p associated with Π, K(Π).

Existence (Uniqueness) of Intrinsic means

[Karchar, H. 1977]

Suppose all sectional curvatures on M bounded above by $\mathbf{C} \geq 0$. Q is a probability distribution on M with finite Frechet function. Support of Q contained in a ball of radius 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 and hence the intrinsic mean exists (as a unique minimizer of F).

Asymptotic distribution of sample intrinsic mean

[Bhattacharya, R. and Patrangenaru, Vic. 2003]

- Assume support of Q contained in a closed geodesic ball $\overline{B}_r(p)$ with center p and radius r, r as above; which is disjoint from the cutlocus $C(p)$.

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

- 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. $\mu = \phi(\mu_I)$, μ_I being the intrinsic mean of Q. $\mu_n = \phi(\mu_{n,I})$, $\mu_{n,I}$ being a measurable selection from the sample mean set.

- $\Lambda = E((D_r D_s h(Y_1, \mu)))_{r,s=1}^{d}$;
  $\Sigma = Cov((D_r h(Y_1, \mu)))_{r=1}^{d}$.
  Then Λ and Σ are positive definite and

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

Asymptotic confidence set for the intrinsic mean

- The earlier result can be used to construct an assymptotic $1-\alpha$ confidence set for μ which is

  $$\{\mu_I : n(\mu_n - \mu)^t(\hat{\Lambda}^{-1}\hat{\Sigma}\hat{\Lambda}^{-1})^{-1}(\mu_n - \mu) \leq \mathcal{X}_d^2(1-\alpha)\}$$

- where $\mathcal{X}_d^2(1-\alpha)$ is the upper $(1-\alpha)^{\text{th}}$ quantile of the chi-squred distribution with d degrees of freedom,

- d is the dimension of the manifold and

  $$\hat{\Lambda} = \frac{1}{n}\sum_{i=1}^{n} D_r D_s h(Y_i, \mu_n)$$
  $$\hat{\Sigma} = \frac{1}{n}\sum_{i=1}^{n} D_r h(Y_i, \mu_n) D_s h(Y_i, \mu_n)$$

  are the sample estimates of Λ and Σ respectively.

Expressions for Λ and Σ

[Bhattacharya, A. 2006]

- M has all sectional curvatures bounded above by $C \geq 0$.

- Q has intrinsic mean μ_I, and for any geodesic γ, $\gamma(0) = \mu_I$; there exists $s_0 > 0$ such that the cut-locus of $\gamma\mid_{[0,s_0]}$, $C(\gamma\mid_{[0,s_0]})$ has Q measure 0.

- $Y_j = exp_{\mu_I}^{-1} X_j$ be the normal coordinates of the sample around μ_I. Then we have

$$D_r h(Y_1, \mu) = -2Y_1^r$$
$$E(D_r h(Y_1, \mu)) = 0$$
$$\Sigma = 4Cov(Y_1^r, Y_1^s)$$

Also when M has constant sectional curvature C,

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

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

where

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

Extrinsic Mean and Variation

- Let $\phi : M \to \mathbb{R}^k$ be an isometric map of M onto $\tilde{M} = \phi(M) \subset \mathbb{R}^k$: $\rho(x,y) = \|\phi(x) - \phi(y)\|$, 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 $Q_n = \sum_{i=1}^{n} \delta_{X_i}$ be the corresponding empirical distribution, then the Frechet mean (set) of Q_n (wrt ρ) is the **Extrinsic sample mean(set)**. The sample Frechet variation is the **sample extrinsic variation**.

Existence of Extrinsic Mean

[Bhattacharya, R. and Patrangenaru, Vic. 2003]

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

- 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 $P_{\tilde{M}}\mu$, 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 a **non focal point** of $\mathbb{R}^k$ (w.r.t. M̃); o.w. it is a **focal point**.

Consistency of the sample mean

[Bhattacharya, R. and Patrangenaru, Vic. 2003]

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

Asymptotic Distribution of the sample extrinsic mean

- In a neighborhood of a nonfocal point, μ, the projection map, $P(.)$ is smooth.

- Then $\sqrt{n}[P(\bar{Y}) - P(\mu)]$ has an asymptotic Gaussian distribution on the tangent space of $\phi(M)$ at $P(\mu)$, with mean vector zero.

- This can be used to construct a assymptotic confidence region for $P(\mu)$.

Assymptotic distribution of sample variation

[Bhattacharya, R. and Bhattacharya, A. 2006]

- Suppose Q has support in a single coordinate patch, (U, ϕ). [$\phi : U \longrightarrow \mathbb{R}^d$ smooth] and the Frechet Mean μ of Q exists. Let μ_n be a measurable selection from the sample mean set. Then

$$\sqrt{n}(F_n(\mu_n) - F(\mu)) \xrightarrow{\mathcal{L}} N(0, V\rho^2(X_1, \mu))$$

- This gives an asymptotic level α Confidence set for the population variation, $F(\mu)$ which is:

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

- where s^2 is the sample estimate of $V(\rho^2(X_1, \mu))$ which is

$$\frac{1}{n} \sum_{i=1}^{n} (\rho^2(X_j, \mu_n) - F_n(\mu_n))^2$$

The Planer shape space of k-ads, Σ_2^k

- Consider k points on the plane not all same. Assume $k > 2$ and refer to such a set as a k-ad.

- A k-ad can be represented by a complex k-vector.

- Shape of a complex k-ad $z = (z_1, \ldots z_k)$ is the orbit of z under translation, scaling and rotation. Can be represented by $[\mathbf{z}] = \{e^{i\theta}\frac{(\mathbf{z}-\langle\mathbf{z}\rangle)}{\|\mathbf{z}-\langle\mathbf{z}\rangle\|} \colon 0 \le \theta < 2\pi\}$.

- Thus the shape space of k-ads, Σ_2^k is the set of all complex lines on the hyperplane, $H_{k-1} = \{\mathbf{w} \in C^k \setminus \{0\} \colon \sum_1^k w_j = 0\}$, so has the structure of the complex projective space $\mathbb{C}P^{k-2}$ (the space of all complex lines through the origin in $\mathbb{C}^{k-1}$).

- Let $[\mathbf{z}]$ and $[\mathbf{w}]$ be two shapes. Let $\mathbf{u} = \frac{(\mathbf{z}-\langle\mathbf{z}\rangle)}{\|\mathbf{z}-\langle\mathbf{z}\rangle\|}$, $\mathbf{v} = \frac{(\mathbf{w}-\langle\mathbf{w}\rangle)}{\|\mathbf{w}-\langle\mathbf{w}\rangle\|}$. The **Procrustes coordinates** of $\mathbf{v}$ onto $\mathbf{u}$ is

$$\mathbf{v}^P = (\hat{a} + i\hat{b})1_k + \hat{\beta}e^{i\hat{\theta}}\mathbf{v}.$$

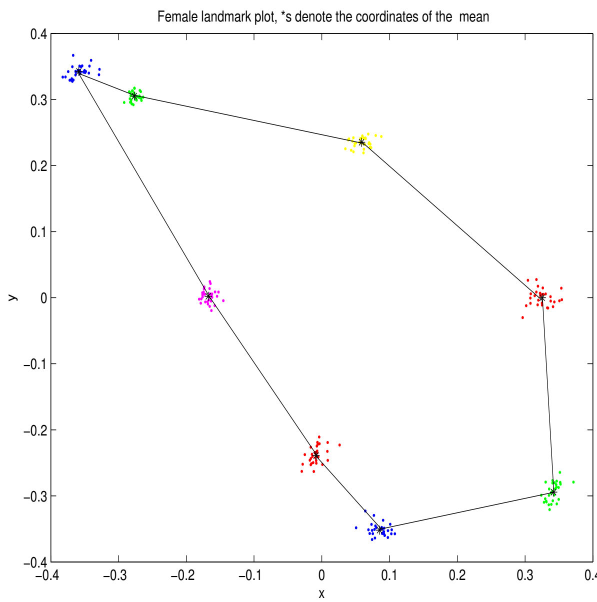
where (β, θ, a, b) are chosen to minimize

$$D^2(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \beta e^{i\theta}\mathbf{v} - (a + ib)\mathbf{1}_k\|^2$$

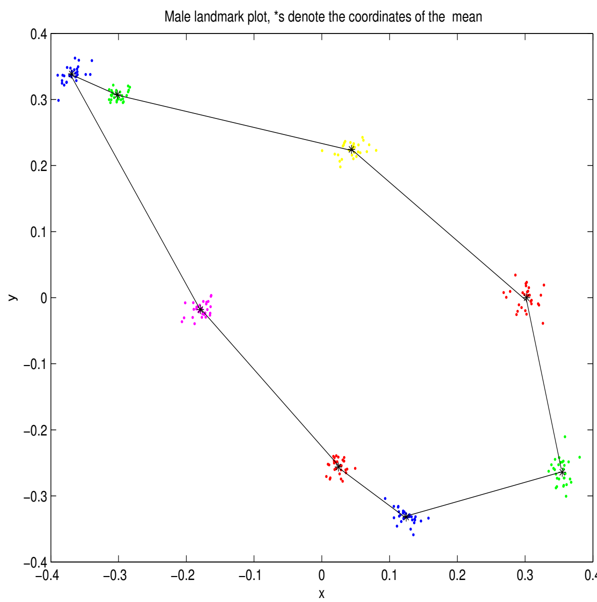
Application

- As an example, consider 8 locations on a skull projected on a plane. There are 30 female and 29 male samples. [Source: Statistical Shape Analysis - Dryden & Mardia].

- This is the plot of the procrustes coordinates of the female and male samples onto their extrinsic means.



Procrustes Coordinates of Female landmarks.
* denotes coordinates of the mean shape.



Procrustes Coordinates of Male landmarks.
* denotes coordinates of the mean shape.

Intrinsic Mean on Σ_2^k

- Σ_2^k can be embedded into $\mathbb{C}^{k^2}$ via the **Veronese-Whitney embedding** $\phi([\mathbf{u}]) = \mathbf{u}\overline{\mathbf{u}}'$. This induces a Riemannian structure on this space, known as the **Fubini-Study** metric.

- The geodesic distance for the Fubini-Study metric is

$$d_g([\mathbf{z}], [\mathbf{w}]) = \text{arccos}(|\mathbf{u}'\overline{\mathbf{v}}|) \in [0, \frac{\pi}{2}]$$
$$u = \frac{(\mathbf{z} - \langle \mathbf{z} \rangle)}{\|\mathbf{z} - \langle \mathbf{z} \rangle\|}, \;\; v = \frac{(\mathbf{w} - \langle \mathbf{w} \rangle)}{\|\mathbf{w} - \langle \mathbf{w} \rangle\|}.$$

- All sectional curvatures bounded between 1 and 4.

- Intrinsic mean of a probability distribution Q exists if its support contained in a geodesic ball of radius atmost $\frac{\pi}{8}$.

- Then any measurable selection from the sample mean set is consistent.

Extrinsic Mean on Σ_2^k

- The extrinsic distance wrt the Vernose-Whitney embedding, ϕ is:

$$\begin{aligned}\rho^2([\mathbf{z}],[\mathbf{w}]) &= \|\mathbf{u}\overline{\mathbf{u}}' - \mathbf{v}\overline{\mathbf{v}}'\|^2 \\ &= 2(1 - |\overline{\mathbf{u}}'\mathbf{v}|^2)\end{aligned}$$

- Let μ_0 denote the mean vector of $Q_0 \doteq Q \circ \phi^{-1}$, regarded as a probability measure on $\mathbb{C}^{k^2}$.

- The extrinsic mean μ_E, say, of Q is unique iff the eigenspace for the largest eigenvalue of $\mu_{\mathbf{0}}$ is (complex) one dimensional, and then $\mu_E = [\mathbf{w}]$, $\mathbf{w}(\neq 0) \in$ eigenspace of the largest eigenvalue of $\mu_{\mathbf{0}}$.

Sample Mean

- The extrinsic sample mean set of a iid sample $X_1, \ldots, X_n$ is the the class of the eigen space of the largest eigen value of $\bar{Y}$, where $Y_j = \phi(X_j); j = 1, \ldots, n$.

- If the largest eigen value of $\mu_{\mathbf{0}}$ is simple, then any measurable selection from the sample mean set is consistent.

Extrinsic Variation on Σ_2^k

- The extrinsic variation of Q is $2(1-\lambda_k)$ where λ_k is the largest eigenvalue of $\mu_{\mathbf{0}}$.

- The extrinsic sample variation is $2(1-\hat{\lambda}_k)$ where $\hat{\lambda}_k$ is the largest eigen value of $\bar{Y}$.

- The sample variation is a consistent estimator of the true variation (even when λ_k is not simple).

Testing equality of Extrinsic Means

- 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 $x_1, \dots, x_n$ and $y_1, \dots, y_n$ are random samples from Q_1 and Q_2 respectively.

- Define $X_i = \phi(x_i)$, $Y_i = \phi(y_i)$ as their images which are random samples from $Q_1 \circ \phi^{-1}$ and $Q_2 \circ \phi^{-1}$ respectively.

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

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

- Using the assymptotic distributions of the sample means, one can construct an assymptotic $\mathcal{X}^2$ test for the problem.

- For the male and female skull sample, value of the test statistic is 40.

- P-value for the test $= 7.18 \times 10^{-5}$.

- So we reject H_0, ie the male and females have different mean shapes.

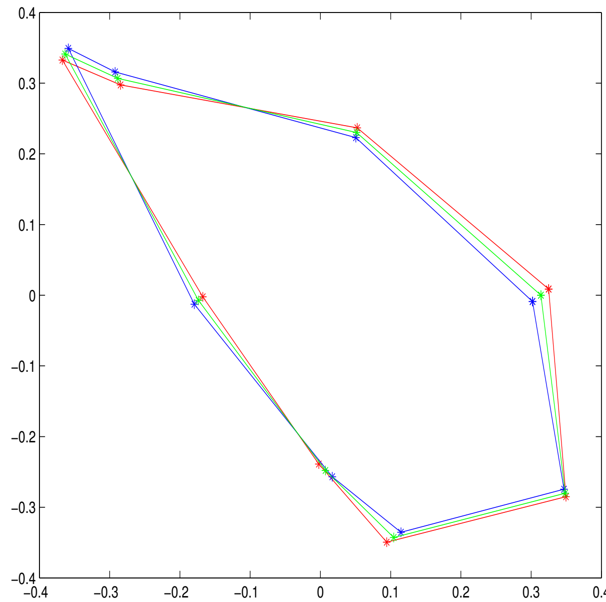


Figure 3: Procrustes coordinates of sample extrinsic means.
Red: Female, Blue: Male, Green: Pooled sample mean

Testing equality of Extrinsic Variations

- Let V_1 and V_2 be the extrinsic variations of Q_1 and Q_2 respectively.

- Let $\hat{V}_1$ and $\hat{V}_2$ be the respective sample variations.

- Then assuming that the populations have unique means μ_1 and μ_2 respectively,

  $\sqrt{n}(\hat{V}_1 - V_1) \xrightarrow{\mathcal{L}} N(0, \sigma_1^2)$
  $\sqrt{n}(\hat{V}_2 - V_2) \xrightarrow{\mathcal{L}} N(0, \sigma_2^2)$
  where $\sigma_1^2 = V\rho^2(x_1, \mu_1),\ \sigma_2^2 = V\rho^2(y_1, \mu_2)$

- To test if the two populations have the same spread around their respective centers, i.e.

$$H_0 : V_1 = V_2$$

- Under H_0,

$$T = \sqrt{n}\frac{\widehat{V}_1 - \widehat{V}_2}{s_1^2 + s_2^2} \xrightarrow{\mathcal{L}} N(0,1)$$

  where s_1^2 and s_2^2 are the sample estimates of σ_1^2 and σ_2^2 respectively.

- Reject H_0 if $|T| > Z_{1-\frac{\alpha}{2}}$. This is an asymptotic level α test.

- For the skull sample, the test statistic, $T = -0.9226$ and the assymptotic P-value of the test is 0.3562.

- So we accept H_0, the spreads are the same.

- This is also obvious from the 95% confidence intervals of the variations which are:
  Males: $(0.0031, 0.0046)$
  Females: $(0.0034, 0.0056)$.

Future work

- To come up with more general results on the existence (uniqueness) of Intrinsic means.
- To derive the assymptotic distribution of the sample variation when the mean does not exist.
- To compare the performance of the non-parametric tests with that of parametric tests under different shape distributions.
- To extend my theory to **Projective Shapes** which finds applications in Machine Vision.
- To combine Shape and Image Analysis for example in Medical Images.