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

• \((M, \rho)\) a metric space and \(Q\) a probability measure on \(M\).

• The Frechet function of \(Q\),
  \[ F(p) = \int_M \rho^2(p, x)Q(dx), \quad p \in M. \]

• 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 common distribution \(Q\), and \(Q_n = \frac{1}{n} \sum_{j=1}^{n} \delta_{X_j}\) 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.

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

• If this set is a singleton, $u$ is a nonfocal point of $\mathbb{R}^k$ (w.r.t. $\tilde{M}$); 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 $\mathbb{R}^k$: $\tilde{Q} = Q \circ \phi^{-1}, \tilde{Q}_n = Q_n \circ \phi^{-1}$.

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

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

• Example 1 ($S^{k-1}$): 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** \( \mathbb{RP}^{k-1} \): \( \mathbb{RP}^{k-1} = \) 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** \( \phi \) into the space of all \( k \times k \) matrices identified with \( \mathbb{R}^{k^2} \), \( \phi([u]) = uu', (u = (u_1, .., u_k)' \in S^{k-1}) \).

• \( \phi \) is an **Equivariant Embedding** of \( \mathbb{RP}^{k-1} \).

• **Metric** \( \rho \) on \( \mathbb{RP}^{k-1} \), \( \rho^2([u], [v]) = \|uu' - vv'\|^2 = \text{Trace}(uu' - vv')^2 \).
• $Q$ be a probability measure on $\mathbb{RP}^{k-1}$, and $\tilde{\mu}$ the mean of $\tilde{Q} = Q \circ \phi^{-1}$ 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 eigenvector corresponding to the largest eigenvalue of $\tilde{\mu}$. 
Example 3 (Planer Shape Space of k-ads, $\Sigma_2^k$). 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 $z = (z_1, z_2, \ldots, z_k)$, the equivalence class, or orbit of $z$ under translation, rotation and scaling.

To remove translation, substract $\langle z \rangle \equiv (\langle z \rangle, \langle z \rangle, \ldots, \langle z \rangle)$ ($\langle z \rangle = \frac{1}{k} \sum_{j=1}^{k} z_j$) from $z$ to get $z - \langle z \rangle$. Rotation of the $k$-ad by an angle $\theta$ and scaling by a factor $r > 0$ achieved by multiplying $z - \langle z \rangle$ by $\lambda = r \exp i\theta$. Hence the shape of the $k$-ad, the complex line passing through $z - \langle z \rangle$. 
• Structure of the complex projective space \( \mathbb{C}P^{k-2} \). Represent the element of \( \Sigma^k_2 \) corresponding to a k-ad \( z \) by the curve \( \gamma(z) = [z] = \{ e^{i\theta} \frac{(z-\langle z \rangle)}{\|z-\langle z \rangle\|} \mid 0 \leq \theta < 2\pi \} \) on the unit sphere in \( H_{k-1} = \{ z \in \mathbb{C}^k : z.1 = 0 \} \approx \mathbb{C}^{k-1} \).

• The Veronese-Whitney embedding of \( \Sigma^k_2 \) given by \( \phi : \Sigma^k_2 \rightarrow \mathbb{C}^k, \phi([z]) = uu^*, \) where \( u = \frac{(z-\langle z \rangle)}{\|z-\langle z \rangle\|}. \)

• The distance \( \rho \) on \( \Sigma^k_2 \), \( \rho^2([z],[w]) = \|uu^*-vv^*\|^2 \).
• $Q$ a probability measure on $\Sigma_{-2}^{k}$, and $\mu_0$ the mean vector of $Q_0 \doteq Q \circ \phi^{-1}$, regarded as a probability measure on $C_{2}^{k}$ (or, $\mathbb{R}^{2k^2}$).

• The **extrinsic mean** $\mu_E$, of $Q$ is unique iff the eigenspace for the largest eigenvalue of $\mu_0$ is (complex) one dimensional, and then $\mu_E = [w]$, $w(\neq 0) \in$ eigenspace of the largest eigenvalue of $\mu_0$.

• Then it follows that any measurable selection from the sample extrinsic mean set is a consistent estimator of $\mu_E$. 
• **Example 4 (Size and Shape of Planer k-ads, $SΣ^k_2$)** Comprised of all equivalence classes $[z]$ of landmarks $z = (z_1, z_2, \ldots, z_k) \in \mathbb{C}^k$, defined by $[z] = \{e^{i\theta}(z - <z>): 0 \leq \theta < 2\pi\}$.

• The **Veronese-Whitney** embedding $\phi$ of $SΣ^k_2$ into $\mathbb{C}^{k^2}$ (identified with the set of all $k \times k$ matrices with complex elements):

\[
\phi([z]) = ((u_j \bar{u}_{j'}))_{1 \leq j, j' \leq k} = uu^*
\]
\[
u = (u_1, \ldots, u_k)', \quad u^* = \bar{u}'
\]
\[
u_j = \frac{z_j - <z>}{\sqrt{r[z]}} \quad (1 \leq j \leq k)
\]
\[
r^2[z] = \|z - <z>\|^2 = \sum_{j=1}^{k} |z_j - <z>|^2
\]

• $\rho^2([z], [w]) = Trace(uu^* - vv^*)^2$. 


• $\phi(S\Sigma^k_2)$ 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 $S\Sigma^k_2$, $Q \circ \phi^{-1}$, 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 $\tilde{\mu}$, $\lambda_k$ is simple, then the Extrinsic mean of $Q$ is $\mu_E = [\lambda_k u_0]$, where $u_0$ is a unit eigen vector in the eigenspace of $\lambda_k$. 
Intrinsic Mean

• $(M, \rho)$ a Riemannian manifold, with $\rho$ being the \textbf{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 $\rho$ is called the \textbf{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 scalar product at $p$ of the tangent space of $\mathbb{R}^k$ : $g_p(v_1, v_2) = v_1 \cdot v_2$. $g$ 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 $\rho_g$) 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 $\Sigma^k_2$ of Planer k-Ads)** For $v_1, v_2 \in T_{[z]} \Sigma^k_2$, the metric tensor on $T \Sigma^k_2$ is taken to be the Euclidean scaler product $v_1.v_2$.

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

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

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

$$nF_n(\tau) \equiv \sum_{r=1}^{n} \arccos^2(|z'_r \bar{\tau}|^2), \quad (||\tau|| = 1; \tau' \in H_{k-1})$$
• **Example 3** (Axial Space $\mathbb{R}P^{k-1}$) The Geodesic distance is

$$\rho_g([x],[y]) = \arccos(|u^tv|)$$

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

$$nF_n([y]) = \sum_{r=1}^{n} \arccos^2(|x'_r y|)$$
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
Further Work done

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

- Worked it out explicitly in case of $\Sigma_2^k$.

- Computed asymptotic Confidence Intervals for the mean of $Q$.

- Test for equality of means in two sample problems.

- Worked out the corresponding Bootstrap statistics.