Differentially Private Geodesic Regression

Aditya Kulkarni*, Carlos Soto

Department of Physics, UMASS Amherst

Department of Mathematics and Statistics, UMASS Amherst

February 9, 2026

MOTIVATION

- A manifold $\mathcal{M}$ is a topological space that is locally equivalent to Euclidean space $\mathbb{R}^n$ (eg: 2D Sphere $\mathbb{S}^2$). A smooth manifold that admits a Riemannian metric ($g_p(U,V) = g_p(V,U)$, $g_p(U,U) \geq 0$, $U, V \in T_p\mathcal{M}$) is called a Riemannian manifold.
- In modern statistical practices, it is common to encounter data that inherently live in curved, non-Euclidean spaces such as spherical data $\mathbb{S}^n$ (e.g., directional wind data, spatial data), symmetric positive definite matrices $\mathrm{SPD}(n)$(e.g., covariance matrices, brain tensors), Kendall shape space $\mathbb{C}P^{k-2}$ (Corpus callosum images).
- **Geodesic regression** (Fletcher 2011) provides a framework for modeling the relationship between a real-valued independent variable and a manifold-valued dependent variable.

$$E(p, v) = \frac{1}{2\,n} \sum_{i=1}^{n} d(\mathrm{Exp}(p, x_i\, v), y_i)^2, \tag{1}$$

$$(\hat{p}, \hat{v}) = \mathrm{argmin}_{(p,v)}\ E(p, v). \tag{2}$$

DP Geodesic Regression

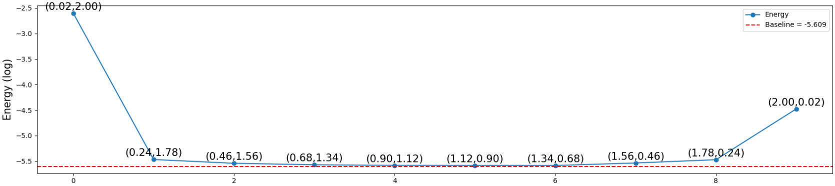
- A randomised mechanism $f(z;D)$ is said to satisfy *pure differential privacy*(Dwork et al. 2006) if $P(f(z;D) \in A) \leq \exp(\epsilon) P(f(z;D') \in A)$, for given privacy budget $\epsilon > 0$, all $D \sim D'$, and A is any measurable set in $\mathcal{M}$.
- We use the K-norm gradient mechanism on manifolds which takes the form: $f(z;D) \propto \exp\{-\sigma^{-1} \|\nabla E(z;D)\|_z\}$.
- KNG satisfies pure differential privacy when the noise scale is given by $\sigma = \Delta/\epsilon$ where $\Delta = \sup_{D\sim D'} \|\nabla E(z;D) - \nabla E(z;D')\|_z$ is the global sensitivity.

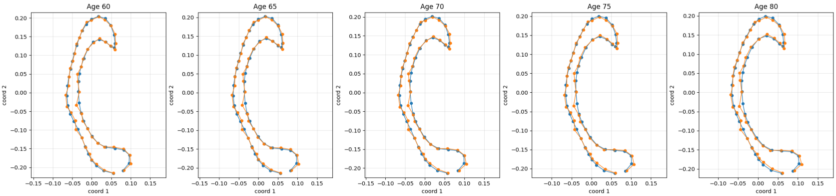
Table 1. Bounds on Δ_p and $\Delta_{\vec{v}}$.

Sensitivity	$\kappa_l \geq 0$	$\kappa_l < 0$
Δ_p	$\frac{2\tau}{n}$	$\frac{2\tau}{n} \cosh\left(2\sqrt{-\kappa_l}(\tau_m + \tau)\right)$
$\Delta_{\vec{v}}$	$\frac{2\tau}{n}$	$\frac{\tau}{n} \frac{\sinh(2\sqrt{-\kappa_l}(\tau_m + \tau))}{\sqrt{-\kappa_l}(\tau_m + \tau)}$

Example Results - Kendal Shape Space

- We analyze corpus callosum shapes from the Alzheimer's Disease Neuroimaging Initiative (ADNI) (Cornea et al. 2017) which lie on Kendall Shape Space $\mathbb{C}P^{k-2}$ with $k = 50$ landmarks.





The independent variable is age and the dependent is the corpus callosum image.

Selected References

Cornea, Emil et al. (2017). "Regression models on Riemannian symmetric spaces". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 79.2, pp. 463–482.

Dwork, Cynthia et al. (2006). "Calibrating noise to sensitivity in private data analysis". In: *Theory of cryptography conference*. Springer, pp. 265–284.

Fletcher, Thomas (2011). "Geodesic regression on Riemannian manifolds". In: *Proceedings of the Third International Workshop on Mathematical Foundations of Computational Anatomy-Geometrical and Statistical Methods for Modelling Biological Shape Variability*, pp. 75–86.

Thank You For Listening