DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF COPENHAGEN

CLT and Smeariness for Means on Riemannian Manifolds

Pernille Emma Hartung Hansen, *pehh@di.ku.dk*, University of Copenhagen.

Introduction

- In search of a well-defined mean point for random variables on Riemannian manifolds, inspiration from the Euclidean mean value has been crucial.
- Most commonly studied is the Fréchet mean whose definition originates from the least squared distance property of the Euclidean mean.
- Existence, uniqueness and estimators of the Fréchet mean are still heavily studied.
- Under certain conditions, the Central Limit Theorem can be extended to Fréchet means and the extended definition of smeariness can further reveal behavior of the Fréchet estimator.
- Less studied are Maximum Likelihood means on Riemannian Manifolds for which CLT results have not been extended.

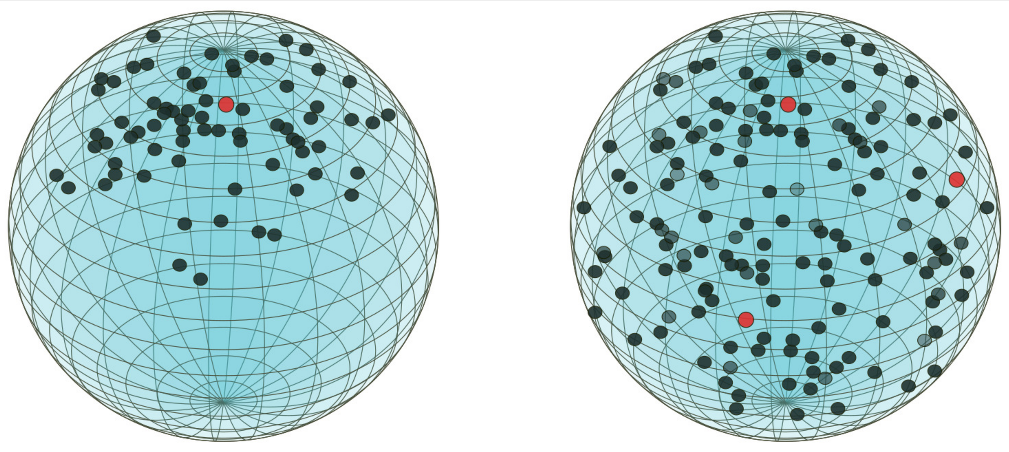

Figure 1: The illustration exemplifies two different distributions on the sphere $\mathcal{S}^2$ and the difficulty in finding a unique mean point.

Riemannian manifolds

A Riemannian manifold is a pair of a smooth manifold $\mathcal{M}$ and a Riemmanian metric g ie. a collection of inner products $g_p(v,u) = \langle v,u\rangle_p$ for each $p \in \mathcal{M}$. The Riemannian metric induces a distance on $\mathcal{M}$

$$\text{dist}(p,q) = \min_{\gamma} \int_0^1 ||\dot\gamma(t)||_{\gamma(t)} dt \tag{1}$$

where γ is a continuous differentiable curve from p to q. The exponential $\exp_p : T_p\mathcal{M} \to \mathcal{M}\backslash\mathcal{C}(p)$ are maps on each tangent space $T_p\mathcal{M}$ defined by

$$\exp_p(v) = \gamma_v(1) \tag{2}$$

where γ_v is the unique maximal geodesic with $\gamma_v(0) = p$ and $\dot\gamma_v(0) = v$. These are diffeomorphisms with inverse denoted by $\log_p$.

Fréchet means

A point $\mu \in \mathcal{M}$ is a **Fréchet mean** of the random point $\mathbf{x} : \Omega \to \mathcal{M}$ if

$$\mu \in \arg\min \mathbb{E}[\text{dist}(\mathbf{x}, y)^2] \tag{3}$$

Note: There might not exist such a point or there might be more than one!

Let $F_n(x) = \frac{1}{N}\sum_{i=1}^N \text{dist}(x, X_i)^2$ be the Empirical Fréchet function. We want to consider the behavior of the mean estimator μ_n which is a measurable collection where

$$\mu_n \in \arg\min_{x\in\mathcal{M}} F_n(x) \tag{4}$$

Central Limit Theorem

For the CLT for Fréchet means of random points $\mathbf{x}$ on Riemmanian $m-$manifold $\mathcal{M}$ we assume:

1. $\mathbf{x}$ has a unique Fréchet mean $\mu \in \mathcal{M}$ and $\mu_n \xrightarrow{\mathbb{P}} \mu$.
2. Almost Surely Locally Lipschitz and Differentiable at Mean of $\text{dist}(p,q)^2$
3. For some $r \geq 2$, rotation matrix $R \in SO(m)$ and $T_1, ..., T_m$ it holds that
$$F(x) = F(0) + \sum_{j=1}^m T_j |(Rx)_j|^r + o(||x||^r)$$
where $F(x) = \mathbb{E}[\text{dist}(\mathbf{x}, \exp_p(x))^2]$.

Theorem 1 (Eltzner and Huckemann [2]) *Under assumption 1. − 3. it holds that*

$$\sqrt{n}(V_1^n, ..., V_m^n)^T \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma) \tag{5}$$

where

1. $V_n^i = (R\log_\mu(\mu_n))_i |(R\log_\mu(\mu_n))_i|^{r-2}$

2. $\Sigma = \frac{1}{r^2} T^{-1} Cov[grad|_{x=0} dist(x,X)^2] T^{-1}$

3. $T = diag(T_1, ..., T_m)$

Smeariness

When generalizing the classical CLT theorem, the definition of smeariness arises to fit other convergence rates.

Definition 1 *A sequence $\mu_n \xrightarrow{\mathbb{P}} \mu$ is said to be k-smeary with limiting distribution X if*

$$n^{\frac{1}{2(k+1)}} \log_\mu(\mu_n) \xrightarrow{\mathcal{D}} X \tag{6}$$

Note: Theorem 1 can be restated as $(V_1^n, ..., V_m^n)^T$ being $0-$smeary.

Example: An odd distribution

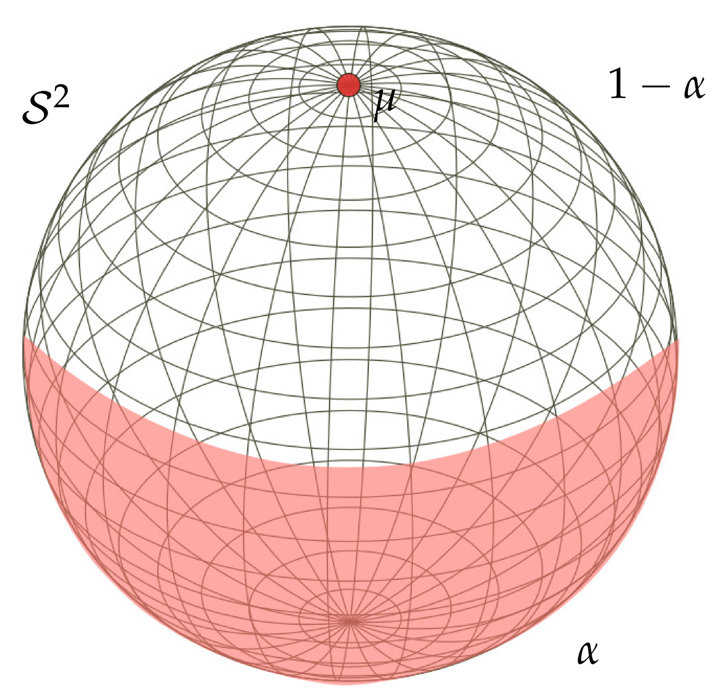

Figure 2: The distribution of X.

Let X be a random point on $\mathcal{S}^2$ with the following distribution (see Figure 4):

1. uniformly distributed on the lower hemisphere $\mathbb{L} = \{(q_1,q_2,q_3) \in \mathcal{S}^2 | q_2 \leq 0\}$ with total mass $0 < \alpha < 1$
2. point mass in the north pole $\mu = (0,1,0)$ with probability $1-\alpha$.

Theorem 2 (Eltzner and Huckemann [2]) *Let μ_n be the Fréchet estimator of X. For $\alpha = \left(1 + \frac{vol(\mathcal{S}^3)}{2vol(\mathcal{S}^2)}\right)^{-1} \approx 0.56$ the only Fréchet mean of X is μ and there exists a full-rank matrix Σ such that*

$$\left(n^{1/6}\log_\mu(\mu_n)\right)^3 \to \mathcal{N}(0,\Sigma). \tag{7}$$

ie. the Fréchet estimator μ_n is $2-$smeary.

1. For $1 > \alpha > 0.56$ the Fréchet mean is not unique.
2. For $0 \leq \alpha < 0.56$ the Fréchet mean is unique, but as seen in Figure 3 the Fréchet estimator slowly approaches 0-smeariness when α decreases.

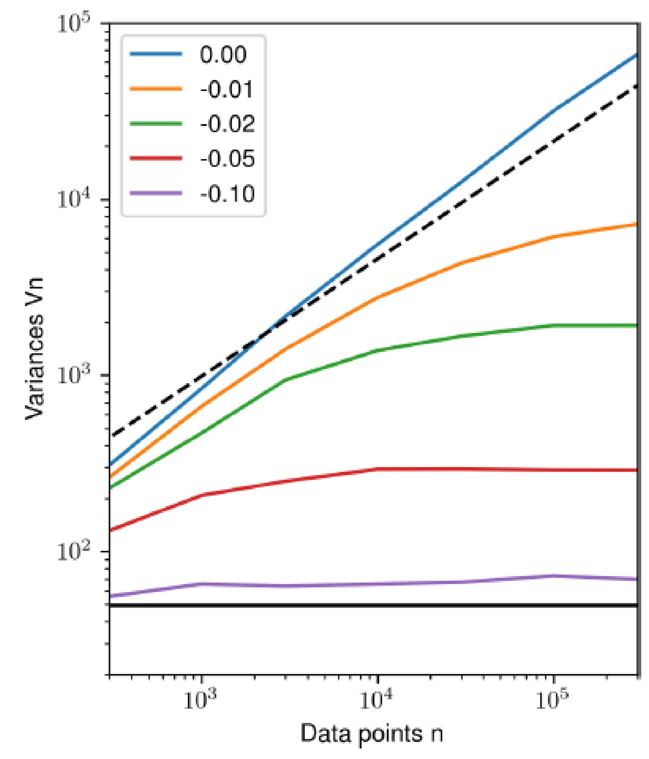

Figure 3: The variance of the Fréchet estimator for different values of α. Black line indicates 0-smeary and dotted line indicates 2-smeary.

Maximum Likelihood Mean

The heat kernel $p(x,\mu,t)$ is the transition density of the Brownian motion B_t with initial value μ.

Definition 2 *A point $\mu \in \mathcal{M}$ is a Maximum Likelihood mean of X if $\mu \in \arg\min_{y\in\mathcal{M}} \mathbb{E}[-\ln p(X,y,t)]$ for $t \in \mathbb{R}_+$*

As for the Fréchet mean, we are interested in a CLT or smeareness result for the estimator

$$\mu_n \in \arg\min_{y\in\mathcal{M}} -\frac{1}{n}\sum_{i=1}^n \ln p(X_i, y, T) \tag{8}$$

Goals

1. Existence and uniqueness results for this mean
2. CLT result for MLE under reasonable assumptions
3. Smeariness result for MLE for certain distributions

The odd distribution revisited

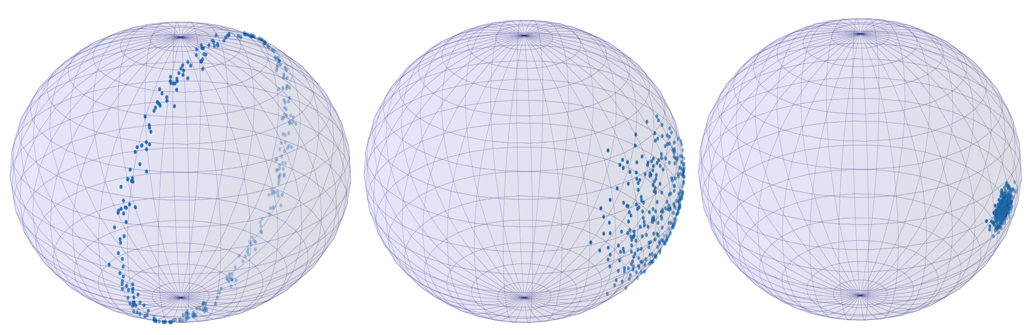

Figure 4: Estimated Maximum likelihood means for X with $t = 0.1, 0.5, 1$ [1].

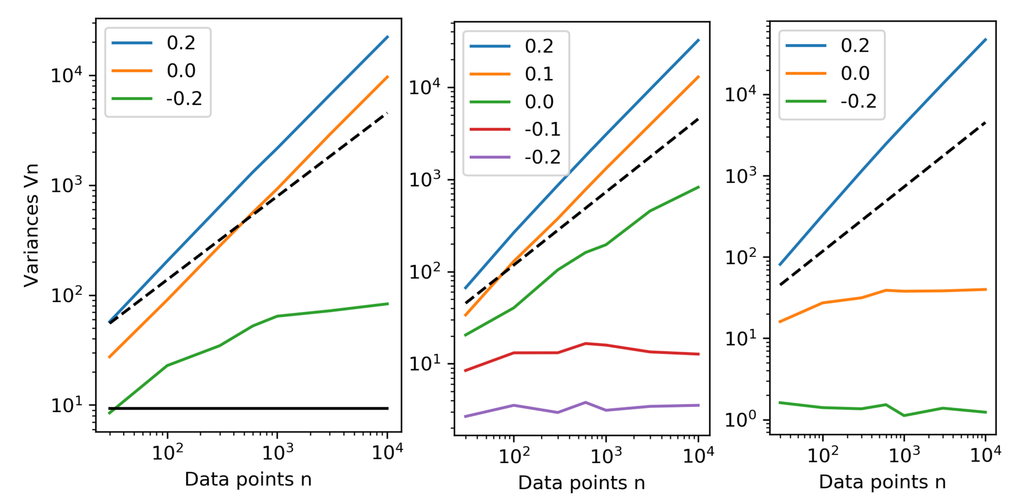

Figure 5: Variance for Maximum Likelihood estimator for $t = 0.1, 0.5, 1$[1]

References

[1] Benjamin Eltzner. Illustration generated from code.

[2] Benjamin Eltzner and Stephan F. Huckemann. A Smeary Central Limit Theorem for Manifolds with Application to High Dimensional Spheres. *arXiv:1801.06581 [math, stat]*, January 2018. arXiv: 1801.06581.