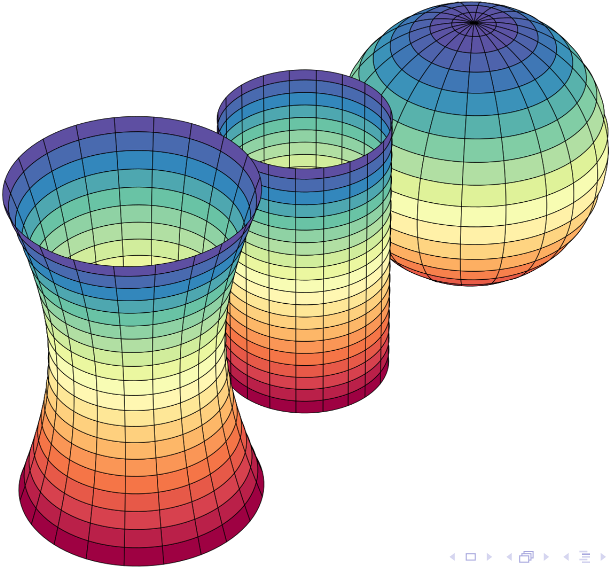
Functional Principal Component Analysis for Manifold Data

February 6, 2019

Types of manifolds or manifolds with additional structures

- Topological Manifold: spheres are same as cubes.
- Differential Manifold: spheres and ellipsoids are same.
- Riemannian Manifold: equipped with Riemannian metric for defining angles and distances.

Example: manifolds



Example: n-sphere

The n-sphere $\mathbb{S}^n$ is an n-dimensional manifold that can be embedded in Euclidean (n + 1)-space.

$$\mathbb{S}^n = \{x \in \mathbb{R}^{n+1} : ||x|| = r\}$$

A pair of points on the real line is 0-sphere.

A circle on $\mathbb{R}^2$ is 1-sphere.

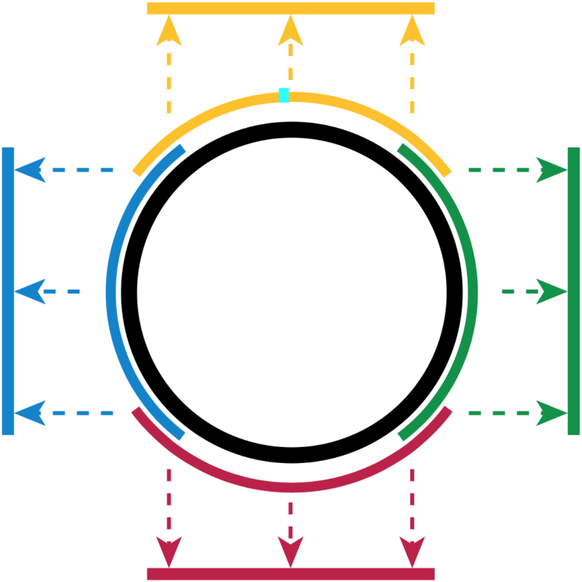
Surface of a ball is 2-sphere.

Manifolds

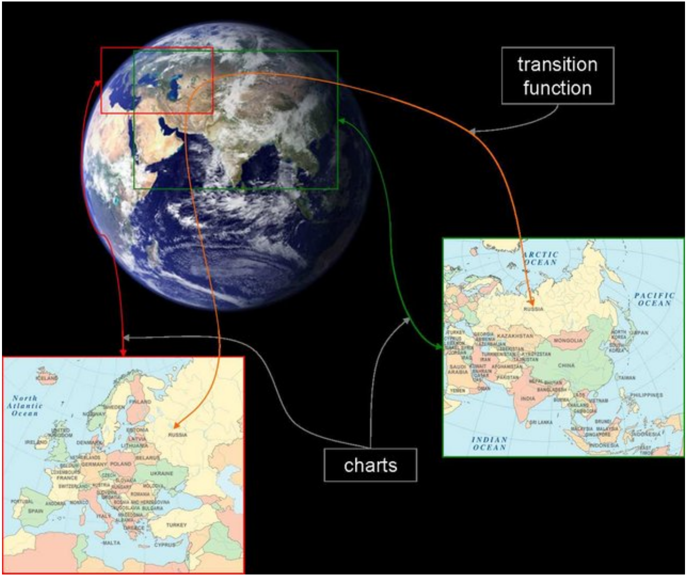
Importance

Manifolds locally resemble Euclidean space near each point, i.e., point of n-manifold has a neighbourhood that is homeomorphic to the Euclidean space of dimension n.

Example 1



Example 2

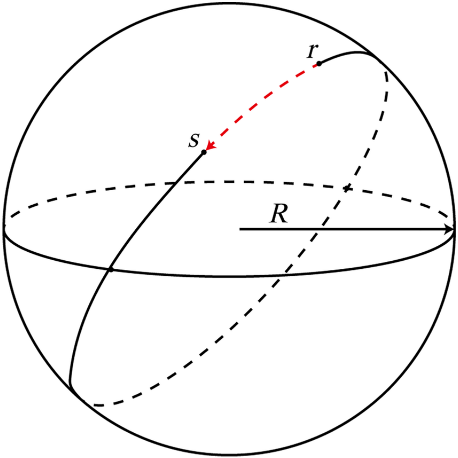

transition function
charts
ARCTIC OCEAN
PACIFIC OCEAN
INDIAN OCEAN
North Atlantic Ocean

Geodesics

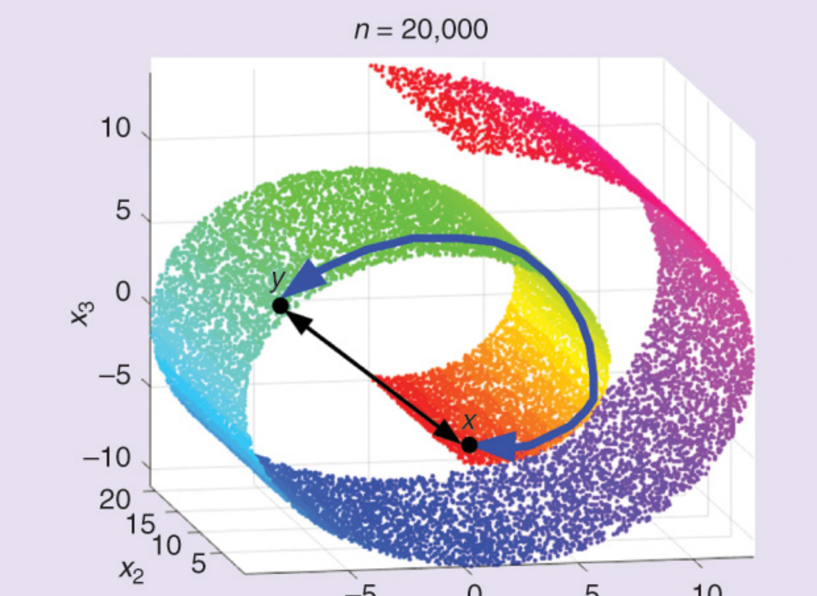
Geodesics: generalization of straight lines

- Like straight lines in Euclidean spaces, we need a 'straight line' to quantify distance.
- Geodesics are the curves on a smooth manifold that locally yield the shortest distance between two points.

Geodesics: example



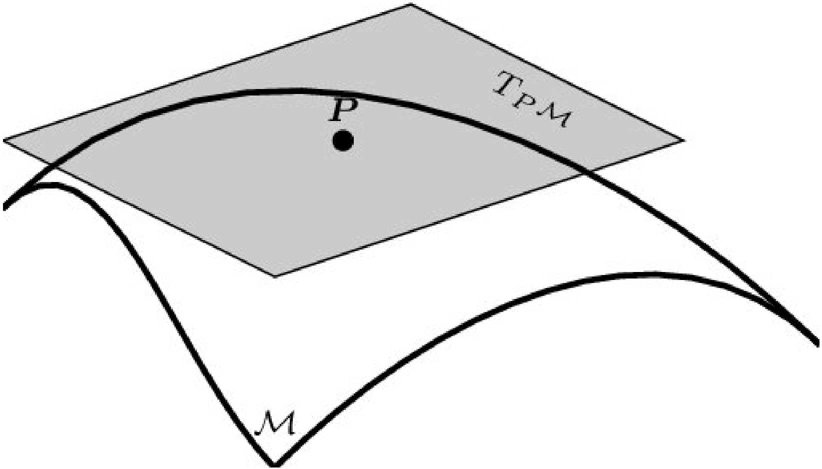
Geodesic distance



Tangent space T_pM

- A tangent space T_pM is a real vector space containing all vectors passing a point p on a n-manifold M.
- T_pM is also n-dimensional and same as the manifold M.
- Tagent space provides the best linear approximation to the manifold around p (Taylor expansion).

Tangent space: Example



Tangent space: Linear approximation

Assume we have a 2-dimensional manifold parametrized as $z = f(x, y)$. Taking the Taylor expansion

$$z = f(x, y) \approx f(p) + \nabla f(p) \cdot ((x, y) - p) + \cdots$$

which gives the tangent space or plane.

Exponential map

Let $\mathcal{D}(p)$ be an open subset of T_pM defined by

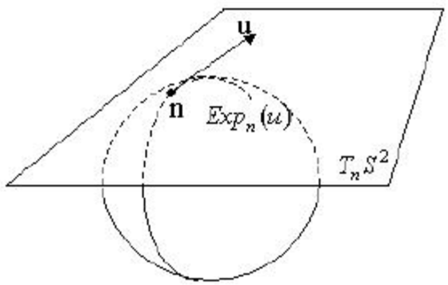
$$\mathcal{D}(p) = \{v \in T_pM | \gamma_v(1) \text{ is defined}\}$$

where γ_v is the unique geodesic with $\gamma_v(0) = p$ and $\gamma_v'(0) = v$.

Exponential map

$$\exp_p(v) : \mathcal{D}(p) \to M, \text{ given by } \exp_p(v) = \gamma_v(1)$$

Exponential map

- Local parametrization of the manifold M.
- A map from the subspace of T_pM to M.
- 

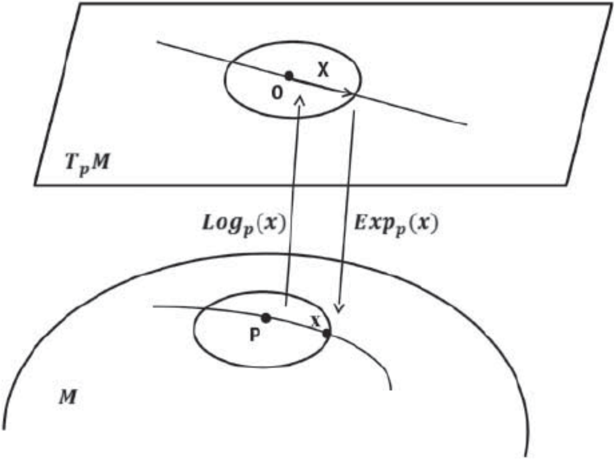
Radius of injectivity

Injectivity radius inj_p

Let M be a smooth manifold. $\forall p \in M$, the injectivity radius of M at p is the supremum such that $\exp_p$ is a diffeomorphism on the open ball $B(0, r) \subseteq T_pM$.

Diffeomorphism means that there is a smooth inverse function of $\exp_p$ which is the logarithm map $\log_p$.

example



Sphere example

The geodesic distance between any points p_1 and p_2 on a sphere is the great-circle distance,

$$d_M(p_1, p_2) = \cos^{-1}(p_1^T p_2)$$

The exponential map given a point p on the manifold is

$$\exp_p(\boldsymbol{v}) = \cos(||\boldsymbol{v}||_E)p + \sin(||\boldsymbol{v}||_E)\frac{\boldsymbol{v}}{||\boldsymbol{v}||_e}$$

The logrithm map $\log_p : M \setminus \{-p\} \to T_pM$ is the inverse of the exponential map

$$\log_p(q) = \frac{q - (p^T q)p}{||q - (p^T q)p||} d_M(p, q)$$

Notations

Let $\mathcal{M}$ be a d-dimensional complete Riemannian manifold embedded in a Euclidean space $\mathbb{R}^{d_0}$ for $d \leq d_0$. $\mathcal{X}$ denotes the sample space of $\mathcal{M}$-valued function where

$$\mathcal{X} = \{x : \mathcal{T} \to \mathcal{M} | x \in C(\mathcal{T})\}$$

for some compact interval $C(\mathcal{T}) \subset \mathbb{R}$.

Frechet mean

Like variables in Euclidean space, we need to define the mean or the expectation for the $\mathcal{M}$-valued function $X(t)$. The Frechet mean is introduced as the global minimizer of the aggregate energy function or the sum of squared distance for each $t \in \mathcal{T}$. That is,

$$M(p,t) = \sum_{j=1}^{J} d^2_{\mathcal{M}}(X_j(t), p) \text{ and } \mu_{\mathcal{M}}(t) = \underset{p \in \mathcal{M}}{\operatorname{argmin}} M(p,t)$$

Idea of RFPCA

The idea of RFPCA is to represent the variation of the inifite dimensional object X around the mean function $\mu_{\mathcal{M}}$ in a lower dimensional submanifold.

Basis expansion

The K-dimensional submanifold is defined in the following way

- Given a system of K orthonomral basis functions $\psi_k(t) \in T_{\mu_{\mathcal{M}(t)}}$.
- The submanifold

$$\mathcal{M}_K(\boldsymbol{\Psi}_K) := \{x \in \mathcal{X}, x(t) = \exp_{\mu_M(t)}(\sum_{k=1}^{K} a_k\psi_k(t))|t \in \mathcal{T}, a_k \in \mathbb{R}\}$$

The best K-dimensional approximation to X

Let $\Pi(x, \mathcal{M}_K)$ be the projection of $x \in \mathcal{X}$ on $\mathcal{M}_K$ which is defined as

$$\Pi(x, \mathcal{M}_K) := \underset{y \in \mathcal{M}_K}{\operatorname{argmin}} \int_{\mathcal{T}} d_{\mathcal{M}}(y(t), x(t))^2 dt$$

The best K-dimensional approximation to X is then find a submanifold $\mathcal{M}_K$ that minimizes

$$F_S(\mathcal{M}_K) = E \int_{\mathcal{T}} d_{\mathcal{M}}(X(t), \Pi(X, \mathcal{M}_K)(t))^2 dt$$

where each manifold is generated by K basis functions.

Logarith mapping

Instead of directly optimizing $F_S(\mathcal{M}_K)$ over a family of manifolds, the objective function is modifed by invoking tangent space approximation. Recalling injectivity radius and exponential map with its inverse logarithm map,

$$V(t) = log_{\mu_{\mathcal{M}}(t)}(X(t))$$

is well-defined for all $t \in \mathcal{T}$ as long as $X(t)$ stays within $\text{inj}_{\mu_{\mathcal{M}}(t)}$.

Optimization

A practical and tractable optimality criterion is therefore defined as

$$F_V(\mathcal{V}_K) = E(||V - \Pi(V, \mathcal{V}_K)||^2)$$

over all K-dimensional linear subspaces

$$\mathcal{V}_K(\psi_1, ..., \psi_K) = \{\sum_{k=1}^{K} a_k \psi_k(t) | a_k \in \mathbb{R}\}$$

for $\psi_k \in \mathbb{H}$ and $\psi_k(t) \in T_{\mu_{\mathcal{M}(t)}}$. This is equivalent to a multivartiate functional principal component analysis.

Covariance $G(t, s)$ ov $V(t)$

Let $G(t, s)$ be the covariance function of $V(t)$ in L^2 sense. Then,

$$G(t, s) = cov(V(t), V(s)) = E(V(t)V(s)^T)$$

since the log-mapped data $V(t) = log_{\mu_{\mathcal{M}}(t)}(X(t))$ is zero at the intrinsic mean of data on manifold.

Karhunen-Loeve decomposition of $V(t)$

Let

$$G(t,s) = \sum_{k=}^{\infty} \lambda_k \phi_k(t) \phi_k(s)^T$$

where ϕ_k's are the orthonormal vector-valued eigenfucntions with eigenvalues λ_k. Then, the Karhunen-Loeve decomposition of $V(t)$ is

$$V(t) = \sum_{k=1}^{\infty} \xi_k \phi_k(t)$$

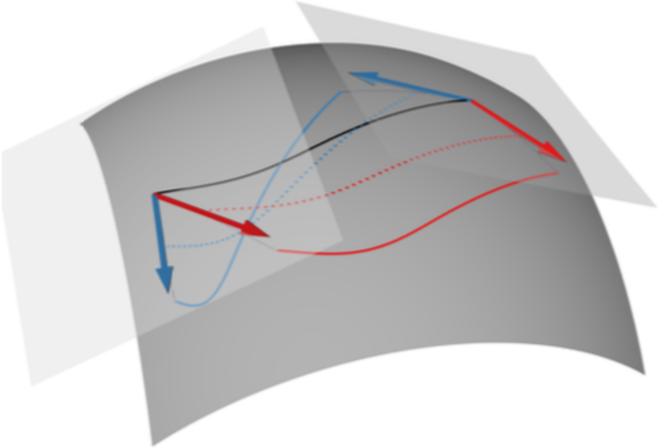
where $\xi_k = \int_{\mathcal{T}} V(t)\phi_k(t)dt$ is the k-th RFPC score.

Best K-dimensional approximation

The Best K-dimensional approximation to V is V_k that minimizes

$$E(||V - \Pi(V, \mathcal{V}_K)||^2)$$

Illustration



Choosing a suitable K

When $K = 0$, $V_0(t) = 0$ and $X_0(t) = \mu_{\mathcal{M}}(t)$. Define U_K to be the residual variance as

$$U_K = E \int_{\mathcal{T}} d_{\mathcal{M}}(X(t), X_K(t))^2 dt$$

The fraction of varinance explained by the first K components as

$$\text{FVE}_K = \frac{U_0 - U_k}{U_0}$$

$K^\star$ is chosen as the smallest K with largest FVE.

Estimation I

Given a Riemannian manifold $\mathcal{M}$ with n independent observations $X_1, ..., X_n$, which are $\mathcal{M}$-valued random functions that are distributed as X.

- Sample Frechet mean $\hat{\mu}_{\mathcal{M}}(t)$ is obtained by minimizing $M_n(p, t) = \frac{1}{n}\sum_{i=1}^{n} d_{\mathcal{M}}(X_i(t), p)^2$.
- The log-mapped data V_i: $\hat{V}_i(t) = \log_{\hat{\mu}_{\mathcal{M}}(t)}(X_i(t))$.
- Sample covariance function $\hat{G}(t, s) = \frac{1}{n}\sum_{i=1}^{n} \hat{V}_i(t)\hat{V}_i(s)$

Estimation II

- Obtain k-th eigenvalue and eigenfunction $(\hat{\lambda}_k, \hat{\phi}_k)$ of $\hat{G}$.
- Calculate kth RFPC score of each subject

$\xi_{ik} = \int_{\mathcal{T}} V_i(t)\phi_k(t)dt$.

Hence, the K-dimensional representation becomes

$$\hat{V}_{iK}(t) = \sum_{k=1}^{K} \hat{\xi}_{ik}\hat{\phi}_k(t), \hat{X}_{iK}(t) = \exp_{\hat{\mu}_{\mathcal{M}(t)}}(\sum_{k=1}^{K} \hat{\xi}_{ik}\hat{\phi}_k(t))$$

Example: Vancouver wind data

Finding Frechet mean

$$M(p,t) = \sum_{j=1}^{J} d_{\mathcal{M}}^2(X_j(t), p) \text{ and } \mu_{\mathcal{M}}(t) = \underset{p \in \mathcal{M}}{\text{argmin}} M(p,t)$$

The above can be thought of as a function of angles, and we can develope a naive gradient descent to search for the minimizer given the uniqueness in 1-sphere case.

$$p^{(i+1)} = p^{(i)} - \frac{dM(p,t)}{dp^{(i)}}$$

Go to R...

Thank You