Functional models for time-varying random objects

Paromita Dubey and Hans-Georg Müller

University of California at Davis, USA

[Read before The Royal Statistical Society at a meeting organized by the Research Section on Wednesday, October 16th, 2019, Professor A. Doucet in the Chair]

Summary. Functional data analysis provides a popular toolbox of functional models for the analysis of samples of random functions that are real valued. In recent years, samples of time-varying object data such as time-varying networks that are not in a vector space have been increasingly collected. These data can be viewed as elements of a general metric space that lacks local or global linear structure and therefore common approaches that have been used with great success for the analysis of functional data, such as functional principal component analysis, cannot be applied. We propose metric covariance, a novel association measure for paired object data lying in a metric space \((\Omega, d)\) that we use to define a metric autocovariance function for a sample of random \(\Omega\)-valued curves, where \(\Omega\) generally will not have a vector space or manifold structure. The proposed metric autocovariance function is non-negative definite when the squared semimetric \(d^2\) is of negative type. Then the eigenfunctions of the linear operator with the autocovariance function as kernel can be used as building blocks for an object functional principal component analysis for \(\Omega\)-valued functional data, including time-varying probability distributions, covariance matrices and time dynamic networks. Analogues of functional principal components for time-varying objects are obtained by applying Fréchet means and projections of distance functions of the random object trajectories in the directions of the eigenfunctions, leading to real-valued Fréchet scores. Using the notion of generalized Fréchet integrals, we construct object functional principal components that lie in the metric space \(\Omega\). We establish asymptotic consistency of the sample-based estimators for the corresponding population targets under mild metric entropy conditions on \(\Omega\) and continuity of the \(\Omega\)-valued random curves. These concepts are illustrated with samples of time-varying probability distributions for human mortality, time-varying covariance matrices derived from trading patterns and time-varying networks that arise from New York taxi trips.

Keywords: Dynamic networks; Fréchet integral; Functional data analysis; Metric covariance; Object data; Principal component analysis; Stochastic processes

1. Introduction

Time-varying data where one collects a sample of independent and identically distributed random functions, which take values in a general object space that does not have a linear structure, are increasingly common, whereas the statistical methodology for the analysis of such data has been lagging behind. We aim to introduce techniques that will help to fill this gap. For the case where observations consist of samples of random trajectories that take values in \(\mathbb{R}^p\), the methodology of choice is often functional data analysis (FDA) (Ramsay and Silverman, 2005; Horvath and Kokoszka, 2012; Wang et al., 2016), where methodology for one-dimensional \((p = 1)\) functional data is readily available. Models for functional data that consist of vector-valued processes \((p > 1)\) have been studied more recently (Zhou et al., 2008; Berrendero et al., 2011; Chiou et al., 2014; Claeskens et al., 2014; Verbeke et al., 2014; Chiou et al., 2016) as well as

Address for correspondence: Paromita Dubey, Department of Statistics, University of California at Davis, 399 Crocker Lane, Davis, CA 95616, USA.

E-mail: pdubey@ucdavis.edu

© 2019 Royal Statistical Society
the case where at each time point one records a random function, i.e. function-valued stochastic processes (Park and Staicu, 2015; Chen and Müller, 2012; Chen et al., 2017). In these models, the responses are in a linear space, either the Euclidean space $\mathbb{R}^p$ or the Hilbert space $L^2$. Functions of objects in spaces that can be locally approximated by linear spaces such as Riemannian manifolds such as spheres have also been considered more recently (Lin et al., 2017; Dai and Müller, 2018). The major objective of this paper is to overcome the global or local linearity assumptions that are inherent in these previous approaches. The challenge is that existing FDA methodology relies on vector operations and inner products, which are no longer available.

Functional principal component analysis (FPCA) (Kleffe, 1973; Dauxois et al., 1982) has emerged as the method of choice to represent and interpret samples of random functions that take values in linear spaces. It also provides dimension reduction by expanding an underlying random process into the basis functions given by the eigenfunctions of the autocovariance operator and then truncating this expansion at a finite number of expansion terms. A related tool is the modes of variation, which enable exploration of the effects of single eigendirections (Castro et al., 1986; Jones and Rice, 1992) and are useful in practical applications (Dong et al., 2018). FPCA also provides a starting point for many theoretical investigations and FDA techniques such as functional clustering (Chiou and Li, 2007; Jacques and Preda, 2014; Suarez and Ghosal, 2016) or regression and classification (Yao et al., 2005a; Dai et al., 2017).

As we enter the era of ‘big data’, it has become increasingly common to observe more complex, often non-Euclidean, data on a time grid. Technological advances have made it possible to record and store efficiently time courses of image, network, sensor or other complex data. For example, neuroscientists are interested in dynamic functional connectivity, where one essentially observes samples of time-varying covariance or correlation matrices obtained from functional magnetic resonance imaging data for each subject in a sample. Time-varying network data arise in various forms, e.g. road or Internet traffic networks or time evolving social networks, and it is of interest to extract structure and patterns from such data.

To obtain efficient and interpretable summaries of the information that is contained in samples of complex observations is a major task in modern statistics that has led for example to the development of methods such as geodesic principal component analysis in the space of probability distributions on $\mathbb{R}$ (Bigot et al., 2017) and on more general Hilbert spaces (Seguy and Cuturi, 2015) that utilize optimal transport geometry and geodesic curves under the Wasserstein metric. These approaches utilize geodesics to connect the random distributions with the Wasserstein barycentres. We aim here at identifying dominant directions of variation in a sample of time-varying random object trajectories, where the random objects are indexed by time and are in a general metric space. The time-varying aspect provides for a novel and little-explored setting, and to develop tools that are supported by theory and are useful for the further exploration and analysis of such data is the main motivation for this paper.

Although FPCA for samples of functions taking values in smooth Riemannian manifolds has been studied both practically and theoretically (Anirudh et al., 2017; Dai and Müller, 2018), these approaches critically depend on the local Euclidean property of Riemannian manifolds and cannot be extended to functional data objects that take values in more general metric spaces that do not have a tractable and relatively simple Riemannian geometry. FPCA for doubly functional data, where the observations at each time point are functions rather than scalars (Chen et al., 2017), is based on a tensor product representation of the underlying function-valued stochastic process. The functions need to be Hilbert space valued, so this approach cannot be applied to non-Hilbertian data. Because of the lack of linear structure, developing a form of FPCA for random functions taking values in a metric space, which we refer to as \textit{functional random objects}, is a major challenge.
Consider a totally bounded metric space \((\Omega, d)\) and a random sample of fully observed \(\Omega\)-valued functional data. Aiming to extend key tools of FDA to cover such data, we first revisit the well-established FPCA for the case of real-valued functional data. The essence of FPCA is contained in the autocovariance structure of the underlying random functions, i.e. their covariance at different time points. This leads to the question how to quantify correlation between random objects in general metric spaces that correspond to the values of the random function at different time points. An example for such an extension of Pearson correlation to the case of multivariate data is the RV-coefficient (Robert and Escoufier, 1976), which is 0 if all the vector components are uncorrelated and is strictly positive otherwise.

In this paper we introduce metric covariance, which is a novel association measure for paired data in general metric spaces. Metric covariance differs in key aspects from distance correlation, which is another measure of dependence between metric space data (Lyons, 2013; Székely and Rizzo, 2017), the latter being primarily suited to measure probabilistic independence rather than for quantifying the strength of ‘positive’ or ‘negative’ association, which is the primary goal of the former. Unlike distance covariance, the magnitude of metric covariance quantifies the degree of association between paired data objects. The key component of FPCA is to decompose the variation in a sample of trajectories into orthogonal directions. An important difference between metric covariance and distance covariance, which is specifically relevant in this context, arises when considering the associated notion of variance. In contrast with distance correlation, metric covariance of a random object with itself leads to an interpretable notion of variance for data objects, as we shall demonstrate below. We also show that metric covariance is symmetric and non-negative definite whenever the squared distance \(d^2\) is a semimetric of negative type (Sejdinovic et al., 2013; Lyons, 2013; Schoenberg, 1938). The notion of metric correlation can then be easily derived from metric covariance and random objects will be considered to be uncorrelated if they have a metric correlation of size 0.

In FPCA for \(\mathbb{R}\)-valued functional data, once the autocovariance function has been determined, one defines a linear Hilbert–Schmidt operator whose eigenfunctions represent the orthonormal directions of variance for the functional data in the Hilbert space \(L^2\). The corresponding eigenvalues represent the fraction of variance explained by the respective FPCs, which are the lengths of the projections of the functional data in the direction of each eigenfunction. How can we extend these ideas to object-valued functional data, where we do not have a linear structure or inner product? We proceed by constructing a linear Hilbert–Schmidt operator by using the proposed metric covariance as its kernel and utilize its eigenfunctions and eigenvalues. For real-valued functional data, we obtain the FPCs by the Karhunen–Loève expansion of centred functional data in the eigenbasis, where the FPCs are the inner products of the centred functional data with respect to the eigenfunctions. Unfortunately it is not possible to ‘centre’ object functional data living in general metric spaces and also we do not have an inner product. In the case of FDA in the Hilbert space \(L^2\), the inner products can be expressed as integrals. Although due to the lack of linear structure there is no integral for functional random objects, the interpretation of inner products as integrals nevertheless provides a way forward that we develop in this paper. We propose two approaches for obtaining FPCs for object functional data: one in which the FPCs are scalar irrespective of the nature of the metric space in which the random objects live, and an alternative approach in which the FPCs themselves are random objects, i.e. \(\Omega\) valued.

To obtain FPCs in object space, we introduce the notion of a generalized Fréchet integral of an \(\Omega\)-valued curve with respect to a real-valued function, where this integral resides in \(\Omega\). Generalized Fréchet integrals depend on the underlying metric \(d\) in \(\Omega\) and are defined under the constraint that the real-valued function in the integrand integrates to 1. We draw inspiration
from the covariance integral for multivariate functional data that was previously introduced as a Fréchet integral (Petersen and Müller, 2016). This previous integral is a special case of the generalized Fréchet integral that is introduced here; it corresponds to the special case where $\Omega$ is the space of covariance matrices and the real-valued function in the integrand is the constant function 1. We demonstrate that the resulting object FPCs, which reside in $\Omega$, provide useful insights about the structure of the underlying functional random objects.

For an alternative scalar approach, we extract relatively simple characteristics from the object functional data. A first step is to define a ‘mean’ function by using the notion of Fréchet means (Fréchet, 1948). This mean function resides in the object function space and serves as a ‘central’ trajectory for the object functional data. To obtain a representative scalar FPC for a specific random object trajectory and eigenfunction, we utilize projections of the distance function between the specific random object trajectory and the Fréchet mean trajectory on each of the eigenfunctions. The resulting Fréchet scores encapsulate variation in the departures of functional random objects from the Fréchet mean trajectory. As we illustrate in simulations and data analysis, plotting these Fréchet scores against each other often illustrates meaningful patterns in the sample of object functional data that are generally difficult to capture visually, because of their complexity and non-linearity. For example, such plots can aid in detecting the presence of clusters or outliers in functional random objects.

In this paper, we have three major objectives. First, we lay out a framework for extending FPCA to general metric-space-valued functional data. The population target parameters are the metric autocovariance operator, its eigenvalues and eigenfunctions and the population Fréchet mean function, which are introduced in Section 2; additionally, the object FPCs, which are generalized Fréchet integrals and the Fréchet scores (Section 3). Second, we provide sample-based estimators of these population targets and establish their asymptotic properties under mild restrictions on the metric entropy of the metric space $\Omega$ and the continuity of the object functional data (Section 4). Proofs of all results are in section S1 of the on-line supplement. Third, we illustrate our results through simulations (Section 5) and various data examples (Section 6), which include samples of time-varying probability distributions of age at death obtained from human mortality data of 32 countries, time-varying yellow taxi trip networks of different regions in Manhattan observed daily during the year 2016, and of changing trade patterns between countries that can be represented as time-varying covariance matrices, followed by a brief discussion (Section 7).

2. Metric covariance

2.1. Covariance and correlation for random objects

We consider a totally bounded metric space $(\Omega, d)$ where $d$ is a metric and an $\Omega$-valued stochastic process $X = \{X(t)\}_{t \in [0,1]}$ on the interval $[0, 1]$. With $P$ denoting the probability measure of the random process $X$, we are given a sample $\{X_i = (X_i(t))_{t \in [0,1]} : i = 1, 2, \ldots, n\}$ of random $\Omega$-valued functions on $[0, 1]$ generated by $P$. The simplest case is $\Omega = \mathbb{R}$ with the intrinsic Euclidean metric, where $\{X_1, X_2, \ldots, X_n\}$ is a sample of real-valued functional data. For general metric spaces $\Omega$, we refer to $\{X_1, X_2, \ldots, X_n\}$ as a sample of functional random objects. Inspired by the approach to obtain FPCA for real-valued functional data, our first goal is to quantify the association between random objects $X(s)$ and $X(t)$ in $\Omega$, where $s$ and $t$ are two arbitrary points in the domain $[0, 1]$.

For motivation, consider first the case of real random variables $(U, V)$ with finite covariance. Imagine for a moment that we cannot add, subtract or multiply these random variables and are restricted to compute their distances $d_E(U, V) = |U - V|$. As is well known, we then can
write the variance of $U$ by using an independent and identically distributed copy $U'$ of $U$ by
\[
\text{var}(U) = \frac{1}{2} E\{d^2_E(U, U')\}.
\]

Interestingly, this non-algebraic construction can be extended to the covariance of $U, V$: let $(U', V')$ be an independent and identically distributed copy of $(U, V)$. We then obtain an alternative formulation of $\text{cov}(U, V)$ in terms of pairwise distances as follows:
\[
\text{cov}(U, V) = E\{(U - E(U))(V - E(V))\} = \frac{1}{2} E\{d^2_E(U, V') + d^2_E(U', V) - 2d^2_E(U, V)\}.
\]

If $(U, V)$ are $\mathbb{R}^d$-valued random variables with $d_E(\cdot, \cdot)$ denoting the Euclidean distance in $\mathbb{R}^d$, a simple calculation shows that, in this case,
\[
\frac{1}{2} E\{d^2_E(U, V') + d^2_E(U', V) - 2d^2_E(U, V)\} = E\{(U - E(U))^T (V - E(V))\},
\]
which is the inner product in the Hilbert space of $\mathbb{R}^d$-valued random variables with finite $E(U^T U)$. Next consider the case where $(U, V)$ are $\mathcal{H}$-valued random variables, where $\mathcal{H}$ is a Hilbert space and $d_E(\cdot, \cdot)$ is replaced by $d_{\mathcal{H}}(U, V) = \|U - V\|_{\mathcal{H}}$, the metric that arises from the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ of the Hilbert space. If the metric $d_{\mathcal{H}}(\cdot, \cdot)$ is bounded then $E(\|U\|_{\mathcal{H}}^2) < \infty$. One can show with some simple algebra and utilizing the Riesz representation theorem that
\[
\frac{1}{2} E\{d^2_{\mathcal{H}}(U, V') + d^2_{\mathcal{H}}(U', V) - 2d^2_{\mathcal{H}}(U, V)\} = E\{(U - E(U), V - E(V))_{\mathcal{H}}\},
\]
which is the inner product in $L^2(\mathcal{H})$: the Hilbert space of $\mathcal{H}$-valued random variables $U$ such that $E(\|U\|_{\mathcal{H}}^2) < \infty$.

What happens if $(U, V)$ are $\Omega$-valued random variables and we replace $d_{\mathcal{H}}$ by $d$ where $(\Omega, d)$ is a general metric space with no vector space structure to rely on? Or, for which spaces does the function $\frac{1}{2} E\{d^2(U, V') + d^2(U', V) - 2d^2(U, V)\}$ retain desirable properties? Proposition 3 of Sejdinovic et al. (2013) implies that, whenever $d$ is a semimetric of negative type, there is a Hilbert space $\mathcal{H}$ and an injective map, say $f : \Omega \rightarrow \mathcal{H}$, with
\[
d^2(U, V) = \|f(U) - f(V)\|_{\mathcal{H}}^2,
\]
and therefore it follows from equation (1) that, for some ‘remote’ Hilbert space $\mathcal{H}$ and the unknown map $f(\cdot)$,
\[
\frac{1}{2} E\{d^2(U, V') + d^2(U', V) - 2d^2(U, V)\} = E\{(f(U) - E\{f(U)\}, f(V) - E\{f(V)\})_{\mathcal{H}}\}.
\]

Here a space $(Z, \rho)$ with a semimetric $\rho$ is of negative type if for all $n \geq 2$, $z_1, z_2, \ldots, z_n \in Z$ and $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ with $\sum_{i=1}^n \alpha_i = 0$ we have
\[
\sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \rho(z_i, z_j) \leq 0.
\]

These considerations motivate the following definition of a generalized version of covariance $\text{cov}_\Omega(U, V)$ for paired random objects $(U, V)$ that take values in $\Omega \times \Omega$, where $(\Omega, d)$ is a metric space:
\[
\text{cov}_\Omega(U, V) = \frac{1}{2} E\{d^2(U, V') + d^2(U', V) - 2d^2(U, V)\},
\]
where as above $(U', V')$ is an independent and identically distributed copy of $(U, V)$. We refer to $\text{cov}_\Omega(U, V)$ as the metric covariance of $U$ and $V$. The metric covariance is always finite if the underlying metric space is bounded and coincides with the usual notion of covariance in Euclidean spaces.
We also define the metric correlation between two $\Omega$-valued random variables as

$$\rho_{\Omega}(U, V) = \frac{\text{cov}_{\Omega}(U, V)}{\sqrt{\text{cov}_{\Omega}(U, U) \text{cov}_{\Omega}(V, V)}}.$$ 

By the Cauchy–Schwarz inequality we have $-1 \leq \rho_{\Omega}(U, V) \leq 1$. Metric covariance or metric correlation depends on the choice of the metric $d$ and different choices of $d$ might reveal different aspects of association between random objects, depending on the underlying geometry of the metric.

### 2.2. Metric autocovariance operators

As in the real-valued Euclidean case, we define the metric autocovariance function $C(s, t)$ for functional random objects $\{X_1, X_2, \ldots, X_n\} \in \Omega$ as

$$C(s, t) = \text{cov}_{\Omega}\{X(s), X(t)\},$$

for all $(s, t) \in [0, 1] \times [0, 1]$. Obviously, $C(s, t)$ is a symmetric kernel and therefore has real eigenvalues when used as the kernel of a linear Hilbert–Schmidt operator. The following result shows that, for metric spaces $(\Omega, d)$ for which the squared distance function $d^2$ is of negative type, the metric autocovariance operator is positive semidefinite.

**Proposition 1.** If $d^2$ is of negative type, then $C(s, t)$ is a non-negative definite kernel.

By proposition 3 in Sejdinovic et al. (2013) and equation (3), $\text{cov}_{\Omega}(U, V) = 0$ implies that there is an abstract Hilbert space $\mathcal{H}$ and an injective map $f : \Omega \rightarrow \mathcal{H}$ such that $f(U)$ and $f(V)$ are orthogonal in $L^2(\mathcal{H})$. Note that $\text{var}_{\Omega}(U) = \text{cov}_{\Omega}(U, U) = \frac{1}{2}E\{d^2(U, U')\}$, which for real-valued random variables equals $\text{var}(U)$.

Formally, we can define the metric autocovariance operator as a linear Hilbert–Schmidt integral operator $T_C$ that operates on functions $g \in L^2(0, 1]$ and utilizes the metric autocovariance kernel,

$$(T_C g)(s) = \int_0^1 C(s, t) g(t) dt.$$ 

We note that for example theorem 4.6.4 of Hsing and Eubank (2015) implies the non-negative definiteness of the kernel $C(s, t)$, in the sense that $(T_C f, f) \geq 0$ for all $f$.

By Mercer’s theorem there is an orthonormal basis $\{\phi_j\}_{j=1}^{\infty}$ of $L^2([0, 1])$ consisting of eigenfunctions of $T_C$ such that the corresponding sequence of eigenvalues $\{\lambda_j\}_{j=1}^{\infty}$, which are ordered in declining order, is non-negative, since $C(s, t)$ is positive semidefinite. The eigenfunctions corresponding to non-zero eigenvalues are continuous on $[0, 1]$ and $C$ has the representation

$$C(s, t) = \sum_{j=1}^{\infty} \lambda_j \phi_j(s) \phi_j(t),$$

where the convergence is absolute and uniform; see, for example, lemma 4.6.1 and theorems 4.5.2, 4.6.2, 4.6.5 and 4.6.7 of Hsing and Eubank (2015).

We have thus accomplished the first step of extending FPCA from Euclidean-valued functional data to general metric space-valued functional data. The eigenfunctions $\{\phi_j\}_{j=1}^{\infty}$ can be interpreted as principal directions of variation of the functional object process and will be ordered according to the size of the associated eigenvalues. We can view the eigenvalues as representing a metric version of the ‘fraction of variance explained’, which is their common in-
interpretation in the real-valued case. The only requirement for this extension is that the squared metric $d^2$ is of negative type but this is not a severe restriction and in the light of proposition 3 of Sejdinovic et al. (2013) is true for the following examples:

(a) $(\Omega, d)$ where $\Omega$ is the space of univariate probability distributions on a common compact support in $T \subset \mathbb{R}$—choices of $d$ include the popular 2-Wasserstein metric or the $L^2$-metric;
(b) $(\Omega, d)$ where $\Omega$ is the space of correlation matrices of a fixed dimension $r$, where the choice of metrics includes the Frobenius metric, log-Frobenius metric, power Frobenius metric and Procrustes metric (Dryden et al., 2009; Pigoli et al., 2014; Tavakoli et al., 2016);
(c) $(\Omega, d)$ where $\Omega$ is the space of networks with a fixed number, say $r$, of nodes—one can view networks as adjacency matrices or graph Laplacians equipped with the Frobenius metric (Ginestet et al., 2017) or as resistance matrices equipped with the resistance perturbation metric (Monnig and Meyer, 2018).

We conclude that in most cases of interest the autocovariance operator and its eigenfunctions will be well defined.

2.3. Interpretation of metric covariance

When $X$ and $Y$ are real valued, classical Pearson correlation captures the strength and sign of linear (also monotone) associations between $X$ and $Y$. From a geometrical perspective, Pearson correlation can be interpreted as the cosine of the angle between $X$ and $Y$. In $\mathbb{R}^d$, angles between vectors are defined by using inner products, which can also be used for data in Hilbert space to characterize dependence. Specifically, for random functions in the metric space $L^2$ this idea leads to the notion of ‘dynamic correlation’ in FDA (Dubin and Müller, 2005), which was found to be useful for data analysis in genetics (Opgen-Rhein and Strimmer, 2006) and psychology (Liu et al., 2016). Dynamic correlation turns out to be equivalent to metric covariance when the random objects are in the Hilbert space $L^2([0, 1])$, equipped with the usual $L^2$-metric. Metric covariance then provides a generalization beyond Hilbert spaces.

For general metric spaces, under the weak assumption that the squared metric is of negative type, the map $f$ from object to Hilbert space in equation (2) implies that metric covariance can be derived from the inner product in an abstract Hilbert space, whereas metric correlation is obtained by standardizing metric covariance and is thus tied to the notion of an angle in an abstract space. Hence its magnitude can be interpreted as the strength of association between random objects. Although we use the existence of a map $f$ and an associated abstract Hilbert space, we do not require knowledge about $f$. Metric covariance is thus a natural extension of Pearson covariance to general metric spaces.

In recent work (Petersen and Müller, 2019a), Wasserstein covariance for pairs of univariate probability distributions was introduced and was shown to have an appealing interpretation as an expected value of an inner product of optimal transport maps. More specifically, if $f_1$ and $f_2$ are the components of a random bivariate density process and $F_1^{-1}(\cdot)$ and $F_2^{-1}(\cdot)$ the corresponding random quantile functions, the squared Wasserstein distance between $f_1$ and $f_2$ is given by

$$d^2_W(f_1, f_2) = \int_0^1 \{Q_1(t) - Q_2(t)\}^2 dt$$

and the Wasserstein covariance between $f_1$ and $f_2$ was introduced as

$$\text{cov}_W(f_1, f_2) = E\left(\int_0^1 [Q_1(t) - E\{Q_1(t)\}][Q_2(t) - E\{Q_2(t)\}] dt\right).$$
Wasserstein covariance is then easily seen to be a special case of metric covariance when the metric-space-valued random objects are probability distributions and the Wasserstein metric is used. This Wasserstein version of metric covariance was found to quantify the degree of synchronization of the movement of probability mass from the marginal Fréchet means of the probability distributions to the random components of a multivariate density process. In applications to functional magnetic resonance imaging data, this Wasserstein version led to new findings and insights about differences in brain connectivity of normal versus Alzheimer disease patients, which is a topic of special interest in neuroimaging (Petersen and Müller, 2019a). The examples of dynamic correlation for Hilbert-space-valued random variables in FDA and of Wasserstein covariance or Wasserstein correlation demonstrate the utility of metric covariance or metric correlation in non-standard spaces and its interpretability in applications. This provides evidence that metric covariance and metric correlation are indeed useful tools for data analysis in general metric spaces.

A word of caution is in order. Although metric covariance can be universally applied and in the space of distributions with the Wasserstein metric has an interpretation as an inner product of transport maps, such interpretations hinge on the specific metric space in which the random objects are located and may not be available for all spaces. In practice, interpretations for specific scenarios can be important. The choice of the metric also matters and should be considered carefully, as it will affect the interpretation of metric covariance.

Apart from the interpretation of covariance as the expectation of an inner product, the diagonal elements of the metric autocovariance surface reflect a natural notion of variance of metric-space-valued objects, as

$$\text{var}_\Omega = \frac{1}{2} E\{d^2(U, U')\},$$

where $U'$ is an independent copy of $U$. This provides a variation measure that is tied to the average squared distance of objects that are independently sampled from the underlying population, which is a natural and interpretable measure of spread that is well known to coincide with conventional variance in the Euclidean case.

Since it is sensible to define variance for metric-space-valued random objects as

$$\frac{1}{2} E\{d^2(U, U')\} = \frac{1}{2} E\{d^2(U, U') - d^2(U, U)\},$$

it is then natural to extend this to a covariance measure between random objects $(U, V)$ that reflects the difference between squared distances when sampling independently from the marginal distributions of $U$ and $V$ and when sampling from the joint distribution of $(U, V)$. This simple idea provides another avenue to suggest

$$\tilde{\text{cov}}_{\Omega}(U, V) = E\{d^2(U, V') - d^2(U, V)\}.$$

Symmetrizing this expression and adding the factor 0.25 to match the usual definition of covariance in the Euclidean case then leads to formula (4). These arguments also lead to an interpretation of the total variance that corresponds to the trace of the proposed metric covariance operator $C(s, t)$, as an integrated squared distance between the functional random objects $X$ and an independent copy $X'$,

$$\sum_{j=1}^{\infty} \lambda_j = \int_0^1 \text{cov}_\Omega\{X(t), X(t)\} dt = \frac{1}{2} \int_0^1 E[d^2\{X(t), X'(t)\}] dt,$$

We find in our examples and applications that the eigenfunctions that are derived from metric covariance lead to useful and often well interpretable modes of variation of the time-varying
metric random objects in the sense of Jones and Rice (1992), adding to the practical appeal of metric covariance for the analysis of functional random objects.

To conclude this discussion, we note that metric covariance differs substantially from distance correlation (Székely et al., 2007; Lyons, 2013). A distinguishing feature of distance correlation is that it is equivalent to probabilistic independence between the distributions of $U$ and $V$ when it is 0, but we find that it is not suitable as a covariance or correlation measure for random objects in the situations that we study here. Specifically, the autocovariance operator that it generates is not useful for our purposes. For further details on this, see section S2 in the on-line supplement.

3. Functional principal components: generalized Fréchet integrals and Fréchet scores

3.1. Generalized Fréchet integrals and object functional principal components

FPCs in the case of real-valued functional data are projections of the centred process onto the directions of the eigenfunctions and therefore summarize how a function differs from the mean function along orthonormal eigenfunction directions. Formally the FPC of the $i$th process $X_i(t)$ and the $k$th eigenfunction $\phi_k(t)$ is

$$
\xi_{ik} = \int_0^1 \{ X_i(t) - \mu(t) \} \phi_k(t) dt,
$$

where $\mu(\cdot)$ is the mean process. The part of the score contributing to the variability of the functional data is $\int_0^1 X_i(t) \phi_k(t) dt$, which is just a horizontal shift of the actual scores, so centring is not needed when our goal is to decompose the variability of the random processes $X$, which is fortuitous as one cannot ‘centre’ object data to obtain an analogue of $X(t) - \mu(t)$, as algebraic operations such as subtraction are not feasible in metric spaces.

In the Euclidean case for any function $\phi(\cdot)$ on $[0, 1]$, whenever $\int_0^1 \phi(t) dt \neq 0$, we can obtain a scaled version of the integral of $X(t)$ with respect to $\phi(t)$ as follows:

$$
\int_0^1 X(t) \frac{\phi(t)}{\int_0^1 \phi(t) dt} dt = \arg \inf_{\omega \in \mathbb{R}} \int_0^1 d^2_\mathbb{E} \{ \omega, X(t) \} \frac{\phi(t)}{\int_0^1 \phi(t) dt} dt.
$$

This suggests defining an integral of an $\Omega$-valued function $S(\cdot)$ with respect to a real-valued function $\phi(\cdot)$ which integrates to 1. For any real-valued function $\phi(\cdot)$ with $\int_0^1 \phi(t) dt = 1$, we define the generalized Fréchet integral of $S(\cdot)$ with respect to $\phi(\cdot)$ as

$$
\int_\Omega S(t) \phi(t) dt = \arg \inf_{\omega \in \Omega} \int_0^1 d^2 \{ \omega, S(t) \} \phi(t) dt,
$$

provided that the integral $\int_0^1 d^2 \{ \omega, S(t) \} \phi(t) dt$ exists as a limit of Riemann sums for all $\omega \in \Omega$ and the minimizer of the integrals over $\omega \in \Omega$ exists and is unique. A special case of the integral in equation (7) was introduced as the Fréchet integral in Petersen and Müller (2016), where an integral for the space of covariance matrices was constructed for $\phi(t)1$.

The Fréchet integrals that are defined here are far more general. Generalized Fréchet integrals can be interpreted as an extension of weighted Fréchet means (Fréchet, 1948). We omit the additional term ‘generalized’ in what follows and note that Fréchet integrals can be interpreted as projections of functional random objects onto functions $\phi$, by weighing the elements $S(t)$ according to the value of $\phi(t)$, in direct analogy to projections in the linear function space $L^2$.

This feature motivates us to employ Fréchet integrals to obtain object FPCs.

For fixed $\omega \in \Omega$ consider the Fréchet integral function

$$
\int_0^1 X(t) \phi(t) dt = \arg \inf_{\omega \in \mathbb{R}} \int_0^1 d^2_\mathbb{E} \{ \omega, X(t) \} \phi(t) dt.
$$
\[ I(\omega) = \int_0^1 d^2\{\omega, S(t)\} \phi(t) dt, \]

which, if it exists, is the limit of Riemann sums. A sufficient condition for its existence is that \( d^2\{\omega, S(t)\} \phi(t) \) is a continuous function of \( t \in [0, 1] \). If the metric is bounded and \( S(\cdot) \) and \( \phi(\cdot) \) are continuous for \( t \in [0, 1] \), the function \( d^2\{\omega, S(t)\} \phi(t) \) is a continuous function of \( t \in [0, 1] \) and the integral \( I(\omega) \) exists for all \( \omega \). Note that, for any \( \omega \in \Omega \), \( I(\omega) \) is finite by the Cauchy–Schwarz inequality whenever the metric space is bounded and the \( L^2 \)-norm of the function \( \phi(\cdot) \) is finite.

If the integrals \( I(\omega) \) exist as limits of Riemann sums, the question arises under which conditions the minimizers of the Riemann sums converge and whether the limit of the minimizers coincides with the Fréchet integral \( \int_{\Omega} S\phi \). Proposition 2 below addresses this question. Let \( 0 = x_0 < x_1 < x_2 < \ldots < x_k = 1 \) be a partition \( \mathcal{P} \) of \([0, 1]\), where the \([x_j, x_{j+1}]\) are the subintervals of the partition and the length of the \( j \)th subinterval is \( \Delta_j = x_{j+1} - x_j \). The mesh size \( \epsilon_\mathcal{P} \) of the partition is given by \( \epsilon_\mathcal{P} = \max \Delta_j \). We select \( t_0, t_1, \ldots, t_{k-1} \) such that, for each \( j, t_j \in [x_j, x_{j+1}] \). For each \( \omega \in \Omega \), the Riemann sum \( I_\mathcal{P}(\omega) \) corresponding to the partition \( \mathcal{P} \) and \( t_0, t_1, \ldots, t_{k-1} \) is given by

\[ I_\mathcal{P}(\omega) = \sum_{j=0}^{k-1} d^2\{\omega, S(t_j)\} \phi(t_j) \Delta_j \]

and the Riemann integral \( I(\omega) \) is obtained as a limit of Riemann sums as the partition becomes finer. Formally, \( I(\omega) = \lim_{\epsilon_\mathcal{P} \to 0} I_\mathcal{P}(\omega) \).

We shall invoke the following assumptions for the integral function \( I(\omega) \). For ease of notation, we suppress \( t \) in \( \int_{\Omega} S(t) \phi(t) dt \), writing \( \int_{\Omega} S\phi \) in what follows.

**Assumption (i).** The integrand function \( H(\omega, t) = d^2\{\omega, S(t)\} \phi(t) \) is uniformly equicontinuous in \( t \in [0, 1] \) and \( \omega \in \Omega \).

**Assumption (ii).** \( \int_{\Omega} S\phi = \arg \min_{\omega \in \Omega} I(\omega) \) exists and is unique, and \( \inf_{\delta > 0} \int_{d(\omega, \int S\phi) > \delta} I(\omega) > 0 \).

**Assumption (iii).** There are constants \( \beta > 0, \nu > 0 \) and \( C > 0 \) such that

\[ I(\omega) - I \left( \int_{\Omega} S\phi \right) \geq C d^\beta \left( \omega, \int_{\Omega} S\phi \right) \]

whenever \( d(\omega, \int_{\Omega} S\phi) < \nu \).

Define \( \Sigma_{\mathcal{P}, \Omega} S\phi = \arg \min_{\omega \in \Omega} I_\mathcal{P}(\omega) \).

**Proposition 2.**

(a) Under assumption (i), \( I_\mathcal{P}(\omega) \) converges to \( I(\omega) \) uniformly in \( \omega \) as \( \epsilon_\mathcal{P} \to 0 \).

(b) Under assumptions (i) and (ii), \( \lim_{\epsilon_\mathcal{P} \to 0} d(\Sigma_{\mathcal{P}, \Omega} S\phi, \int_{\Omega} S\phi) = 0 \).

(c) If \( \lim_{\epsilon_\mathcal{P} \to 0} h(\epsilon_\mathcal{P}) \sup_{\omega \in \Omega} |I_\mathcal{P}(\omega) - I(\omega)| = 0 \) for a function \( h \) with \( h(\delta) \to \infty \) as \( \delta \to 0 \), then, under assumption (iii), \( \lim_{\epsilon_\mathcal{P} \to 0} h(\epsilon_\mathcal{P}) d(\Sigma_{\mathcal{P}, \Omega} S\phi, \int_{\Omega} S\phi) = 0 \).

As a continuous function on a compact interval is uniformly continuous, whenever \( S(\cdot) \) is continuous and \( \phi(\cdot) \) is bounded and continuous, assumption (i) holds since, for \( D = \text{diam}(\Omega) \),

\[
|H(\omega, t_1) - H(\omega, t_2)| = |d^2\{\omega, S(t_1)\} \phi(t_1) - d^2\{\omega, S(t_2)\} \phi(t_1) + d^2\{\omega, S(t_2)\} \phi(t_1) - d^2\{\omega, S(t_2)\} \phi(t_2)| \leq 2Dd\{S(t_1), S(t_2)\}|\phi(t_1)| + D^2|\phi(t_1) - \phi(t_2)|.
\]
Assumption (i) is sufficient to guarantee that the Fréchet integrals are well defined, whereas assumption (iii) is a restriction on the curvature of the function $I(\omega)$ near its minimizer, implying convergence rates of the approximations of the Fréchet integrals. A few examples of spaces that satisfy assumptions (ii) and (iii) are as follows.

(a) Let $(\Omega, d_\Omega)$ be the space of univariate probability distributions on a common support $T \subset R$. For any $\omega \in \Omega$, denote the corresponding random distribution and quantile functions by $Q(\omega)$. The squared 2-Wasserstein metric between distributions $\omega_1$ and $\omega_2$ is

$$d^2_W(\omega_1, \omega_2) = d^2_L\{Q(\omega_1), Q(\omega_2)\} = \int_0^1 \{Q(\omega_1)(u) - Q(\omega_2)(u)\}^2 du.$$ 

For any $S(t)$ taking values in $\Omega$, where we view $Q\{S(t)\}$ as the quantile function of the distribution at time $t \in [0, 1]$, writing $Q(S(t))(u)$ for the $u$th quantile of the distribution at time $t$, define $Q^*(u) = \int_0^1 Q\{S(t)\}(u) \phi(t) dt$. Since $\int_0^1 \phi(t) dt = 1$, a simple calculation shows that, for any $\omega \in \Omega$,

$$\arg \inf_{\omega \in \Omega} I(\omega) = \arg \inf_{\omega \in \Omega} d^2_L\{Q(\omega), Q^*\};$$

therefore the minimizer exists and is unique by the convexity of the space of univariate quantile functions. By the orthogonal projection theorem the minimizer $\tilde{\omega}$ is uniquely characterized by

$$(Q(\tilde{\omega}) - Q^*, \omega)_{L^2} = 0,$$

for all $\omega \in \Omega$, and therefore it is enough to choose $\nu = C = 1$ and $\beta = 2$ in assumption (iii).

(b) Consider the space of graph Laplacians or graph adjacency matrices of connected, undirected and simple graphs with a fixed number $r$ of nodes $(\Omega, d_F)$, equipped with the Frobenius metric $d_F$. For any $\omega \in \Omega$,

$$d^2_F(\omega_1, \omega_2) = \sum_{j=1}^r \sum_{k=1}^r (\omega_1, jk - \omega_2, jk)^2.$$ 

For any $S(t)$ taking values in $\Omega$, let $S_{jk}(t)$ be the $(j, k)$th entry of the graph Laplacian or the graph adjacency matrix. Define $S^*_{jk} = \int_0^1 S_{jk}(t) \phi(t) dt$. Since $\int_0^1 \phi(t) dt = 1$, it can be easily seen that, for any $\omega \in \Omega$,

$$\arg \inf_{\omega \in \Omega} I(\omega) = \arg \inf_{\omega \in \Omega} d^2_F(\omega, S^*),$$

and so the minimizer exists and is unique by the convexity of the space of graph Laplacians (Ginestet et al., 2017) and the space of graph adjacency matrices. Again, by the orthogonal projection theorem, the minimizer $\tilde{\omega}$ is uniquely characterized by

$$\sum_{j=1}^r \sum_{k=1}^r (\tilde{\omega}, jk - S^*_{jk}) \omega_{jk} = 0,$$

for all $\omega \in \Omega$ and therefore it is enough to choose $\nu = C = 1$ and $\beta = 2$ in assumption (iii).

(c) The same arguments also imply that $(\Omega, d_F)$ satisfies assumptions (ii) and (iii) when $\Omega$ is the space of correlation matrices of a fixed dimension $r$.

As we have seen, for general metric spaces $\Omega$, under mild assumptions on the boundedness of the metric and continuity of the functions $S(\cdot)$ and $\phi(\cdot)$, the Fréchet integral has nice properties if it exists and is unique. Moreover, when $\Omega$ is bounded and $\int_0^1 |\phi(t)| dt < \infty$,
\[ |I(\omega_1) - I(\omega_2)| \leq 2Dd(\omega_1, \omega_2) \int_0^1 |\phi(t)|dt, \]

and therefore \( I(\omega) \) is a continuous function of \( \omega \in \Omega \). This ensures that the Fréchet integral always exists when \( \Omega \) is compact.

We now define the FPCs corresponding to the bounded continuous eigenfunctions \( \phi_k \) of the metric autocovariance operator in the object space by using Fréchet integrals. For this, we assume that all trajectories \( \{X_i(t)\}_{t \in [0,1]} \) have continuous sample paths almost surely and the metric space \( \Omega \) is bounded, and furthermore that the following assumptions hold.

**Assumption 1.** \( \int_0^1 \phi_k(t)dt \neq 0. \)

**Assumption 2.** \( \int_\Delta X_i \phi_k^* \) exists and is unique almost surely for all \( i = 1, \ldots, n \), where \( \phi_k^*(t) = \phi(t)/\int_0^1 \phi(t)dt. \)

Then object FPCs for \( X_i \) and \( \phi_k \) are defined as the Fréchet integrals

\[ \psi_{ik} = \int_\Delta X_i \phi_k^*, \tag{8} \]

which are random objects in \( \Omega \). Similarly to ordinary FPCA we can choose various basis functions aiming to explain a desired percentage of variation in the data utilizing the eigenvalues of the metric autocovariance operator. If \( \Omega = \mathbb{R} \), the object FPCs correspond to a location- and scale-shifted version of the ordinary FPCs.

### 3.2. Fréchet scores

Exploratory data analysis such as checking for clusters or outliers often benefits from plotting the FPCs against each other for the case of real-valued functional data. FPCs defined by using Fréchet integrals live in the object space \( \Omega \) and therefore visualizing them is non-trivial. One approach is to obtain their projections to a lower dimensional real space by using multi-dimensional scaling or its variants (Kruskal, 1964; Belkin and Niyogi, 2002) and then visualizing the projections. Here we propose another approach for obtaining a scalar version of object FPCs. The resulting scalar FPCs are interpretable and can be plotted against each other and are thus useful for exploratory data analysis.

In the real-valued case, one obtains projections of the deviations of the observed random curves from the mean curve onto dominant eigenfunctions. Although the concept of a mean function can be generalized to object functional data by using Fréchet means (Fréchet, 1948), one cannot centre object data and does not have directional information. Nevertheless, it is possible to study how distances of sample curves from the mean curve project onto a few dominant eigenfunctions, in analogy to the real-valued case. Formally, given a random object process \( \{X(t)\}_{t \in [0,1]} \), the population Fréchet mean function is

\[ \mu_\oplus(t) = \arg \min_{\omega \in \Omega} E[d^2\{\omega, X(t)\}], \]

where we assume existence and uniqueness of the minimizer. For real-valued functional data under the Euclidean metric the Fréchet mean function coincides with the usual pointwise mean function. Defining distance functions

\[ D_i(t) = d\{X_i(t), \mu_\oplus(t)\} \]

for sample trajectories \( X_i \), we represent the scalar functions \( D_i \) in the eigenbasis of the metric autocovariance operator, obtaining the coefficients
\[
\beta_{ik} = \int_0^1 D_i(t) \phi_k(t) \, dt = \int_0^1 d\{X_i(t), \mu_{\oplus}(t)\} \phi_k(t) \, dt. \tag{9}
\]

We refer to the scalars \(\beta_{ik}\) as the Fréchet scores.

The Fréchet scores can be interpreted as decomposition of the departures of the sample elements from the ‘central’ Fréchet mean curve in predominant directions of variation. They can be plotted against each other and have the potential to provide interesting insights, as we shall illustrate in the data applications. Considering the existence of the Fréchet scores, with \(D\) denoting as before the diameter of the totally bounded metric space \(\Omega\), continuity of the Fréchet mean function implies that, for any \(t_1, t_2 \in [0, 1]\),

\[
\| d^2\{X_i(t_1), \mu_{\oplus}(t_1)\} - d^2\{X_i(t_2), \mu_{\oplus}(t_2)\} \| = |d^2\{X_i(t_1), \mu_{\oplus}(t_1)\} - d^2\{X_i(t_1), \mu_{\oplus}(t_2)\}| \\
+ d^2\{X_i(t_1), \mu_{\oplus}(t_2)\} - d^2\{X_i(t_2), \mu_{\oplus}(t_2)\}| \\
\leq 2D \| d\{\mu_{\oplus}(t_1), \mu_{\oplus}(t_2)\} + d^2\{X_i(t_1), X_i(t_2)\} \|.
\]

Thus, for continuous eigenfunction \(\phi_k(\cdot)\), the function \(d^2\{X_i(t), \mu_{\oplus}(t)\} \phi_k(t)\) is a continuous function of \(t \in [0, 1]\) almost surely and therefore the Fréchet scores will exist. Proposition 3 shows that under the following assumption 3 the Fréchet mean function is indeed continuous.

**Assumption 3.** For each \(t \in [0, 1]\), the pointwise Fréchet mean \(\mu_{\oplus}(t)\) exists and is unique, and

\[
\inf_{d\{\omega, \mu_{\oplus}(t)\} > \gamma} E[d^2\{\omega, X(t)\}] > E[d^2\{\mu_{\oplus}(t), X(t)\}]
\]

for any \(\gamma > 0\).

**Proposition 3.** If the random object process \(\{X(t)\}_{t \in [0, 1]}\) has almost surely continuous paths, then \(\mu_{\oplus}(\cdot)\) is continuous under assumption 3.

Assumption 3 is satisfied for the space \((\Omega, d_W)\) of univariate probability distributions with the 2-Wasserstein metric and also for the space \((\Omega, d_F)\), where \(\Omega\) is the space of covariance matrices or alternatively graph Laplacians of fixed dimension with the Frobenius metric \(d_F\) (Dubey and Müller, 2017; Petersen and Müller, 2019b).

4. Estimation and theory

Having defined suitable population targets, our goal now is to construct appropriate estimators, starting with a sample of functional random objects. An empirical estimator of the metric autocovariance operator \(C(s, t)\) as defined in Section 2 is given by

\[
\hat{C}(s, t) = \frac{1}{4n(n-1)} \sum_{i \neq j} f_{s,t}(X_i, X_j), \tag{10}
\]

where

\[
f_{s,t}(X_i, X_j) = d^2\{X_i(s), X_j(t)\} + d^2\{X_j(s), X_i(t)\} - d^2\{X_i(s), X_i(t)\} - d^2\{X_j(s), X_j(t)\}.
\]

Observe that, for each \(s, t \in [0, 1]\), \(\hat{C}(s, t)\) is a \(U\)-statistic and the class \(\{\hat{C}(s, t) : s, t \in [0, 1]\}\) is a family of \(U\)-statistics.

Noting that \(\hat{C}(s, t)\) can be viewed as a stochastic process indexed by the function class \(\mathcal{F} = \{f_{s,t}(\cdot, \cdot) : s, t \in [0, 1]\}\), where \(f_{s,t}(x, y) = d^2\{x(s), y(t)\} + d^2\{y(s), x(t)\} - d^2\{x(s), x(t)\} - d^2\{y(s), y(t)\}\) enables us to apply the theory of \(U\)-processes (Nolan and Pollard, 1987, 1988; Arcones and Giné, 1993) for weak convergence (Billingsley, 1968; van der Vaart and Wellner, 1996).
For the uniform convergence of \( \{ \hat{C}(s, t) : s, t \in [0, 1] \} \), we need an assumption on the rate of continuity of the functional random objects.

**Assumption 4.** The process \( X(\cdot) \) is almost surely \( \alpha \)-Hölder continuous for some \( 0 < \alpha \leq 1 \), where the Hölder constant has a finite second moment, i.e. for some non-negative function \( G(X) \) we have

\[
d\{X(s), X(t)\} \leq G(X)|s - t|^\alpha,
\]

where \( E\{G(X)\}^2 < \infty \).

**Theorem 1.** Under assumption 4, the sequence of stochastic processes

\[
U_n(s, t) = \sqrt{n}\{\hat{C}(s, t) - C(s, t)\}
\]

converges weakly to a Gaussian process with mean 0 and covariance function

\[
R(s, t, u, v) = \text{cov}\{f_{s, t}(X, Y), f_{u, v}(X, Y)\}.
\]

Writing \( \hat{\lambda}_j \) and \( \hat{\phi}_j \) for the eigenvalues and eigenfunctions of \( \hat{C}(s, t) \), uniform convergence and rates of convergence of these estimates of the eigenvalues and eigenfunctions of the metric autocovariance operator to their targets are obtained as a direct consequence of proposition 1 under the following assumption on the spacings of the eigenvalues.

**Assumption 5.** For each \( j \geq 1 \), the eigenvalue \( \lambda_j \) has multiplicity 1, i.e. it holds that \( \delta_j > 0 \), where

\[
\delta_j = \min_{1 \leq l \leq j} (\lambda_l - \lambda_{l+1})^{-1}.
\]

**Corollary 1 (Bosq, 2000).** Under assumptions 4 and 5,

\[
|\hat{\lambda}_j - \lambda_j| = O_P(1/\sqrt{n}),
\]

\[
\sup_{s \in [0, 1]} |\hat{\phi}_j(s) - \phi_j(s)| = O_P(1/\delta_j\sqrt{n}).
\]

As in classical FDA, the eigenfunctions \( \phi_j \) are uniquely identifiable only up to a sign change. For theoretical considerations such as the convergence in corollary 1, we may always assume that true and estimated eigenfunctions are aligned in the sense that \( \langle \hat{\phi}_j, \phi_j \rangle \geq 0 \). Our next objective is to obtain sample estimators for the object FPCs (8) that were defined in Section 3.1. For each \( j \), consider the following estimators of \( \phi_j^*(t) \):

\[
\hat{\phi}_j^*(t) = \frac{\hat{\phi}_j(t)}{\int_0^1 \hat{\phi}_j(t)dt}.
\]

A natural estimator for the Fréchet integral \( \psi_{ik}^* \) is then

\[
\psi_{ik}^* = \int_X \hat{\phi}_j^* = \arg \min_{\omega \in \Omega} \int_0^1 d^2\{\omega, X_i(t)\}\hat{\phi}_j^*(t)dt.
\]

To obtain convergence of \( \psi_{ik}^* \) to its population target, we make the following assumptions.

**Assumption 6.** For every \( i \) and \( k \), \( \psi_{ik}^* \) and \( \psi_{ik} \) exist and are unique almost surely. Moreover, for any \( \epsilon > 0 \), \( c_\epsilon = \inf_{d(\omega, \psi_{ik}) > \epsilon} \int_0^1 d^2\{\omega, X_i(t)\}\phi_j^*(t)dt - \int_0^1 d^2\{\psi_{ik}, X_i(t)\}\phi_j^*(t)dt \geq 0 \) almost surely.
**Assumption 7.** There are constants $\beta_1 > 1$, $\nu' > 0$ and $C' > 0$ such that, almost surely,
\[
\left[ \int_0^1 d^2 \{ \omega, X_i(t) \} \phi^*(t) dt - \int_0^1 d^2 \{ \psi_{ik}', X_i(t) \} \phi^*(t) dt \right] \geq C'd^{\beta_1}(\omega, \psi_{ik}'),
\]
whenever $d(\omega, \psi_{ik}') < \nu'$.

Assumption 6 on the existence and uniqueness of the Fréchet integrals is used to establish consistency. Assumption 7 is a restriction on the local behaviour of the integrals around the minimizer and determines the rate of convergence.

**Theorem 2.** Under assumptions 1, 2, 4 and 6,
\[ d(\hat{\psi}_{ik}, \psi_{ik}') = O_P(1). \]

If additionally assumption 7 holds, then
\[ d(\hat{\psi}_{ik}, \psi_{ik}') = O_P(n^{-1/(2\beta_1)}). \]

Here we choose $\hat{\phi}_j$ to be such that $\langle \hat{\phi}_j, \phi_j \rangle \geq 0$ which ensures matching signs for the true and estimated eigenfunctions in the computation of $\hat{\psi}_{ik}$ and $\psi_{ik}'$.

Next we provide estimates of the Fréchet scores and study their asymptotics. The starting point is the following estimator of the population Fréchet mean function:
\[ \hat{\mu}(t) = \arg \min_{\omega \in \Omega} \frac{1}{n} \sum_{i=1}^n d^2 \{ X_i(t), \omega \}. \]  
(12)

We need the following assumptions.

**Assumption 8.** The Fréchet mean function estimate $\hat{\mu}(t)$ exists and is unique almost surely for all $t \in [0, 1]$. Additionally, for every $\varepsilon > 0$, there exists $\tau(\varepsilon) > 0$ such that
\[ \lim_{n \to \infty} P \left( \inf_{\hat{s} \in [0,1]} \inf_{d(\omega, \mu_{\hat{s}}(s)) > \varepsilon} \frac{1}{n} \sum_{i=1}^n \left[ d^2 \{ X_i(s), \omega \} - d^2 \{ X_i(s), \hat{\mu}(s) \} \right] \geq \tau(\varepsilon) \right) = 1. \]

**Assumption 9.** There are a sufficiently small $\delta > 0$ and constants $0 < \nu_0 \leq 1$ and $H_\delta > 0$, such that for all $\Omega$-valued functions $\omega(\cdot)$ with $d_{\infty}(\omega, \mu_{\hat{s}}) < \delta$, where $d_{\infty}(\omega, \mu_{\hat{s}}) = \sup_{s \in [0,1]} d_{\infty}(\omega(s), \mu_{\hat{s}}(s))$, the functions $\omega(\cdot)$ are $\nu_0$ Hölder continuous with Hölder constant bounded above by $H_\delta$, i.e.
\[ d(\omega(s), \omega(t)) \leq H_\delta |s-t|^{\nu_0}. \]

**Assumption 10.** For $I(\delta) = \int_0^1 \sup_{s \in [0,1]} \sqrt{\log(N[\alpha, \delta, B_\delta \{ \mu_{\hat{s}}(s) \}, d])} d\varepsilon$, it holds that $I(\delta) = O(1)$ as $\delta \to 0$ for all sufficiently small $\delta > 0$ and for any constant $A > 0$. Here $B_\delta \{ \mu_{\hat{s}}(s) \} = \{ \omega \in \Omega : d(\omega, \mu_{\hat{s}}(s)) < \delta \}$ is the $\delta$-ball around $\mu_{\hat{s}}(s)$ and $N[\varepsilon, B_\delta \{ \mu_{\hat{s}}(s) \}, d]$ is the covering number, i.e. the minimum number of balls of radius $\varepsilon$ required to cover $B_\delta \{ \mu_{\hat{s}}(s) \}$ (van der Vaart and Wellner, 1996).

**Assumption 11.** There are $\alpha > 0$, $D > 0$ and $\beta_2 > 1$ such that
\[ \inf_{s \in [0,1]} \inf_{d(\omega, \mu_{\hat{s}}(s)) < \alpha} \left( E[d^2 \{ X(s), \omega \}] - E[d^2 \{ X(s), \mu_{\hat{s}}(s) \}] - D d^2 \{ \omega, \mu_{\hat{s}}(s) \} \right) \geq 0. \]

**Proposition 4.** Under assumptions 3 and 8,
\[
\sup_{t \in [0,1]} d\{\hat{\mu}_\oplus(t), \mu_\oplus(t)\} = o_P(1).
\]

Assumptions 4 and 9–11 are required to obtain an entropy condition for the space of functional random objects (lemma 1 below), which is used to establish the rate of convergence of the sample Fréchet mean function. We note that assumption 9, where we assume that in a sufficiently close neighbourhood of the true Fréchet mean function \(\mu_\oplus(t)\) all object functions have a common rate of Hölder continuity and a common Hölder constant, is weaker than assumptions that have been required in classical FDA (see for example Müller et al. (2006)), where one deals with real-valued random functions. Assumption 10 is a bound on the covering number of the object metric space and is satisfied by common instances for random objects that include the examples that were discussed at the end of Section 3.2.

We write \(\omega(\cdot)\) for \(\Omega\)-valued functions \([0, 1] \to \Omega\) and define
\[
V_n(\omega, s) = \frac{1}{n} \sum_{i=1}^{n} [d^2\{X_i(s), \omega(s)\} - d^2\{X_i(s), \mu_\oplus(s)\}],
\]
\[
V(\omega, s) = E[d^2\{X(s), \omega(s)\} - d^2\{X(s), \mu_\oplus(s)\}].
\]
Here \(\hat{\mu}_\oplus(\cdot)\) is the minimizer of \(V_n(\omega, s)\) and \(\mu_\oplus(\cdot)\) is the minimizer of \(V(\omega, s)\). We refer to \(\hat{\mu}_\oplus(\cdot), \mu_\oplus(\cdot)\) and \(\omega(\cdot)\) as \(\hat{\mu}_\oplus, \mu_\oplus\) and \(\omega\) in what follows. To derive the rate of convergence of \(\hat{\mu}_\oplus\), we first obtain a bound for \(E\{\sup_{s \in [0,1]} d(\mu_\oplus(s), \hat{\mu}_\oplus(s))\}\) for small \(\delta \geq 0\), where \(d(\omega, \mu_\oplus) = \sup_{s \in [0,1]} d\{\omega(s), \mu_\oplus(s)\}\). For this, we define function classes
\[
\mathcal{F}_\delta = \{f_{\omega, \delta}(s) = d^2\{x(s), \omega(s)\} - d^2\{x(s), \mu_\oplus(s)\} : s \in [0,1], d(\omega, \mu_\oplus) < \delta\}. \tag{13}
\]
It is easy to see that an envelope function for this class is the constant function \(F(\chi) = 2M\delta\), where \(M\) is the diameter of \(\Omega\). The \(L^2\)-norm of this envelope function is \(\|F\|_2 = 2M\delta\). By theorem 2.14.2 of van der Vaart and Wellner (1996) we have
\[
E\left\{\sup_{s \in [0,1]} \sup_{d(\omega, \mu_\oplus) < \delta} |V_n(\omega, s) - V(\omega, s)|\right\} \leq \frac{2M\delta}{\sqrt{n}} \cdot J[1, \mathcal{F}_\delta, L^2(P)], \tag{14}
\]
where \(J[1, \mathcal{F}_\delta, L^2(P)]\) is the bracketing integral of the function class \(\mathcal{F}_\delta\):
\[
J[1, \mathcal{F}_\delta, L^2(P)] = \int_{0}^{1} \sqrt{1 + \log[N\{\|F\|_2, \mathcal{F}_\delta, L^2(P)\}]} \, d\varepsilon.
\]
Here \(N\{\|F\|_2, \mathcal{F}_\delta, L^2(P)\}\) is the minimum number of balls of radius \(\varepsilon\|F\|_2\) required to cover the function class \(\mathcal{F}_\delta\) under the \(L^2(P)\) norm. Lemma 1 provides the behaviour of the bracketing integral of the function class \(\mathcal{F}_\delta\): a key step for the proof of theorem 3.

**Lemma 1.** Under assumptions 4, 9 and 10, it holds for the function class \(\mathcal{F}_\delta\) as defined in expression (13) that \(J[1, \mathcal{F}_\delta, L^2(P)] = O\{\sqrt{\log(1/\delta)}\}\) as \(\delta \to 0\).

**Theorem 3.** Under assumptions 3, 4 and 8–11,
\[
\sup_{s \in [0,1]} d\{\hat{\mu}_\oplus(s), \mu_\oplus(s)\} = O_P\left[\left\{\frac{\sqrt{\log(n)}}{n}\right\}^{1/\beta_2}\right].
\]
Setting \(\hat{D}_i(t) = d\{X_i(t), \hat{\mu}_\oplus(t)\}\), an application is the convergence of the estimated Fréchet scores.
\[ \hat{\beta}_{ik} = \int_0^1 \hat{D}_i(t)\hat{\phi}_k(t) \, dt. \]

**Corollary 2.** Under assumptions 3–5 and 8–11,
\[
|\hat{\beta}_{ik} - \beta_{ik}| = O_P\left[ n^{-1/2} + \left\{ \frac{\sqrt{\log(n)}}{n} \right\}^{1/\beta_2} \right].
\]

Following the widely adopted convention, we assume throughout that true and estimated eigenfunctions are aligned in the sense that \( \langle \hat{\phi}_j, \phi_j \rangle \geq 0 \), as the scores are identifiable only up to a change in sign.

## 5. Simulations

We illustrate the utility of the proposed methods through simulations for two settings. In the first setting, the space \( \Omega \) consists of univariate probability distributions equipped with the 2-Wasserstein metric and, in the second setting, \( \Omega \) consists of networks with fixed number of nodes, represented as graph adjacency matrices and equipped with the Frobenius metric.

### 5.1. Time-varying probability distributions

We generated random samples of sizes \( n = 25, 50, 100 \) of ‘distribution’-valued curves on the domain \([0, 1]\), where, for each \( t \in [0, 1] \), \( X_i(t) \) is a normal distribution with mean \( \mu_i(t) \) and variance \( \sigma_i^2(t) \) with
\[ 
\mu_i(t) = 1 + U_i \phi_1(t) + V_i \phi_3(t), \quad U_i \sim N(0, 12), \quad V_i \sim N(0, 1),
\]
\[ 
\sigma_i^2(t) = 3 + W_i \phi_2(t) + Z_i \phi_3(t), \quad W_i \sim \sqrt{72}U(0, 1), \quad Z_i \sim \sqrt{9}U(0, 1),
\]
with \( \phi_1(t) = (t^2 - 0.5)/0.3416, \phi_2(t) = \sqrt{3}t, \phi_3(t) = (t^3 - 0.3571t^2 - 0.6t + 0.1786)/0.0895 \), where \( \phi_1, \phi_2 \) and \( \phi_3 \) are orthonormal on \([0, 1]\). We use the 2-Wasserstein metric for the distribution space \( \Omega \). For these specifications, the metric autocovariance function is
\[ 
C(s, t) = 12 \phi_1(s) \phi_1(t) + 6 \phi_2(s) \phi_2(t) + 1.75 \phi_3(s) \phi_3(t),
\]
and \( \phi_1(\cdot), \phi_2(\cdot) \) and \( \phi_3(\cdot) \) are the first three eigenfunctions.

We applied the proposed method to estimate the metric autocovariance operator to the simulated data and obtained its eigenvalues and eigenfunctions. Denoting the estimated metric autocovariance surface and the estimated \( j \)th eigenvalue and eigenfunction obtained at the \( k \)th simulation run by \( \hat{C}_k(s, t) \) respectively \( \hat{\lambda}_{j,k} \) and \( \hat{\phi}_{j,k} \), we computed mean integrated squared errors (MISE)
\[
\text{MISE}(C) = \frac{1}{100} \sum_{k=1}^{100} \int_0^1 \int_0^1 \{ \hat{C}_k(s, t) - C(s, t) \}^2 \, ds \, dt,
\]
\[
\text{MISE}(\phi_j) = \frac{1}{100} \sum_{k=1}^{100} \int_0^1 \{ \hat{\phi}_{j,k}(s) - \phi_j(s) \}^2 \, ds,
\]
\[
\text{MISE}(\lambda_j) = \frac{1}{100} \sum_{k=1}^{100} (\hat{\lambda}_{j,k} - \lambda_j)^2.
\]
Fig. 1 shows the true and estimated metric autocovariance surfaces and their eigenfunctions for one randomly chosen simulation run for \( n = 25 \) and \( n = 100 \). We find that the method proposed has negligible bias as the sample size increases. The MISEs are reported in Table 1 and are seen to decrease with increasing sample sizes.

To illustrate the nature of the simulated random density trajectories, four density-valued random functions that are part of a sample of density-valued random functions as generated in one Monte Carlo run are displayed in Fig. 2, reflecting variation in means and variances of the Gaussian distributions as a function of time for the four selected subjects. The estimated object FPCs, i.e. the Fréchet integrals of the object curves along the first two eigenfunctions, from one Monte Carlo run are in Fig. 3 for sample size 50. Here the first object FPCs reflect variation in location of the distributions and the second object FPCs variation in the variance of the distributions, which is what we expect in view of how these data were generated. The object FPCs are found to be useful for discovering the underlying modes of variation for distributions as functional random objects.

### 5.2. Time-varying networks

In each iteration, we generated random samples of sizes \( n = 25, 50, 100 \) of time-varying random networks with 10 nodes each in the time interval \([0, 1]\). For generating the edge weights, we...
Table 1. MISEs for the estimators of the metric autocovariance kernel $C$ and the eigenfunctions $\phi_1$ and $\phi_2$ in dependence on sample size when the functional random objects are distributions

| $n$ | $C$   | $\phi_1$ | $\phi_2$ | $\phi_3$ | $\lambda_1$ | $\lambda_2$ | $\lambda_3$ |
|-----|-------|----------|----------|----------|-------------|-------------|-------------|
| 25  | 12.2709 | 1.2841  | 1.5351  | 1.5202  | 10.8634     | 7.8471     | 3.5798     |
| 50  | 8.6598  | 0.0504  | 0.0201  | 0.0030  | 0.9748      | 0.6482     | 0.3680     |
| 100 | 4.0697  | 0.0158  | 0.0084  | 0.0047  | 0.1239      | 0.0607     | 0.0314     |

Fig. 2. Four randomly chosen observations of density-valued trajectories, selected from the sample of distributions generated by one of the Monte Carlo runs: the densities are plotted as a function of time.

followed the model that is described below. We assumed that the network has two communities: the first five nodes belonging to one community and the second five nodes to the other. For each fixed time $t$, the edge weights within each community and also those between the communities are the same, where the latter are smaller than the within-community edge weights. Formally, if $p_{1,i}(t)$, $p_{2,i}(t)$ and $p_{12,i}(t)$ denote the edge weight at time $t \in [0, 1]$ for the first community, the second community and between communities, for the $i$th network-valued curve we generated

$$
\begin{align*}
    p_{1,i}(t) &= 0.5 + U_i \phi_1(t) + V_i \phi_3(t), \\
    p_{2,i}(t) &= 0.5 + W_i \phi_2(t) + Z_i \phi_3(t), \\
    p_{12,i}(t) &= 0.1.
\end{align*}
$$

(17)
Here the $U_i$, $V_i$, $W_i$ and $Z_i$ were generated from the uniform distributions $U(0, 0.4)$, $U(0, 0.1)$, $U(0, 0.3)$ and $U(0, 0.1)$ respectively. The functions $\phi_1(t)$, $\phi_2(t)$ and $\phi_3(t)$ are orthonormal polynomials derived from Jacobi polynomials $P_n^{(\alpha, \beta)}(x)$ (Totik, 2005), which are classical orthogonal polynomials for $\alpha, \beta > 1$. They are orthogonal with respect to the basis $(1 + x)^{\beta}(1 - x)^{\alpha}$ on $[-1, 1]$. With a suitable change of basis, one can obtain a version of the Jacobi polynomials on $[0, 1]$ which are orthonormal with respect to the weight function $x^\beta(1 - x)^\alpha$ on $[0, 1]$. We selected $\phi_1(t)$, $\phi_2(t)$ and $\phi_3(t)$ as

$$\phi_j(t) = \frac{P_j^{(3, 3)}(2t - 1)t^{1.5}(1 - t)^2}{\left[\int_0^1 P_j^{(3, 3)}(2t - 1) t^{1.5}(1 - t)^2 dt\right]^{1/2}} \quad \text{for } j = 1, 2, 3.$$ 

The weighted networks are represented as graph adjacency matrices with the Frobenius metric. Here the true metric autocovariance function is

$$C(s, t) = 0.266 \phi_1(s)\phi_1(t) + 0.15 \phi_2(s)\phi_2(t) + 0.0417 \phi_3(s)\phi_3(t),$$

and $\phi_1(\cdot)$, $\phi_2(\cdot)$ and $\phi_3(\cdot)$ are the first three eigenfunctions.

We estimated the metric autocovariance operator from the simulated data and obtained its eigenfunctions for different sample sizes. Fig. 4 displays the true and estimated metric autocovariance surfaces and corresponding eigenfunctions for one randomly chosen simulation run for $n = 25$ and $n = 100$. The MISEs were computed as described for the previous simulation setting and are reported in Table 2. They decrease with increasing sample sizes. The method proposed is seen to work very well.

The object FPCs were obtained by using Fréchet integrals (11). For visualization they are presented as ‘networks.mov’ in the on-line supplementary materials. In the movie the leftmost plot corresponds to Fréchet integrals for the first eigenfunction which, as expected because of the true model, shows variation only in the edge weights of the first community. The middle plot...
corresponds to Fréchet integrals for the second eigenfunction and indicates variation only in the edge weights of the second community. The rightmost plot corresponds to Fréchet integrals for the third eigenfunction and where variation in both the first and the second community edge weights can be discerned.

6. Data applications

6.1. Mortality data

The human mortality database provides life table data differentiated by gender and is available from www.mortality.org. Currently the mortality database contains life table data

---

**Table 2.** MISEs for the estimators of the metric autocovariance kernel \( C \) and eigenfunctions–eigenvalues \( \phi_j, \lambda_j, j = 1, 2, 3 \), in dependence on sample size for samples of functional random objects that correspond to time-varying networks

| \( n \) | \( C \) | \( \phi_1 \) | \( \phi_2 \) | \( \phi_3 \) | \( \lambda_1 \) | \( \lambda_2 \) | \( \lambda_3 \) |
|------|-------|-------|-------|-------|-------|-------|-------|
| 25   | 0.0039 | 0.0039 | 0.0017 | 0.0007 | 0.0025 | 0.0010 | 0.0007 |
| 50   | 0.0017 | 0.0093 | 0.0046 | 0.0021 | 0.0007 | 0.0003 | 0.0001 |
| 100  | 0.0010 | 0.0130 | 0.0063 | 0.0028 | 0.0001 | 0.0001 | 0.0001 |
for 37 countries spanning over five decades. One can obtain histograms from life tables and smooth these with local least squares to obtain estimated probability density functions for age at death. We carried this out for the age interval [0, 80] years. The mortality data can then be viewed as samples of time-varying univariate probability distributions, for a sample of 32 countries, where the time axis corresponds to calendar years between 1960 and 2009 and the observation that is made at each calendar year for each country corresponds to the age-at-death distribution for that year. We included the 32 countries which had complete records over the entire calendar period. For each country and year, we used the Hades package that is available from https://stat.ucdavis.edu/hades/ for smoothing the histograms and used bandwidth = 2 to obtain the age-at-death densities. For illustration, the time-varying age-at-death distributions represented as density functions for the age interval [0, 80] years and indexed by calendar year are displayed for four selected countries, the USA, Ukraine, Russia and Portugal, for males in Fig. 5 and for females in Fig. 6.

Choosing the 2-Wasserstein metric for the probability distributions space, the estimated metric autocovariance surfaces for males and females can be inspected in Fig. 7 and the eigenfunctions of the corresponding autocovariance operators in Fig. 8. The autocovariance functions and eigenfunctions indicate that there are systematic differences between males and females.

The resulting object FPCs, i.e. the Fréchet integrals, are illustrated in Fig. 9 for the first two eigenfunctions. The object FPCs are distributions that are represented as densities for

![Fig. 5](image-url) Time-varying age-at-death density functions for the age interval [0, 80] years for males in (a) the USA, (b) Ukraine, (c) Russia and (d) Portugal
Fig. 6. Time-varying age-at-death density functions for the age interval [0, 80] years for females in (a) the USA, (b) Ukraine, (c) Russia and (d) Portugal.

males and females. The eastern European countries that are included in the database, namely
Belarus, Bulgaria, the Czech Republic, Hungary, Latvia, Lithuania, Poland, Slovakia, Ukraine,
Russia and Estonia, underwent major political upheaval due to the end of Communist rule in
these regions during the period between the late 1980s and early 1990. This is reflected in clear
distinctions between the eastern European countries (red) and the rest (blue) in the Fréchet
integrals for the males but much less so for the females, which indicates that particularly male
mortality was affected by the political upheavals.

The sample Fréchet mean function at a particular calendar year corresponds to the sample
average of the quantile functions of the various countries at that calendar year and is illustrated in
the movies ‘mean_males.mov’ and ‘mean_females.mov’ in the on-line supplementary materials.
Fig. 10 illustrates the scalar FPCs, i.e. the Fréchet scores for the second eigenfunction plotted
against the Fréchet scores for the first eigenfunction for males and females. Russia is an outlier
for the first eigenfunction for males and Portugal is an outlier for the second eigenfunction, even
though it does not belong to the above list of eastern European countries. One could speculate
that this might be related to the fact that Portugal in 1974 moved to a democratic government
after four decades of authoritarian dictatorship. Figs 5 and 6 suggest higher infant mortality
for both males and females in Portugal during the earlier era. Another interesting observation
is that the order of outliers is reversed for females, as Russia turns out to be an outlier for
females for the second eigenfunction and Portugal for the first. The plots of the Fréchet scores
against each indicate that there are clear distinctions between the two groups of countries and
Portugal.
6.2. Time-varying networks for New York taxi data

The New York City Taxi and Limousine Commission provides records on pick-up and drop-off dates and times, pick-up and drop-off locations, trip distances, itemized fares, rate types, payment types and driver-reported passenger counts for yellow and green taxis which are available from http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml. The time resolution of these data is of the order of seconds. Of interest are networks which represent
how many people travelled between places of interest and the evolution of these networks during a typical day. To study this, we constructed samples of time-varying networks where the sample elements are the recordings for each day in the year 2016. Three days (January 23rd and 24th and March 13th) were excluded from the study because of incomplete records.

We focus on the Manhattan area, which has the highest traffic and split the area according to
Fig. 9. Fréchet integrals (11) for the first eigenfunction for (a) males and (b) females and for the second eigenfunction for (c) males and (d) females for the mortality data of eastern European countries, namely Belarus, Bulgaria, the Czech Republic, Hungary, Latvia, Lithuania, Poland, Slovakia, Ukraine, Russia and Estonia; the other countries.
the provided location shape files into 10 zones, which form the regions of interest. Details about the zones are in section S6.1 of the on-line supplement. Yellow taxis provide the predominant taxi service in Manhattan. We divided each day into 20-min intervals, and for each interval constructed a network with nodes corresponding to the 10 selected zones and edge weights representing the number of people who travelled between the zones connecting the edges within the 20-min interval. The edge weights were normalized by the maximum edge weight for each day so that they lie in $[0, 1]$. We thus have a time-varying network for each of the 363 days in 2016 for which complete records are available, where the time points where the network-valued functions are evaluated correspond to the 20-min intervals of a 24-h day. The observations at each time point correspond to a 10-dimensional graph adjacency matrix which characterizes the network between the 10 zones of Manhattan for the particular 20-min interval.

We choose the Frobenius metric as metric between the graph adjacency matrices. The sample Fréchet mean function at a particular time point therefore corresponds to the sample average of the graph adjacency matrices of 363 networks corresponding to different days for that time point. It is illustrated in the movie ‘mean_NY.mov’ in the on-line supplementary materials. Fig. 11 illustrates the estimated autocovariance function and associated eigenfunctions. The plots of the Fréchet scores for the second, third and fourth eigenfunction against the scores for the first eigenfunction can be found in Fig. 12, where the blue dots correspond to Mondays–Thursdays, the green dots to Fridays and the red dots to Saturdays and Sundays. Several interesting patterns emerge: weekdays and weekends form clearly distinguishable clusters. Special holidays show similar patterns to those of weekends. Several outliers can be identified by using the projection scores for the eigenfunctions, which turn out to be special days: for the first eigenfunction, the outliers correspond to New Year’s day and November 6th, 2016, which is the day when daylight saving ends. March 13th, 2016, is the day that the daylight saving begins but was excluded as it did not have complete records. For the second eigenfunction, an outlying point is Independence Day, July 4th, 2016, and, for the
Fig. 11. (a) Estimated metric autocovariance surface (10) and (b) the corresponding eigenfunctions for the New York taxi data, viewed as time-varying networks: red, 44.7%; blue, 32.72%; green, 10.56%; purple, 8.42%

third eigenfunction, February 14th, 2016, which is Valentine’s day. Another day that stands out is September 18th, 2016. On further investigation it was found that between September 17th and 19th, 2016, three bombs exploded and several unexploded bombs were found in the New York metropolitan area (https://en.wikipedia.org/wiki/2016_New_York_and_New_Jersey_bombings).

We then repeated the analysis separately for three groups of days, namely the weekdays Monday–Thursday (group 1), Fridays and weekends (group 2) and holidays (group 3). We present the results in Fig. 17 (in the on-line supplement) and in several movies whose descriptions can be found in sections S3 and S5 of the on-line supplement.
Fig. 12. Fréchet scores (15) for (a) the second, (b) third and (c) fourth eigenfunctions in the $y$-axis plotted against Fréchet scores for the first eigenfunction in the $x$-axis, for the New York taxi data: ●, Mondays–Thursdays; ○, Fridays; ●, Saturdays and Sundays.

6.3. World trade data
The Center for International Data at the University of California, Davis (http://cid.econ.ucdavis.edu/nberus.html), provides detailed documentation of United Nations trade data for the years 1962–2000. The data set, which is publicly available from www.nber.org, contains bilateral trade data during this time period for several commodities and countries. We studied the time period 1970–1999 for 46 actively trading countries and the 26 most common types of commodity. The list of chosen countries and commodities can be found in section S6.2 of the on-line supplement. For each country, commodity and year, we represent current trade as the ratio of the amount of total trade, i.e. import–export value (in thousands of US dollars),
to the amount of total trade recorded for the same commodity and country in the year 2000, yielding a 26-dimensional vector of trade ratios.

Viewing the countries as sampling units, we obtain for each country and calendar year \( t \) a \((26 \times 26)\)-dimensional raw covariance matrix of commodities trade ratios as \( \tilde{\Sigma}(t) = (Q(t) - \bar{Q}(t))(Q(t) - \bar{Q}(t))^\top \), where \( Q(t) \) is the country-specific 26-dimensional vector of commodities trade ratios for year \( t \) and the mean vector \( \bar{Q} \) is obtained as a cross-sectional average over all 46 countries. These raw time-varying raw covariances were then smoothed by using local Fréchet regression with a Gaussian kernel (Petersen and Müller, 2019b; Petersen et al., 2019) to obtain samples of smooth time-varying 26-dimensional covariance matrices between the components.

Fig. 13. (a) Estimated metric autocovariance surface and (b) corresponding first four eigenfunctions for the trade data: \( \quad \), 47.95%; \( \color{blue}{-} \), 20.67%; \( \color{green}{-} \), 15.64%; \( \color{purple}{-} \), 6.79%
Fig. 14. Fréchet scores for various eigenfunctions plotted against each other for the trade data

of commodities trade for each of the 46 countries over the time period 1970–2000, yielding
time-varying covariance matrices over the time period 1970–2000 as functional random objects.

When adopting the Frobenius metric, the sample Fréchet mean function at calendar year \( t \)
corresponds to the sample average of the smoothed covariance across 46 countries for year \( t \). Fig.
13(a) illustrates the estimated metric autocovariance function and Fig. 13(b) its eigenfunctions.
The metric autocovariance and its eigenfunctions provide insights about world trade patterns
over the time period 1970–1999. The first eigenfunction represents increased variability due
to overall expansion in world trade over the years from 1970 to 1999. The slope of the first
eigenfunction is more gradual before 1985 but increases sharply starting from 1985, stagnates a
little around 1990 and then again picks up. This can be connected to the boom in world trade
against the last decade of the new millennium. The second eigenfunction corresponds to a
contrast before 1990 and after 1990. The peak in the second eigenfunction between 1980 and 1985
could be related to a major economic downturn caused by recession affecting several countries
in the data set during the early 1980s. The recession began in the USA in 1981 and continued till
1982 and affected many of the developed western countries. The third eigenfunction captures
effects of the early 1990s recession, which compared with the 1980s recession was much milder.

In Fig. 14, the Fréchet scores for the first four eigenfunctions are plotted against each other.
Thailand and Egypt have high Fréchet scores for the first eigenfunction and Saudi Arabia ranks
the highest for the second eigenfunction. Chile, Israel, Hong Kong and Bulgaria turn out to
figure prominently in the third eigenfunction. Further visualization can be found in section
S4 of the on-line supplement, including a movie that is described in section S5 of the on-line
supplement that demonstrates the object FPCs.

7. Discussion

We propose an extension of functional data methods to the case of functional random objects.
The basis of our approach is metric covariance: a novel covariance measure for paired metric-
space-valued data. Eigenfunctions of the metric covariance operator for time-varying object
data aid in creating a version of object FPCA, where the object FPCs in the metric space \( \Omega \) are
obtained as Fréchet integrals, which are a general and versatile concept. Alternatively, components of variation can be quantified by Fréchet scores, which are real numbers. For the precursor problem, where we have non-functional time-varying object data, i.e. we have observations for just one random object function over time, methods for metric-space-valued regression have been considered previously (Steinke et al., 2010; Faraway, 2014; Petersen and Müller, 2019b), often under the special assumption that the regression responses are on a Riemannian manifold (Shi et al., 2009; Fletcher, 2013; Hinkle et al., 2012; Su et al., 2012; Yuan et al., 2012; Cornea et al., 2017). However, the more general object function case, which is characterized by samples of random functions that are object valued, is considerably more challenging, as the absence of a linear structure in the object space both globally and locally imposes serious limitations on the methods that can be applied.

The tools that we propose here for functional random objects, namely metric covariance, the metric autocovariance operator and its eigenfunctions, the Fréchet integrals and the Fréchet scores, make it possible to obtain compact summaries, visualizations and interpretations of the observed samples of time-varying object data that in themselves are highly complex and difficult to quantify. These tools can provide insights into the patterns of variability of the object trajectories, as we demonstrated in the simulations and data examples. The quantification of functional random objects can also be used for other tasks. For example the object FPCs that we introduce reside in the object space and can serve as responses for a regression model, where predictors are Euclidean vectors and responses are random object trajectories, which are summarized by these object FPCs. Implementing such a regression approach is analogous to the principal component approach for function-to-function regression (Yao et al., 2005b). Various regression models can then be implemented through Fréchet regression (Petersen and Müller, 2019b). For the case where functional random objects feature as predictors in a regression setting, one can employ the vector of Fréchet scores that summarize each random object trajectory as predictors. The ensuing regression, classification and clustering models will be interesting topics for future research.

A core challenge that one faces when modelling and analysing samples of random object trajectories is that, in contrast with the situation for real-valued processes, we cannot expect to represent object-valued processes in terms of an analogue to the Karhunen–Loève expansion, because of the lack of a linear structure in the object space $\Omega$. In some special cases such expansions are possible, e.g. through a transformation method, whenever random objects can be transformed to a linear space, as exemplified for the case of objects that are probability distributions (Petersen and Müller, 2016) or for the case of Riemannian manifold-valued objects (Dai and Müller, 2018). Apart from such special cases, it is an open problem whether more general useful representations of functional random objects can be found. Another open problem is inference for such data, e.g. comparing two groups or testing for structural features of autocovariance. Here the metric autocovariance operator that we introduce in this paper and also the Fréchet mean function could prove useful for the extension of tests that have been considered for real-valued functional data (for some recent examples, see Aston et al. (2017), Constantinou et al. (2017), Chen and Lynch (2018) and Choi and Reimherr (2018)). These and many other open problems in this area indicate that there is ample potential for future research.

Acknowledgements

We thank the reviewers for their comprehensive and helpful comments that led to numerous improvements and new perspectives. This research was supported by National Science Foundation grant DMS-1712864 and National Institutes of Health grant 5UH3OD023313.
References

Anirudh, R., Turaga, P., Su, J. and Srivastava, A. (2017) Elastic functional coding of Riemannian trajectories. 
IEEE Trans. Pattn Anal. Mach. Intell., 39, 922–936.

Arcones, M. A. and Giné, E. (1993) Limit theorems for U-processes. Ann. Probab., 1494–1542.

Aston, J. A., Pigoli, D. and Tavakoli, S. (2017) Tests for separability in nonparametric covariance operators of 
random surfaces. Ann. Statist., 45, 1431–1461.

Belkin, M. and Niyogi, P. (2002) Laplacian eigenmaps and spectral techniques for embedding and clustering. In 
Advances in Neural Information Processing Systems, pp. 585–591.

Berrendero, J., Justel, A. and Svarc, M. (2011) Principal components for multivariate functional data. Computnl 
Statist. Data Anal., 55, 2619–2634.

Bigot, J., Gouet, R., Klein, T. and López, A. (2017) Geodesic PCA in the Wasserstein space by convex PCA. Ann. 
Inst. H. Poincaré B, 53, 1–26.

Billingsley, P. (1968) Convergence of Probability Measures. New York: Wiley.

Bosq, D. (2000) Linear Processes in Function Spaces: Theory and Applications. New York: Springer.

Castro, P. E., Lawton, W. H. and Sylvestre, E. A. (1986) Principal modes of variation for processes with continuous 
sample curves. Technometrics, 28, 329–337.

Chen, K., Delicado, P. and Müller, H.-G. (2017) Modelling function-valued stochastic processes, with applications 
to fertility dynamics. J. R. Statist. Soc. B, 79, 177–196.

Chen, K. and Lynch, B. (2018) A test of weak separability for multi-way functional data, with application to brain 
connectivity studies. Biometrika, 105, 815–831.

Chen, K. and Müller, H.-G. (2012) Modeling repeated functional observations. J. Am. Statist. Ass., 107, 1599–
1609.

Chiou, J.-M., Chen, Y.-T. and Yang, Y.-F. (2014) Multivariate functional principal component analysis: a normalization approach. Statist. Sin., 24, 1571–1596.

Chiou, J.-M. and Li, P.-L. (2007) Functional clustering and identifying substructures of longitudinal data. J. R. 
Statist. Soc. B, 69, 679–699.

Chiou, J.-M., Yang, Y.-F. and Chen, Y.-T. (2016) Multivariate functional linear regression and prediction. J. 
Multiv. Anal., 146, 301–312.

Choi, H. and Reimherr, M. (2012) A geometric approach to confidence regions and bands for functional parame-
ters. J. R. Statist. Soc. B, 80, 239–260.

Claeskens, G., Hubert, M., Slaets, L. and Vakili, K. (2014) Multivariate functional half-space depth. J. Am. Statist. 
Ass., 109, 411–423.

Constantinou, P., Kokoszka, P. and Reimherr, M. (2017) Testing separability of space-time functional processes. 
Biometrika, 104, 425–437.

Cornea, E., Zhu, H., Kim, P. and Ibrahim, J. G. (2017) Regression models on Riemannian symmetric spaces. J. 
R. Statist. Soc. B, 79, 463–482.

Dai, X. and Müller, H.-G. (2018) Principal component analysis for functional data on Riemannian manifolds and 
spheres. Ann. Statist., 46, 3334–3361.

Dai, X., Müller, H.-G. and Yao, F. (2017) Optimal Bayes classifiers for functional data and density ratios. 
Biometrika, 79, 545–560.

Dauxois, J., Pousse, A. and Romain, Y. (1982) Asymptotic theory for the principal component analysis of a vector 
random function: some applications to statistical inference. J. Multiv. Anal., 12, 136–154.

Dong, J. J., Wang, L., Gill, J. and Cao, J. (2018) Functional principal component analysis of glomerular filtration 
rate curves after kidney transplant. Statist. Meth. Med. Res., 27, 3785–3796.

Dryden, I. L., Koloydenko, A. and Zhou, D. (2009) Non-Euclidean statistics for covariance matrices, with appli-
cations to diffusion tensor imaging. Ann. Appl. Statist., 3, 1102–1123.

Dubey, P. and Müller, H.-G. (2017) Fréchet analysis of variance for random objects. Preprint arXiv:1710.02761.

Dubin, J. A. and Müller, H.-G. (2005) Dynamical correlation for multivariate longitudinal data. J. Am. Statist. 
Ass., 100, 872–881.

Faraway, J. J. (2014) Regression for non-Euclidean data using distance matrices. J. Appl. Statist., 41, 2342–2357.

Fletcher, P. T. (2013) Geodesic regression and the theory of least squares on Riemannian manifolds. Int. J. Comput. 
Vis, 105, 171–185.

Fréchet, M. (1948) Les éléments aléatoires de nature quelconque dans un espace distancié. Ann. Inst. H. Poincaré, 
10, 215–310.

Ginestet, C. E., Li, J., Balachandran, P., Rosenberg, S. and Kolaczyk, E. D. (2017) Hypothesis testing for network 
data in functional neuroimaging. Ann. Appl. Statist., 11, 725–750.

Hinkle, J., Muralidharan, P., Fletcher, P. T. and Joshi, S. (2012) Polynomial regression on Riemannian manifolds. 
In Computer Vision, pp. 1–14. New York: Springer.

Horvath, L. and Kokoszka, P. (2012) Inference for Functional Data with Applications. New York: Springer.

Hsing, T. and Eubank, R. (2015) Theoretical Foundations of Functional Data Analysis, with an Introduction to 
Linear Operators. New York: Wiley.

Jacques, J. and Preda, C. (2014) Model-based clustering for multivariate functional data. Computnl Statist. Data 
Anal., 71, 92–106.
Tavakoli, S., Pigoli, D., Aston, J. A. and Coleman, J. (2016) A spatial modeling approach for linguistic object data. *J. Am. Statist. Ass.*, 112, 1261–1273.

Liu, S., Zhou, Y., Palumbo, R. and Wang, J.-L. (2016) Dynamical correlation: a new method for quantifying synchrony with multivariate intensive longitudinal data. *Psychol. Meth.*, 21, 291–308.

Lyons, R. (2013) Distance covariance in metric spaces. *Ann. Probab.*, 41, 3284–3305.

Monnig, N. D. and Meyer, F. G. (2018) The resistance perturbation distance: a metric for the analysis of dynamic networks. *Discr. Appl. Math.*, 236, 347–386.

Müller, H.-G., Stadtmüller, U. and Yao, F. (2006) Functional variance processes. *J. Am. Statist. Ass.*, 101, 1007–1018.

Nolan, D. and Pollard, D. (1987) U-processes: rates of convergence. *Ann. Statist.*, 780–799.

Nolan, D. and Pollard, D. (1988) Functional limit theorems for U-processes. *Ann. Probab.*, 1291–1298.

Opden-Rhein, R. and Strimmer, K. (2006) Inferring gene dependency networks from genomic longitudinal data: a functional data approach. *Revstat*, 4, 53–65.

Park, S. Y. and Staicu, A.-M. (2015) Longitudinal functional data analysis. *Stat.*, 4, 212–226.

Petersen, A., Deoni, S. and Müller, H.-G. (2019) Fréchet estimation of time-varying covariance matrices from sparse data, with application to the regional co-evolution of myelination in the developing brain. *Ann. Appl. Statist.*, to be published.

Petersen, A. and Müller, H.-G. (2016) Fréchet integration and adaptive metric selection for interpretable covariances of multivariate functional data. *Biometrika*, 103, 103–120.

Petersen, A. and Müller, H.-G. (2016) Functional data analysis for density functions by transformation to a Hilbert space. *Ann. Statist.*, 44, 183–218.

Petersen, A. and Müller, H.-G. (2019a) Wasserstein covariance for multiple random densities. *Biometrika*, to be published.

Petersen, A. and Müller, H.-G. (2019b) Fréchet regression for random objects with Euclidean predictors. *Ann. Statist.*, 47, 691–719.

Pigoli, D., Aston, J. A., Dryden, I. L. and Secchi, P. (2014) Distances and inference for covariance operators. *Biometrika*, 101, 409–422.

Ramsay, J. O. and Silverman, B. W. (2005) *Functional Data Analysis*, 2nd edn. New York: Springer.

Robert, P. and Escoufier, Y. (1976) A unifying tool for linear multivariate statistical methods: the RV-coefficient. *Appl. Statist.*, 25, 257–265.

Schoenberg, I. J. (1938) Metric spaces and positive definite functions. *Trans. Am. Math. Soc.*, 44, 522–536.

Seguy, V. and Cuturi, M. (2015) Principal geodesic analysis for probability measures under the optimal transport metric. In *Advances in Neural Information Processing Systems*, pp. 3312–3320.

Sejdinovic, D., Sriperumbudur, B., Gretton, A. and Fukumizu, K. (2013) Equivalence of distance-based and RKHS-based statistics in hypothesis testing. *Ann. Statist.*, 41, 2263–2291.

Shi, X., Styner, M., Lieberman, J., Ibrahim, J. G., Lin, W. and Zhu, H. (2009) Intrinsic regression models for noisy points on nonlinear manifolds. *Medical Image Computing and Computer-assisted Intervention*, pp. 192–199. New York: Springer.

Steinke, F., Hein, M. and Schölkopf, B. (2010) Nonparametric regression between general Riemannian manifolds. *SIAM J. Imaging Sci.*, 3, 527–563.

Su, J., Dryden, I. L., Klassen, E., Le, H. and Srivastava, A. (2012) Fitting smoothing splines to time-indexed, noisy points on nonlinear manifolds. *Im. Visn Comput.*, 30, 428–442.

Suarez, A. J. and Ghosal, S. (2016) Bayesian clustering of functional data using local features. *Baysn Anal.*, 11, 71–98.

Székely, G. J. and Rizzo, M. L. (2017) The energy of data. *A. Rev. Statist. Appl.*, 4, 447–479.

Székely, G. J., Rizzo, M. L. and Bakirov, N. K. (2007) Measuring and testing dependence by correlation of *Weak Convergence and Empirical Processes*. New York: Springer.

Verbeke, G., Fieuws, S., Molenberghs, G. and Davidian, M. (2014) The analysis of multivariate longitudinal data: a review. *Statist. Meth. Med. Res.*, 23, 42–59.

Wang, J.-L., Chiou, J.-M. and Müller, H.-G. (2016) Functional data analysis. *A. Rev. Statist. Appl.*, 3, 257–295.

Yao, F., Müller, H.-G. and Wang, J.-L. (2005a) Functional data analysis for sparse longitudinal data. *J. Am. Statist. Ass.*, 100, 577–590.
Yao, F., Müller, H.-G. and Wang, J.-L. (2005b) Functional linear regression analysis for longitudinal data. *Ann. Statist.*, **33**, 2873–2903.

Yuan, Y., Zhu, H., Lin, W. and Marron, J. S. (2012) Local polynomial regression for symmetric positive definite matrices. *J. R. Statist. Soc. B*, **74**, 697–719.

Zhou, L., Huang, J. and Carroll, R. (2008) Joint modelling of paired sparse functional data using principal components. *Biometrika*, **95**, 601–619.

**Supporting information**

Additional ‘supporting information’ may be found in the on-line version of this article:

‘Online supplement for “Functional models for time-varying random objects”’. 