STATISTICS ON MANIFOLDS

FRECHET MEANS AND THEIR ESTIMATION

Project Submitted by: Abhishek Bhattacharya
Advisor: Dr.Rabi Bhattacharya

Overview

- Frechet mean on Metric Spaces
- Extrinsic Mean
- Examples on some common Manifolds
- Intrinsic Mean
- Examples
- Numerical Calculations
- Other Work

Frechet Mean on Metric Spaces

- (M, ρ) a metric space and Q a probability measure on M.

- The **Frechet function** of Q, $F(p) = \int_M \rho^2(p, x) Q(dx), \ \ p \in M$.

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

- $X_1, X_2, \ldots, X_n$ are iid with commom distribution Q, and $Q_n \doteq \frac{1}{n} \sum_{j=1}^{n} \delta_{X_j}$ is the corresponding empirical distribution.

- The Frechet mean set of Q_n is the **sample (Frechet) mean set**.

- If this set is a singleton, it is the **sample (Frechet) mean**.

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

- If the Frechet mean of Q is unique, then every measurable selection from the Frechet sample mean set is a strongly consistent estimator of the Frechet mean.

Extrinsic Means

- $\phi : M \to \mathbb{R}^k$ an isometric map of M onto $\tilde{M} = \phi(M) \subset \mathbb{R}^k : \rho(x, y) = \|\phi(x) - \phi(y)\|$, is the Euclidean distance.

- $P_{\tilde{M}}\mathbf{u} = \{x \in \tilde{M} : \|x - u\| \leq \|y - u\| \forall y \in \tilde{M}\}$.

- If this set is a singleton, $\mathbf{u}$ is a **nonfocal point** of $\mathbb{R}^k$ (w.r.t. M̃); o.w. it is a **focal point** of $\mathbb{R}^k$.

- The Frechet mean (set) of Q is the **Extrinsic mean(set)** of Q.

- If $X_i (i \geq 1)$ are iid observations from Q, and $Q_n = \sum_{i=1}^{n} \delta_{X_i}$, then the Frechet mean(set) of Q_n is the **Extrinsic sample mean(set)**.

- Let $\tilde{Q}, \tilde{Q}_n$ 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 $\tilde{\mu} = \int_{R^k} \mu \tilde{Q}(du)$ is the mean of $\tilde{Q}$, then the extrinsic mean set of Q is $\phi^{-1}(P_{\tilde{M}}^{\tilde{\mu}})$.

- If $\tilde{\mu}$ is a nonfocal point of R^k (relative to $\tilde{M}$), then the extrinsic sample mean μ_n is a strongly consistent estimator of the extrinsic mean $\mu = \phi^{-1}(P_{\tilde{M}}^{\tilde{\mu}})$.

Examples

- **Example 1 ($\mathbf{S^{k-1}}$)**: The inclusion map $i : S^{k-1} \to \mathbb{R}^k$, $i(x) = x$. The extrinsic mean set of Q on S^{k-1} is 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 on $\mathbb{R}^k$. $\tilde{\mu}$ is non-focal iff $\tilde{\mu} \neq 0$.

- **Example 2 ($\mathbb{R}\mathbf{P}^{\mathbf{k-1}}$)**: $\mathbb{R}P^{k-1}$ = All lines $(\lambda x : \lambda \in \mathbb{R} \setminus \{0\})$ through the origin in $\mathbb{R}^k, x \neq 0$. Can be regarded as the quotient space of S^{k-1} under the relation $u \sim v$ iff $u = -v$.

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

- ϕ is an **Equivariant Embedding** of $\mathbb{R}P^{k-1}$.

- Metric ρ on $\mathbb{R}P^{k-1}$, $\rho^2([u],[v]) = \|uu' - vv'\|^2 = Trace(uu' - vv')^2$.

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

- $\tilde{\mu}$ is **nonfocal** iff its largest eigenvalue is **simple**.

- Then the extrinsic mean of Q is $[\mu_m]$, $\mu_m (\neq 0)$ is a unit eigen vector corresponding to the largest eigenvalue of $\tilde{\mu}$.

- **Example 3 (Planer Shape Space of k-ads,$\boldsymbol{\Sigma_2^k}$)**. Suppose k points on the plane, e.g., k locations on a skull projected on a plane, not all points being the same. Such a set a **k-ad** (or a set of **k landmarks**). Denoted by k complex numbers $(z_j = x_j + iy_j, 1 \leq j \leq k)$. The shape of a k-ad $\mathbf{z} = (z_1, z_2, \ldots, z_k)$, the equivalence class, or orbit of $\mathbf{z}$ under translation, rotation and scaling.

- To remove translation, substract $\langle \mathbf{z} \rangle \equiv (\langle \mathbf{z} \rangle, \langle \mathbf{z} \rangle, \ldots, \langle \mathbf{z} \rangle)$ $(\langle \mathbf{z} \rangle = \frac{1}{k} \sum_{j=1}^{k} z_j)$ from $\mathbf{z}$ to get $\mathbf{z} - \langle \mathbf{z} \rangle$. Rotation of the k-ad by an angle θ and scaling by a factor $r > 0$ achieved by multiplying $\mathbf{z} - \langle \mathbf{z} \rangle$ by $\lambda = r \exp i\theta$. Hence the shape of the k-ad, the complex line passing through $\mathbf{z} - \langle \mathbf{z} \rangle$.

- Structure of the complex projective space $\mathbb{C}P^{k-2}$. Represent the element of Σ_2^k corresponding to a k-ad $\mathbf{z}$ by the curve $\gamma(\mathbf{z}) = [\mathbf{z}] = \{e^{i\theta}\frac{(\mathbf{z}-\langle\mathbf{z}\rangle)}{\|\mathbf{z}-\langle\mathbf{z}\rangle\|} \colon 0 \le \theta < 2\pi\}$ on the unit sphere in $H_{k-1} = \{z \in \mathbb{C}^k : z.\mathbf{1} = 0\} \approx \mathbb{C}^{k-1}$.

- The **Veronese-Whitney embedding** of Σ_2^k given by $\phi\colon \Sigma_2^k \to \mathbb{C}^{k^2}$, $\phi([\mathbf{z}]) = \mathbf{u}\mathbf{u}^*$, $\mathbf{u} = \frac{(\mathbf{z}-\langle\mathbf{z}\rangle)}{\|\mathbf{z}-\langle\mathbf{z}\rangle\|}$.

- The distance ρ on Σ_2^k, $\rho^2([\mathbf{z}],[\mathbf{w}]) = \|\mathbf{u}\mathbf{u}^* - \mathbf{v}\mathbf{v}^*\|^2$.

- Q a probability measure on Σ_2^k, and μ_0 the mean vector of $Q_0 \doteq Q \circ \phi^{-1}$, regarded as a probability measure on C^{k^2} (or, $\mathbb{R}^{2k^2}$).

- The **extrinsic mean** μ_E, of Q is unique iff the eigenspace for the largest eigenvalue of μ_0 is (complex) one dimensional, and then $\mu_E = [\mathbf{w}]$, $\mathbf{w}(\neq 0) \in$ eigenspace of the largest eigenvalue of μ_0.

- Then it follows that any measurable selection from the sample extrinsic mean set is a consistent estimator of μ_E.

- **Example 4 (Size and Shape of Planer k-ads, $\mathbf{S\Sigma_2^k}$)** Comprised of all equivalence classes $[\mathbf{z}]$ of landmarks $\mathbf{z} = (z_1, z_2, \ldots, z_k) \in \mathbb{C}^k$, defined by $[\mathbf{z}] = \{e^{i\theta}(\mathbf{z} - <\mathbf{z}>) : 0 \leq \theta < 2\pi\}$.

- The **Veronese-Whitney** embedding ϕ of $S\Sigma_2^k$ into $\mathbb{C}^{k^2}$ (identified with the set of all $k \times k$ matrices with complex elements):

$$\phi([\mathbf{z}]) = ((u_j \overline{u}_{j'}))_{1 \leq j,j' \leq k} = \mathbf{u}\mathbf{u}^*$$

$$\mathbf{u} = (u_1, \ldots, u_k)', \ \ \mathbf{u}^* = \overline{\mathbf{u}}'$$

$$u_j = \frac{z_j - <z>}{\sqrt{r_{[\mathbf{z}]}}} \ \ (1 \leq j \leq k)$$

$$r^2{}_{[\mathbf{z}]} = \|\mathbf{z} - <\mathbf{z}>\|^2 = \sum_{j=1}^{k} |z_j - <z>|^2$$

- $\rho^2([\mathbf{z}], [\mathbf{w}]) = Trace(\mathbf{u}\mathbf{u}^* - \mathbf{v}\mathbf{v}^*)^2$.

- $\phi(S\Sigma_2^k)$ is a closed subset of $\mathbb{C}^{k^2}(\approx \mathbb{R}^{2k^2})$, but unbounded and, therefore not compact.

- Q a probability measure on $S\Sigma_2^k$, $Q \circ \phi^{-1}$, regarded as a probability measure on $\mathbb{C}^{k^2}(\approx \mathbb{R}^{2k^2})$ has finite second moments.

- $\tilde{\mu}$ the mean ($k \times k$ matrix) of $Q \circ \phi^{-1}$. If the largest eigen value of $\tilde{\mu}$, λ_k is simple, then the Extrinsic mean of Q is $\mu_E = [\lambda_k \mathbf{u}_0]$, where $\mathbf{u_0}$ is a unit eigen vector in the eigenspace of λ_k.

Intrinsic Mean

- (M, ρ) a Riemannian manifold, with ρ being the **geodesic distance** inherited from the natural connection on M.

- If Q is a probability measure on M, the Frechet mean (set) of Q wrt the distance ρ is called the **Intrinsic mean (set)** of Q.

Examples

- **Example 1**(S^{k-1}: Directional Space). At each $p \in S^{k-1}$, the metric tensor $g_p : T_p(S^{k-1}) \times T_p(S^{k-1}) \to \mathbb{R}$ is the restriction of the scaler product at p of the tangent space of $\mathbb{R}^k$: $g_p(v_1, v_2) = v_1.v_2$. g is a smooth metric tensor on the tangent bundle TS^{k-1}.

- The geodesics are the big circles, $\gamma_{p,v}(t) = (\cos t)p + (\sin t)v, -\pi < t \leq \pi)$. $\rho = \rho_g$ is the geodesic distance on S^{k-1}:

$$\rho_g(p,q) = |\cos^{-1}(p.q)| \in [0, \pi]$$

- Q be a probability measure on S^{k-1}. If Q is concentrated in a ball (w.r.t. the distance ρ_g) of radius less than $\pi/4$, then the Frechet mean exists as a unique minimizer. Such a mean is called the **intrinsic mean**.

- Then the sample Frechet mean based on a random sample from Q is consistent.

- **Example 2 (Shape Space Σ_2^k of Planer k-Ads)** For v_1, $v_2 \in T_{[\mathbf{z}]}\Sigma_2^k$, the metric tensor on $T\Sigma_2^k$ is taken to be the Euclidean scaler product $v_1.v_2$.

- The geodesic distance for this metric is (proportional to)

$$d_g([\mathbf{z}],[\mathbf{w}]) = \text{arccos}\,|\mathbf{z}'\overline{\mathbf{w}}|$$

- Given a sample of n k-ads $\mathbf{z}_r (1 \leq r \leq n)$, the intrinsic mean $[\mathbf{z}]_I$, is a minimizer of

$$nF_n(\tau) \equiv \sum_{r=1}^{n} \text{arccos}^2(|\mathbf{z}_r^{'}\overline{\tau}|^2), \quad (\|\tau\| = 1; \tau^{'} \in H_{k-}$$

- **Example 3**(Axial Space $\mathbb{R}P^{k-1}$) The Geodesic distance is

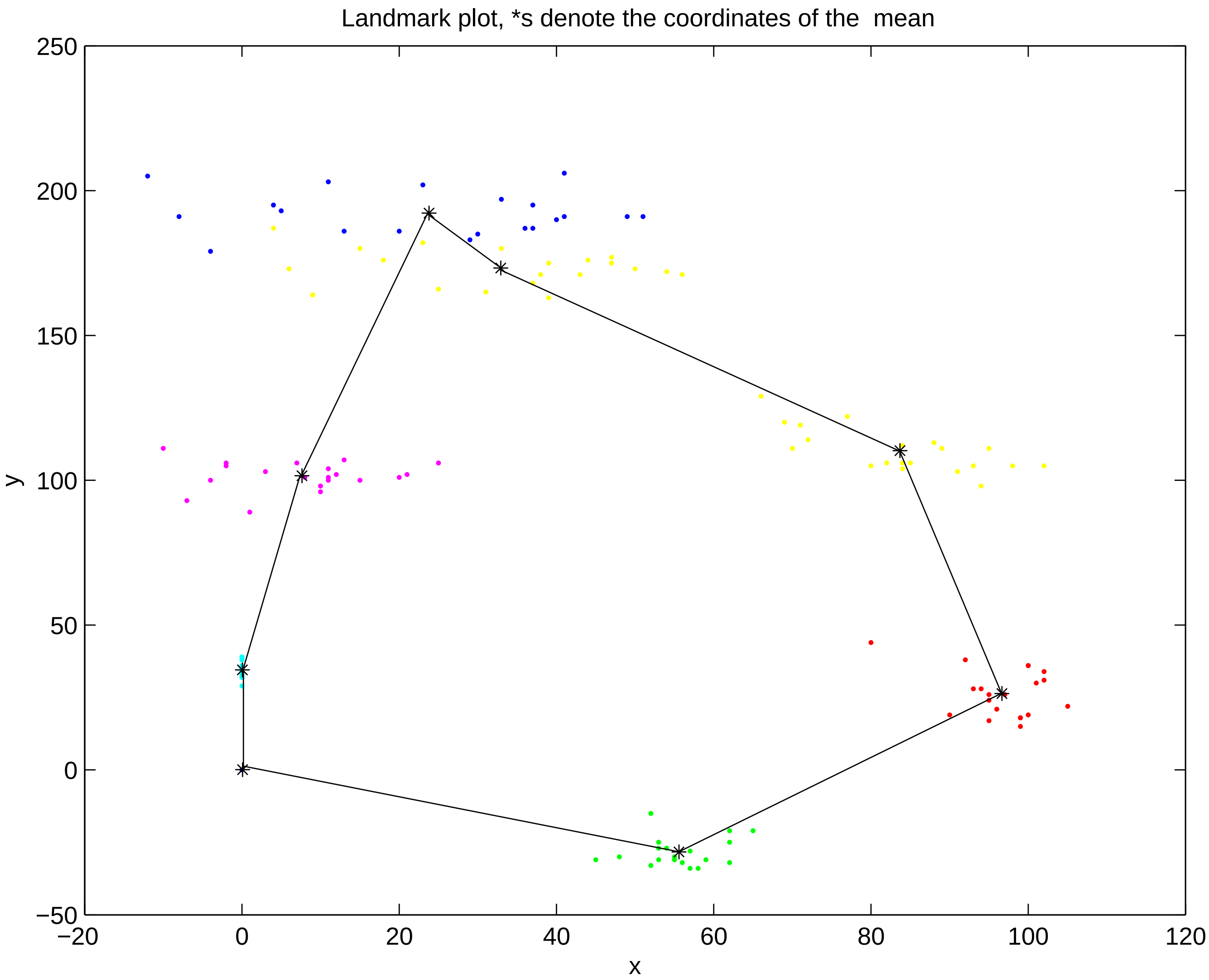
$$\rho_g([x],[y]) = \text{arccos}(|u^{'}v|)$$

- The intrinsic mean based on a sample $x_1, x_2, \ldots, x$ is a minimizer of

$$nF_n([y]) = \sum_{r=1}^{n} \text{arccos}^2(|x_r^{'}y|)$$

Applications

Referring to Example 3 of Extrinsic Mean, this is the plot of 8 landmarks from 20 gorilla skull pictures, along with the Extrinsic mean shape. The mean has been translated, rotated and scaled appropriately.


Landmark plot, *s denote the coordinates of the mean
250
200
150
100
50
0
−50
−20
0
20
40
60
80
100
120
x
y

Further Work done

- Found the assymptotic distribution of the sample Frechet mean, when the Frechet mean of Q is uniquely defined.

- Worked it out explicitly in case of Σ_2^k.

- Computed assymptotic Confidence Intervals for the mean of Q.

- Test for equality of means in two sample problems.

- Worked out the corresponding Bootstrap statistics.