Principal Curves on Riemannian Manifolds — Supplementary Material —

Søren Hauberg

Appendix A Proof of Theorem 2

In this appendix we prove Theorem 2.

Theorem 2. *Let $\mathbf{x} \in T_\mu\mathcal{M}$, where $\mathcal{M}$ is either the D-sphere, the Euclidean plane, or the hyperbolic plane, follow a zero-mean elliptical distribution with support inside the cut-locus and finite covariance Σ. Let $\mathbf{v}_1, \ldots, \mathbf{v}_D$ be the eigenvectors of Σ in order of decreasing eigenvalues. Then, the principal geodesic $t \mapsto \mathrm{Exp}_\mu(t \cdot \mathbf{v}_1)$ is a principal curve.*

Proof for the Euclidean plane: The result is well-known in the Euclidean case [1]. □

Proof for the sphere $\mathbb{S}^D$: A point on the principal geodesic is written as

$$\mathbf{c}_t \equiv \mathbf{c}(t) = \mathrm{Exp}_\mu(t \cdot \mathbf{v}_1) = \cos(t)\mu + \sin(t)\mathbf{v}_1. \tag{1}$$

An orthonormal basis of $T_{\mathbf{c}_t}\mathbb{S}^D$ is given by $\hat{\mathbf{v}}_1, \mathbf{v}_2, \ldots, \mathbf{v}_D$, where $\hat{\mathbf{v}}_1$ denotes the velocity of the principal geodesic at $\mathbf{c}_t$. The collection of points on $\mathbb{S}^D$ that project onto the principal geodesic at $\mathbf{c}_t$ can be written as

$$\hat{\mathbf{c}}_\perp(\hat{t}) = \mathrm{Exp}_{\mathbf{c}_t}(\hat{t} \cdot \hat{\mathbf{v}}_\perp), \quad \hat{t} \in (-\pi/2, \pi/2), \tag{2}$$

where

$$\hat{\mathbf{v}}_\perp = \sum_{d=2}^{D} w_d \mathbf{v}_d, \quad \text{s.t. } \|\hat{\mathbf{v}}_\perp\| = 1. \tag{3}$$

Note that the unit norm constraint can be dropped if the domain of $\hat{t}$ is changed accordingly; for our purposes this is irrelevant. We have

$$\begin{aligned}\hat{\mathbf{c}}_\perp(\hat{t}) &= \cos(\hat{t})\mathbf{c}_t + \sin(\hat{t})\hat{\mathbf{v}}_\perp \\ &= \cos(\hat{t})\cos(t)\mu + \cos(\hat{t})\sin(t)\mathbf{v}_1 + \sin(\hat{t})\hat{\mathbf{v}}_\perp.\end{aligned} \tag{4}$$

To compute the logarithm of $\hat{\mathbf{c}}_\perp(\hat{t})$ we will need

$$\mu^\intercal \hat{\mathbf{c}}_\perp(\hat{t}) = \cos(\hat{t})\cos(t) \equiv \Delta(\hat{t}). \tag{5}$$

• *S. Hauberg is with the Section for Cognitive Systems at the Technical University of Denmark.*
E-mail: `sohau@dtu.dk`

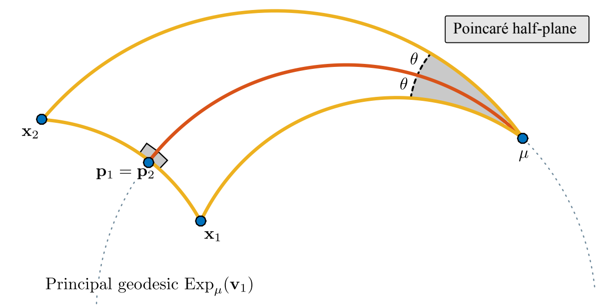

Fig. 1. Notation for the proof of the hyperbolic case.

With this, we compute

$$\begin{aligned}\mathrm{Log}_\mu(\hat{\mathbf{c}}_\perp(\hat{t})) &= \cos^{-1}(\Delta(\hat{t}))\frac{\hat{\mathbf{c}}_\perp(\hat{t}) - \Delta(\hat{t})\mu}{\|\hat{\mathbf{c}}_\perp(\hat{t}) - \Delta(\hat{t})\mu\|} \qquad (6)\\ &= \cos^{-1}(\Delta(\hat{t}))\frac{\cos(\hat{t})\sin(t)\mathbf{v}_1 + \sin(\hat{t})\hat{\mathbf{v}}_\perp}{\|\cos(\hat{t})\sin(t)\mathbf{v}_1 + \sin(\hat{t})\hat{\mathbf{v}}_\perp\|}. \qquad (7)\end{aligned}$$

This is a point in $T_\mu\mathbb{S}^D$. Expressed in the basis $\mathbf{v}_1, \ldots, \mathbf{v}_D$ it has coordinates $\mathrm{Log}_\mu(\hat{\mathbf{c}}_\perp(\hat{t})) = (t, a \cdot w_2, \ldots, a \cdot w_D)^\intercal$, where

$$a = \frac{\cos^{-1}(\Delta(\hat{t}))\sin(\hat{t})}{\|\cos(\hat{t})\sin(t)\mathbf{v}_1 + \sin(\hat{t})\hat{\mathbf{v}}_\perp\|}. \tag{8}$$

Next, we compute

$$\begin{aligned}&\mathrm{Log}_\mu(\hat{\mathbf{c}}_\perp(-\hat{t})) \\ &\quad= \cos^{-1}(\Delta(\hat{t}))\frac{\cos(\hat{t})\sin(t)\mathbf{v}_1 - \sin(\hat{t})\hat{\mathbf{v}}_\perp}{\|\cos(\hat{t})\sin(t)\mathbf{v}_1 + \sin(\hat{t})\hat{\mathbf{v}}_\perp\|},\end{aligned} \tag{9}$$

which has coordinates $\mathrm{Log}_\mu(\hat{\mathbf{c}}_\perp(-\hat{t})) = (t, -a \cdot w_2, \ldots, -a \cdot w_D)^\intercal$. The symmetry of elliptical distributions then implies that

$$p(\hat{\mathbf{c}}_\perp(\hat{t})) = p(\hat{\mathbf{c}}_\perp(-\hat{t})). \tag{10}$$

This symmetry implies that the expectation of all points along $\hat{\mathbf{c}}_\perp(\hat{t})$ is $\mathbf{c}_t$. As this holds for any $\hat{\mathbf{c}}_\perp$, we have

$$\mathbb{E}(\mathbf{x} \mid \mathrm{proj}_{\mathbf{c}}(\mathbf{x}) = \mathbf{c}_t) = \mathbf{c}_t \tag{11}$$

and self-consistency follows. □

Proof for the hyperbolic plane: Let

$$\begin{aligned}\mathbf{y}_1 &= t \cdot \mathbf{v}_1 + u \cdot \mathbf{v}_2 \\ \mathbf{y}_2 &= t \cdot \mathbf{v}_1 - u \cdot \mathbf{v}_2\end{aligned} \tag{12}$$

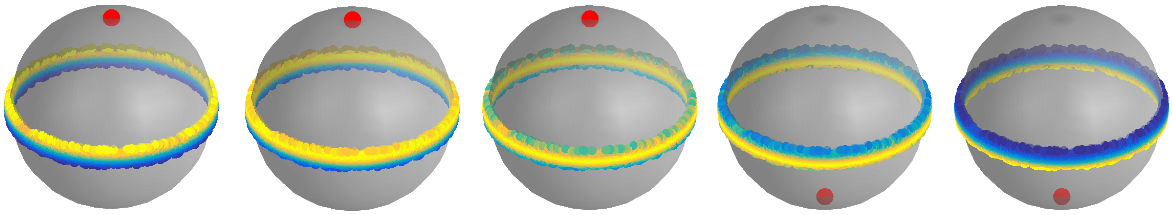

Fig. 2. Data is distributed uniformly in a band around the equator of the unit sphere. Weights (indicated by color) change smoothly according the z-coordinate, yet the intrinsic mean changes discontinuously.

be two points in the tangent space at μ, and let $\mathbf{x}_1 = \mathrm{Exp}_\mu(\mathbf{y}_1)$ and $\mathbf{x}_2 = \mathrm{Exp}_\mu(\mathbf{y}_2)$ be the corresponding points in the hyperbolic plane. Let $\mathbf{p}_1 = \mathrm{Exp}_\mu(\hat{t}_1 \cdot \mathbf{v}_1)$ and $\mathbf{p}_2 = \mathrm{Exp}_\mu(\hat{t}_2 \cdot \mathbf{v}_1)$ denote the projections of $\mathbf{x}_1$ and $\mathbf{x}_2$ onto the principal geodesic. Then $(\mu, \mathbf{x}_1, \mathbf{p}_1)$ and $(\mu, \mathbf{x}_2, \mathbf{p}_2)$ form two right hyperbolic triangles with identical angles θ at μ. From hyperbolic trigonometry it follows that

$$\begin{aligned} \tanh(\hat{t}_1) &= \cos(\theta) \cdot \tanh(\|\mathbf{y}_1\|) \\ \tanh(\hat{t}_2) &= \cos(\theta) \cdot \tanh(\|\mathbf{y}_2\|), \end{aligned} \tag{13}$$

which implies that $\hat{t}_1 = \hat{t}_2$, *i.e.* $\mathbf{x}_1$ and $\mathbf{x}_2$ project to the same point on the principal geodesic. Trigonometry further shows

$$\begin{aligned} \sinh(\mathrm{dist}(\mathbf{x}_1, \mathbf{p}_1)) &= \sin(\theta) \cdot \sinh(\|\mathbf{y}_1\|) \\ \sinh(\mathrm{dist}(\mathbf{x}_2, \mathbf{p}_2)) &= \sin(\theta) \cdot \sinh(\|\mathbf{y}_2\|), \end{aligned} \tag{14}$$

which implies that $\mathbf{x}_1$ and $\mathbf{x}_2$ are equally far away from the principal geodesic. As $p(\mathbf{x}_1) = p(\mathbf{x}_2)$, the average of $\mathbf{x}_1$ and $\mathbf{x}_2$ equals the point at which they project to the principal geodesic. Since this holds for any point pair (12), the principal geodesic is self-consistent. □

Appendix B
Discontinuous Intrinsic Means

In Sec. 3.2 of the manuscript it is noted that a smooth change in the weights of a weighted intrinsic mean can cause a discontinuous change in the intrinsic mean. This issue is best illustrated by an example.

Consider data $\mathbf{x}_{1:N}$ distributed uniformly in a band around the equator of the unit sphere. With equal weights, two equally optimal intrinsic means are available in the form of the poles of the sphere. We assign weights to each observation according to its z-coordinate

$$w_n = \exp\left(-\frac{(z_n - \mu_z)^2}{2\sigma^2}\right). \tag{15}$$

Changing μ_z smoothly then causes a smooth change in the weights. It is, however, easy to see that the intrinsic mean is given by

$$\mu = \mathrm{sign}(\mu_z) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{16}$$

which changes discontinuously with μ_z. This is illustrated in Fig. 2.

References

[1] T. Hastie and W. Stuetzle, "Principal curves," *Journal of the American Statistical Association*, vol. 84, no. 406, pp. 502–516, 1989.