SUPPLEMENT TO "PRINCIPAL COMPONENT ANALYSIS FOR FUNCTIONAL DATA ON RIEMANNIAN MANIFOLDS AND SPHERES"

By Xiongtao Dai and Hans-Georg Müller

University of California, Davis

S1. Additional Proofs.

Proof of Corollary 2. We have

$$\begin{aligned}\sup_{t\in\mathcal{T}}\|\hat{V}_i(t)-V_i(t)\|_E &= \sup_{t\in\mathcal{T}}\|\log_{\mu_{\mathcal{M}}(t)}(X_i(t))-\log_{\hat{\mu}_{\mathcal{M}}(t)}(X_i(t))\|_E\\ &\lesssim \sup_{t\in\mathcal{T}}|d_{\mathcal{M}}(\hat{\mu}_{\mathcal{M}}(t),\mu_{\mathcal{M}}(t))|,\end{aligned}$$

where the last inequality is due to (B3) and the fact that $\log_p(q)$ is continuously differentiable in (p,q) (Theorem I.3.2 in Chavel 2006). □

Proof of Theorem 2. Denote $\tilde{G}(t,s)=\frac{1}{n}\sum_{i=1}^n V_i(t)V_i(s)^T$. Then

$$\begin{aligned}\sup_{t,s\in\mathcal{T}}\left\|\hat{G}(t,s)-G(t,s)\right\|_F &\le \sup_{t,s\in\mathcal{T}}\left\|\hat{G}(t,s)-\tilde{G}(t,s)\right\|_F+\sup_{t,s\in\mathcal{T}}\left\|\tilde{G}(t,s)-G(t,s)\right\|_F\\ &\le \frac{1}{n}\sum_{i=1}^n\sup_{t,s\in\mathcal{T}}\left\|\hat{V}_i(t)\hat{V}_i(s)^T-V_i(t)V_i(s)^T\right\|_F+\sup_{t,s\in\mathcal{T}}\left\|\frac{1}{n}\sum_{i=1}^n V_i(t)V_i(s)^T-G(t,s)\right\|_F\end{aligned} \tag{S1}$$

Since $\sup_{t,s\in\mathcal{T}}\left\|V_i(t)V_i(s)^T\right\|_F<R^2$, viewing $V_i(t)V_i(s)^T$ as random elements in $L_\infty(\mathcal{T}\times\mathcal{T},\mathbb{R}^{d^2})$ the second term is $O_p(n^{-1/2})$ by Theorem 2.8 in Bosq (2000). For the first term, note

$$\begin{aligned}\left\|\hat{V}_i(t)\hat{V}_i(s)^T-V_i(t)V_i(s)^T\right\|_F &\le \left\|(\hat{V}_i(t)-V_i(t))\hat{V}_i(s)^T\right\|_F+\left\|V_i(t)(\hat{V}_i(s)-V_i(s))^T\right\|_F\\ &\le \|\hat{V}_i(s)\|_E\|\hat{V}_i(t)-V_i(t)\|_E+\|V_i(t)\|_E\|\hat{V}_i(s)-V_i(s)\|_E\\ &\lesssim \sup_{t\in\mathcal{T}} d_{\mathcal{M}}(\hat{\mu}_{\mathcal{M}}(t),\mu_{\mathcal{M}}(t)),\end{aligned}$$

where the second inequality is due to the properties of the Frobenius norm, and the last is due to Corollary 2 and (B3). Therefore, by Corollary 1 the first term in (S1) is $O_p(n^{-1/2})$ and (20) follows. Result (21) follows from applying Theorem 4.2.8 in Hsing and Eubank (2015) and from the fact that the operator norm is dominated by the Hilbert-Schmidt norm.

To prove (22), Theorem 5.1.8 in Hsing and Eubank (2015) and Bessel's inequality imply

$$\left\|\hat{\phi}_k - \phi_k\right\| = O_p(n^{-1/2}). \tag{S2}$$

Then note that for any $t \in \mathcal{T}$,

$$\begin{aligned}
\|\hat{\phi}_k(t) - \phi_k(t)\|_E &= \|\int \frac{1}{\hat{\lambda}_k}\hat{G}(t,s)\hat{\phi}_k(s)ds - \int \frac{1}{\lambda_k}G(t,s)\phi_k(s)ds\|_E \\
&= \|(\frac{1}{\hat{\lambda}_k} - \frac{1}{\lambda_k})\int \hat{G}(t,s)\hat{\phi}_k(s)ds + \frac{1}{\lambda_k}\int \hat{G}(t,s)(\hat{\phi}_k(s) - \phi_k(s)) \\
&+ (\hat{G}(t,s) - G(t,s))\phi_k(s)ds\|_E \\
&= O_p\left(\left|\frac{1}{\hat{\lambda}_k} - \frac{1}{\lambda_k}\right| + \left\|\hat{\phi}_k - \phi_k\right\| + \sup_{t,s\in\mathcal{T}}\left\|\hat{G}(t,s) - G(t,s)\right\|_F\right),
\end{aligned}$$

which is of order $O_p(n^{-1/2})$ by (21), (S2), and (20). Since the r.h.s. does not involve t, taking suprema on both sides over $t \in \mathcal{T}$ concludes the proof. □

Proof of Corollary 4. Conditions (A1)–(A2) and (B3) hold for longitudinal compositional data analysis, while (B2) holds by Theorem 2.1 in Afsari (2011). □

S2. Algorithms for the RFPCA of Compositional Data. The following Algorithms 1–3 are provided as examples for RFPCA applied to longitudinal compositional data $Z(t)$ or spherical trajectories $X(t)$. For longitudinal compositional data $Z(t)$, we initialize by defining $X(t)$ as the componentwise square root of $Z(t)$, which then lies on a Euclidean sphere S^d. We assume the trajectories are observed at $t = t_j = (j-1)/(m-1)$ for $j = 1, \ldots, m$, and all vectors are by default column vectors. Very similar algorithms for SFPCA have also been proposed by Anirudh et al. (2017).

The time complexity for Algorithm 1 is $O(nmf(d) + nm^2d^2 + (md)^3)$, where $f(d)$ is the cost for calculating a Fréchet mean in Line 2, which is typically $O(nd)$ or $O(nd^2)$ for gradient descent or quasi-Newton type optimizers per iteration, respectively. The most demanding computational step for the multivariate FPCA is $O(nm^2d^2)$ for Line 7 and $O((md)^3)$ for the eigendecomposition in Line 8. The computational cost for Algorithm 2 is $O(md)$ and that for Algorithm 3 is $O(nmd)$.

Algorithm 1: Spherical functional principal component analysis (SFPCA)

Data: S^d-valued trajectories $X_1(t), \dots, X_n(t)$
Result: $\hat{\mu}_{\mathcal{M}}(t)$, $\hat{V}_i(t)$, $\hat{\xi}_{ik}$, $\hat{\phi}_k(t)$, $\hat{\lambda}_k$, for $i = 1, \dots, n$ and $k = 1, \dots, K$
`// Obtain the intrinsic mean function and tangent vectors`
for $j \in \{1, \dots, m\}$ **do**
  $\hat{\mu}_{\mathcal{M}}(t_j) \leftarrow \arg\min_{p \in S^d} n^{-1} \sum_{i=1}^{n} [\cos^{-1}(p^T X_i(t_j))]^2$
  for $i \in \{1, \dots, n\}$ **do**
    $\hat{V}_i(t_j) = \frac{u}{\sqrt{u^T u}} \cos^{-1}(\hat{\mu}_{\mathcal{M}}(t_j)^T X_i(t_j))$, where $u = X_i(t_j) - (\hat{\mu}_{\mathcal{M}}(t_j)^T X_i(t_j))\mu_{\mathcal{M}}(t_j)$
  end
end
`// A multivariate FPCA. Vec(A) stacks the columns of A.`
$\hat{\mathbf{V}}_i \leftarrow [\hat{V}_i(t_1), \dots, \hat{V}_i(t_m)]^T$, $\hat{\mathbf{G}} \leftarrow n^{-1} \sum_{i=1}^{n} \mathtt{Vec}(\hat{\mathbf{V}}_i)\mathtt{Vec}(\hat{\mathbf{V}}_i)^T$
$[\omega, \mathbf{\Psi}] \leftarrow \mathtt{Eigen}(\hat{\mathbf{V}}_i)$, for eigenvalues $\omega = [\omega_1, \dots, \omega_m]^T$ and eigenvectors $\mathbf{\Psi} = [\psi_1, \dots, \psi_m]$
for $k \in \{1, \dots, K\}$ **do**
  Write $\hat{\mathbf{\Phi}}_k = [\hat{\phi}_k(t_1), \dots, \hat{\phi}_k(t_m)]^T$, $\mathtt{Vec}(\hat{\mathbf{\Phi}}_k) \leftarrow m^{1/2}\psi_k$, $\hat{\lambda}_k \leftarrow m^{-1}\omega_k$, $\hat{\xi}_{ik} \leftarrow m^{-1}\mathtt{Vec}(\hat{\mathbf{V}}_i)^T \mathtt{Vec}(\hat{\mathbf{\Phi}}_k)$
end

Algorithm 2: Truncated K-dimensional representations

Data: $\hat{\mu}_{\mathcal{M}}(t)$, $\{(\hat{\xi}_{ik}, \hat{\phi}_k(t))\}_{k=1}^{K}$
Result: $\hat{X}_{iK}(t)$, $\hat{V}_{iK}(t)$
for $j \in \{1, \dots, m\}$ **do**
  $\hat{V}_{iK}(t_j) \leftarrow \sum_{k=1}^{K} \hat{\xi}_{ik}\hat{\phi}_k(t_j)$
  $\hat{X}_{iK}(t_j) \leftarrow \cos(\|\hat{V}_{iK}(t_j)\|_E)\hat{\mu}_{\mathcal{M}}(t) + \sin(\|\hat{V}_{iK}(t_j)\|_E)\|\hat{V}_{iK}(t_j)\|_E^{-1}\hat{V}_{iK}(t_j)$
end

Algorithm 3: Calculate FVE

Data: Outputs from Algorithm 1
Result: $\widehat{\mathrm{FVE}}_K$
$\hat{U}_0 \leftarrow n^{-1} \sum_{i=1}^{n} \int_{\mathcal{T}} d_{\mathcal{M}}{}^2(X_i(t), \hat{\mu}_{\mathcal{M}}(t))dt$
for $i \in \{1, \dots, n\}$ **do**
  Use Algorithm 2 to obtain $\hat{X}_{iK}(t)$
end
$\hat{U}_K \leftarrow n^{-1} \sum_{i=1}^{n} \int_{\mathcal{T}} d_{\mathcal{M}}{}^2(X_i(t), \hat{X}_{iK}(t))dt$
$\widehat{\mathrm{FVE}}_K = (\hat{U}_0 - \hat{U}_K)/\hat{U}_0$

S3. Additional simulations. We conducted an additional simulation study to investigate the scalability of the RFPCA algorithms to higher dimensions d, on the unit sphere $\mathcal{M} = S^d$ in $\mathbb{R}^{d+1}$ for $d = 5, 10, 15, 20$. Table S1 shows that the RFPCA scales well for larger dimensions in terms of running time, and its relative disadvantage in speed as compared to the L^2 FPCA becomes smaller as d and n get larger.

The samples were generated in the same fashion as in the main text, except for the mean function $\mu_{\mathcal{M}}(t) = \exp_{p_0}(2(d-1)^{-1/2}t, \ldots, 2(d-1)^{-1/2}t, 0.3\pi\sin(\pi t), 0)$, and eigenfunctions $\phi_k(t) = d^{-1/2}R_t[\zeta_k(t/d), \ldots, \zeta_k(t/d+(d-1)/d), 0]^T$, where $p_0 = [0, \ldots, 0, 1]$ and R_t is the rotation matrix from p_0 to $\mu_{\mathcal{M}}(t)$.

TABLE S1

A comparison of mean running time for S^d. The standard errors are below 2% of the means.

	$n=50$				$n=100$				$n=200$				$n=400$			
d	5	10	15	20	5	10	15	20	5	10	15	20	5	10	15	20
RFPCA	1.3	1.7	2.1	2.7	1.9	2.6	3.3	4.3	3.1	4.6	5.7	7.4	5.9	7.8	10.5	13.2
L^2 FPCA	0.4	0.7	1.0	1.5	0.8	1.2	1.8	2.4	1.4	2.5	3.5	4.4	3.0	4.4	6.3	8.2

References.

AFSARI, B. (2011). Riemannian L^p center of mass: Existence, uniqueness, and convexity. *Proceedings of the American Mathematical Society* **139** 655–673.

ANIRUDH, R., TURAGA, P., SU, J. and SRIVASTAVA, A. (2017). Elastic functional coding of Riemannian trajectories. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **39** 922–936.

BOSQ, D. (2000). *Linear Processes in Function Spaces: Theory and Applications.* Springer-Verlag, New York.

CHAVEL, I. (2006). *Riemannian Geometry: A Modern Introduction.* Cambridge University Press, Cambridge.

HSING, T. and EUBANK, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators.* Wiley, Hoboken.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA, DAVIS
DAVIS, CA 95616 USA
E-MAIL: dai@ucdavis.edu
E-MAIL: hgmueller@ucdavis.edu