Asymptotic Consistency of a Karcher Mean Estimator for the Latent Deformation Model

Technical Note

Cody Carroll

Feb. 2026

Setup and Motivation

Carroll and Müller (2023) propose the latent deformation model (LDM)

$$X_{ij}(t) = A_{ij}\,(\lambda \circ \Psi_j \circ H_i)(t), \qquad i = 1, \dots, n, \quad j = 1, \dots, p, \tag{1}$$

where λ is the latent curve, $\Psi_j \in \mathcal{W}$ are component deformation functions, and $H_i \in \mathcal{W}$ are subject-level warping functions. The following standardization conditions are assumed:

$$\frac{1}{p}\sum_{j=1}^{p} \Psi_j^{-1}(t) = t \quad \forall\, t \in \mathcal{T}, \tag{2}$$

$$\mathbb{E}\, H_i^{-1}(t) = t \quad \forall\, t \in \mathcal{T}. \tag{3}$$

Defining the jth *component tempo* $\gamma_j = \lambda \circ \Psi_j$, the normalized curves satisfy

$$X_{ij}^*(t) := \frac{X_{ij}(t)}{\|X_{ij}\|_\infty} = (\gamma_j \circ H_i)(t),$$

where the last equality uses the standard normalization $\|\lambda\|_\infty = 1$ (which implies $\|\gamma_j\|_\infty = 1$ for all j, since warpings preserve the sup-norm). This is a univariate warping problem with template γ_j for each j.

Two estimators of λ

1. A global alignment estimator

For each subject i, select one component $Z_i = X_{i,J_i}$ uniformly at random ($J_i \in \{1, \ldots, p\}$, each with probability $1/p$). Align $\{Z_i^*\}_{i=1}^n$ via pairwise warping and average the aligned curves:

$$\hat{\lambda}_{\text{global}} = n^{-1} \sum_{i=1}^{n} (Z_i \circ \hat{D}_i^{-1}) \,/\, \|Z_i\|_\infty.$$

Theorem 1c of Carroll and Müller (2023) establishes

$$\sup_{t \in \mathcal{T}} |\hat{\lambda}_{\text{global}}(t) - \lambda(t)| = \mathcal{O}_P(n^{-1/2}) + \mathcal{O}_P(\tau_m^{1/2}) + \mathcal{O}(\eta_1^{1/2}).$$

2. A component tempo Karcher mean estimator of λ

The code implementation of Carroll and Müller (2023) (`https://github.com/codycarroll/LDM`) also provides an option for a heuristic estimator of λ as follows. After obtaining $\hat{\gamma}_j$ (the estimated component tempo for each j, averaged over aligned subjects), let $F : (L^\infty(\mathcal{T}))^p \to L^\infty(\mathcal{T})$ denote the Karcher mean functional, so that we can define $\hat{\lambda}_{\text{heuristic}} = F(\hat{\gamma}_1, \ldots, \hat{\gamma}_p)$ as the Karcher mean of the estimated component tempos. Consistency of this estimator is not established in Carroll and Müller (2023) but the goal of this note is to sketch an argument showing that $\hat{\lambda}_{\text{heuristic}}$ achieves the same parametric rate.

Population Level Identification

We first show that at the population level (i.e., replacing $\hat{\gamma}_j$ with the true γ_j), the Karcher mean recovers λ exactly.

Let $\mathcal{Q}(f)(t) = \text{sgn}(f'(t))\sqrt{|f'(t)|}$ denote the square-root velocity (SRV) transform of a function $f \in \mathcal{W}$. The elastic (Fisher–Rao) distance between $f, g \in L^2(\mathcal{T})$ is

$$d_F(f, g) = \inf_{\varphi \in \mathcal{W}} \|\mathcal{Q}(f) - \mathcal{Q}(g \circ \varphi)\|_{L^2}.$$

Two functions f and g satisfy $d_F(f, g) = 0$ if and only if $g = f \circ \varphi$ for some $\varphi \in \mathcal{W}$; they are in the same equivalence class $[f]$ in the quotient space $L^2(\mathcal{T})/\mathcal{W}$.

Lemma 1 (Population identification). *Suppose λ satisfies Assumption (L1) of Carroll and Müller (2023), and the standardization condition* (2) *holds. Then the SRVF Karcher mean of $\{\gamma_j = \lambda \circ \Psi_j\}_{j=1}^p$ is λ.*

Proof sketch.

Since each $\gamma_j = \lambda \circ \Psi_j$, we have $d_F(\gamma_j, \lambda) = 0$ for all j: the infimum in the elastic distance is achieved by the warping $\varphi = \Psi_j^{-1}$, which aligns γ_j exactly to λ. Hence all γ_j belong to the same equivalence class $[\lambda]$ in $L^2(\mathcal{T})/\mathcal{W}$, and the Karcher variance $\frac{1}{p}\sum_j d_F(\cdot, \gamma_j)^2$ equals zero for every $\mu \in [\lambda]$.

We verify directly that λ is a fixed point of the Karcher mean iteration. Setting $\mu = \lambda$, the optimal warping aligning γ_j to μ is $\varphi_j^* = \Psi_j^{-1}$, since $\gamma_j(\Psi_j^{-1}(t)) = (\lambda \circ \Psi_j \circ \Psi_j^{-1})(t) = \lambda(t) = \mu(t)$. The updated mean is then

$$\mu^{\text{new}} = \frac{1}{p}\sum_{j=1}^{p} \gamma_j(\Psi_j^{-1}(t)) = \frac{1}{p}\sum_{j=1}^{p}(\lambda \circ \Psi_j \circ \Psi_j^{-1})(t) = \frac{1}{p}\sum_{j=1}^{p}\lambda(t) = \lambda(t),$$

confirming λ is a fixed point.

For uniqueness: since every $\mu = \lambda \circ \Psi_*$ in $[\lambda]$ achieves Karcher variance zero, the Karcher mean is only determined up to reparametrization within $[\lambda]$ without further constraint. The standardization condition (2) resolves this: it specifies that the Ψ_j^{-1} average to the identity, which identifies λ as the unique representative of $[\lambda]$ consistent with the standardization. Any other representative $\lambda \circ \Psi_*$ with $\Psi_* \neq \text{id}$ would require the optimal alignment warpings $\varphi_j^* = \Psi_j^{-1} \circ \Psi_*$ (which satisfy $\gamma_j(\Psi_j^{-1}(\Psi_*(t))) = \lambda(\Psi_*(t)) = \mu(t)$) to average to Ψ_*, violating (2). Under (L1), which ensures λ is non-degenerate, this representative is unique. $\square$

Asymptotic Consistency

Lemma 1 shows that the target of the heuristic estimator is λ. It remains to transfer the convergence of $\hat{\gamma}_j \to \gamma_j$ to $\hat{\lambda}_{\text{heuristic}} \to \lambda$.

Let $F : (L^\infty(\mathcal{T}))^p \to L^\infty(\mathcal{T})$ denote the Karcher mean functional, so that $F(\gamma_1, \ldots, \gamma_p) = \lambda$ by Lemma 1 and $\hat{\lambda}_{\text{heuristic}} = F(\hat{\gamma}_1, \ldots, \hat{\gamma}_p)$.

Theorem 2 (Consistency of the heuristic estimator)**.** *Under Assumptions (L1), (L2), and (S0–S2) of Carroll and Müller (2023), with* $\tau_m = m^{-(1-\delta)/3}$ *for arbitrarily small* $\delta > 0$,

$$\sup_{t\in\mathcal{T}} |\hat{\lambda}_{\text{heuristic}}(t) - \lambda(t)| = \mathcal{O}_P(n^{-1/2}) + \mathcal{O}_P(\tau_m^{1/2}) + \mathcal{O}(\eta_1^{1/2}).$$

Proof sketch. By Theorem 1b of Carroll and Müller (2023), for each $j = 1, \ldots, p$,

$$\sup_{t\in\mathcal{T}} |\hat{\gamma}_j(t) - \gamma_j(t)| = \mathcal{O}_P(n^{-1/2}) + \mathcal{O}_P(\tau_m^{1/2}) + \mathcal{O}(\eta_1^{1/2}). \tag{4}$$

We claim that F is Lipschitz continuous in a neighborhood of $(\gamma_1, \ldots, \gamma_p)$ with respect to the sup-norm: there exists $C > 0$ such that for all $(f_1, \ldots, f_p)$ sufficiently close to $(\gamma_1, \ldots, \gamma_p)$,

$$\|F(f_1, \ldots, f_p) - F(\gamma_1, \ldots, \gamma_p)\|_\infty \le C \max_{1\le j\le p} \|f_j - \gamma_j\|_\infty. \tag{5}$$

This follows from the implicit function theorem applied to the fixed-point characterization of the Karcher mean: the fixed point equation $\mu = (1/p)\sum_j f_j \circ \varphi_j^*(f_j, \mu)$ defines μ as a smooth function of $(f_1, \ldots, f_p)$ near $(\gamma_1, \ldots, \gamma_p, \lambda)$, provided the derivative of the fixed-point map is non-singular. Non-singularity follows from (L1), which ensures the Hessian of the Karcher variance is positive definite at λ (see, e.g., Bhattacharya & Patrangenaru 2003 for the general Riemannian case; Srivastava & Klassen 2016 for the SRVF setting).

Combining (4) and (5), the rate immediately follows:

$$\begin{aligned}
\sup_{t\in\mathcal{T}} |\hat{\lambda}_{\text{heuristic}}(t) - \lambda(t)| &= \|F(\hat{\gamma}_1, \ldots, \hat{\gamma}_p) - F(\gamma_1, \ldots, \gamma_p)\|_\infty \\
&\le C \max_{1\le j\le p} \|\hat{\gamma}_j - \gamma_j\|_\infty \\
&= \mathcal{O}_P(n^{-1/2}) + \mathcal{O}_P(\tau_m^{1/2}) + \mathcal{O}(\eta_1^{1/2}).
\end{aligned}$$

□

Corollary 3. *Under the assumptions of Theorem 2, if additionally* $\eta_1 \sim \mathcal{O}(n^{-1})$ *and* $m \gtrsim n^{\Delta(1-\delta)^{-1}}$ *for some* $\Delta > 3$, *then*

$$\sup_{t\in\mathcal{T}} |\hat{\lambda}_{\text{heuristic}}(t) - \lambda(t)| = \mathcal{O}_P(n^{-1/2}).$$

That is, the heuristic achieves the same parametric rate as the global alignment estimator.

Discussion

The argument above is largely a consequence of Theorem 1b of Carroll and Müller (2023) plus a continuity property of the Karcher mean functional. The main new ingredient needed for a complete proof was a precise statement of the Lipschitz continuity of the SRVF Karcher mean functional F at the point $(\gamma_1, \ldots, \gamma_p)$. This is a regularity result for the Karcher mean on the infinite-dimensional manifold $(L^2(\mathcal{T})/\mathcal{W}, d_F)$. Results of this type are available in finite dimensions (Bhattacharya & Patrangenaru 2003) and have been extended to function spaces in the SRVF context (Kurtek & Bharath 2015; Srivastava & Klassen 2016), so the main work is verifying the required curvature conditions hold under (L1).

References

Carroll, C. & Müller, H.-G. (2023). Latent deformation models for multivariate functional data and time-warping separability. *Biometrics*, 79(4), 3345–3358.

Bhattacharya, R. & Patrangenaru, V. (2003). Large sample theory of intrinsic and extrinsic sample means on manifolds. *Annals of Statistics*, 31(1), 1–29.

Kurtek, S. & Bharath, K. (2015). Bayesian sensitivity analysis with the Fisher–Rao metric. *Biometrika*, 102(3), 601–616.

Srivastava, A. & Klassen, E. (2016). *Functional and Shape Data Analysis*. Springer.