Convolution Transforms in Manifold Learning and Dimensionality Reduction

Mijail Guillemard and Armin Iske
Department of Mathematics - University of Hamburg

NoKo 2009

Contents

Manifold Learning and Dimensionality Reduction

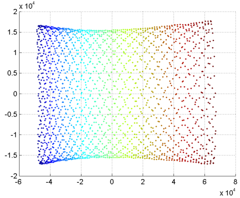
Observed data $X = \{x_1, \ldots, x_m\} \subset \mathcal{M} \subset \mathbb{R}^n$

Hypothesis:

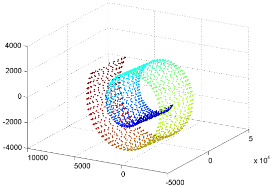
- $Y = \{y_1, \ldots, y_m\} \subset \Omega \subset \mathbb{R}^d$, $d < n$ (Ω domain)
- nonlinear map $\mathcal{A} : \Omega \to \mathbb{R}^n$, $X = \mathcal{A}(Y)$

Task: Recover Y (and Ω)

$Y \subset \Omega \subset \mathbb{R}^2$



$X \subset \mathbb{R}^{1024}$



$$\Omega \subset \mathbb{R}^d \xrightarrow{\mathcal{A}} X \subset \mathcal{M} \subset \mathbb{R}^n$$

$$X \xrightarrow{\mathcal{P}} \Omega' \subset \mathbb{R}^d$$

Motivation and Objectives

Short Term Fourier Transform - Wavelets

- Signal $f \in L^2([0,1])$ (Audio, Signal Processing)

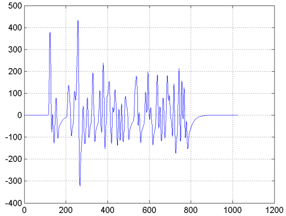
$$\mathcal{G}_g f(b,\omega) := \langle f, g_{b,\omega} \rangle = \int_0^1 f(t)\overline{g_{b,\omega}(t)}\,dt, \quad g_{b,\omega}(t) := g(t-b)e^{2\pi i \omega t}.$$

$$\boxed{X = \{x_i\}_{i=1}^n := \left\{ (f(t_j)g(t_j - k_i))_{j=1}^m \in \mathbb{R}^m \right\}_{i=1}^n}$$

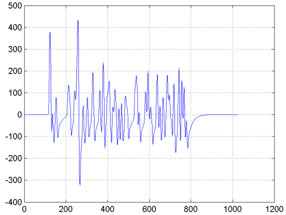
Image Analysis

- Grayscale image $F : [0,1]^2 \to [0,1]$
- Finite covering of small squares $O_i \subset [0,1]^2$

$$\boxed{X := \{x_i\}_{i=1}^n := \{F(O_i) \in \mathbb{R}^m\}_{i=1}^n}$$



$\vdots$



$$X = \{x_1, \dots, x_m\} \subset \mathcal{M} \subset \mathbb{R}^n$$

$$\Omega \subset \mathbb{R}^d \xrightarrow{\mathcal{A}} \mathcal{M} \subset \mathbb{R}^n$$

$$\mathcal{M} \subset \mathbb{R}^n \xrightarrow{T} \mathcal{M}_T \subset \mathbb{R}^n$$

$$\Omega' \subset \mathbb{R}^d \xleftarrow{\mathcal{P}} \mathcal{M}_T \subset \mathbb{R}^n$$

T Signal Transformation (Wavelets, Fourier,...)

$\mathcal{P}$ Dimensionality Reduction

Finite Sampling of a Compact Manifold

Proposition (Niyogi, Smale, Weinberger, 2008)

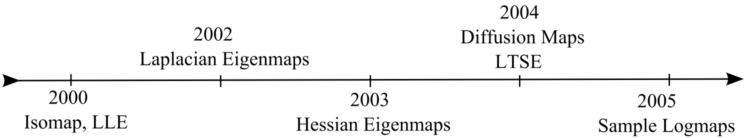
- Let $\mathcal{M}$ be a compact Riemannian submanifold of $\mathbb{R}^m$. Let $X = \{x_i\}_{i=1}^n \subset \mathbb{R}^m$ a finite collection $\epsilon/2$-dense in $\mathcal{M}$, i.e. for each $p \in \mathcal{M}$, there is an $x \in X$ such that $\|p - x\|_{\mathbb{R}^m} < \epsilon/2$. Then for any $\epsilon < \sqrt{\frac{3}{5}}\tau$, we have that $U := \bigcup_{x \in X} B_\epsilon(x)$ deformation retracts to $\mathcal{M}$, and therefore the homology of U equals the homology of $\mathcal{M}$.

- *Condition number* $1/\tau$ of the manifold.
- *Medial axis* of $\mathcal{M}$, and is defined as the closure of the set

$$G := \{x \in \mathbb{R}^m,\ \exists\, p, q \in \mathcal{M}, p \neq q\ :\ d(x, \mathcal{M}) = \|x - p\| = \|x - q\|\}$$

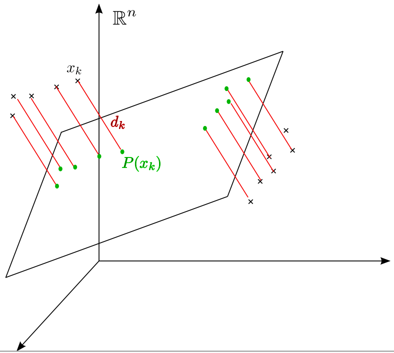
$$\tau := \inf_{p \in \mathcal{M}} d(p, \overline{G})$$

- Principal Component Analysis (PCA)
- Multidimensional Scaling (MDS)
- Isomap - Supervised Isomap
- Whitney Embedding Based Method
- Laplacian Eigenmaps
- Local Tangent Space Alignment
- Riemannian Normal Coordinates



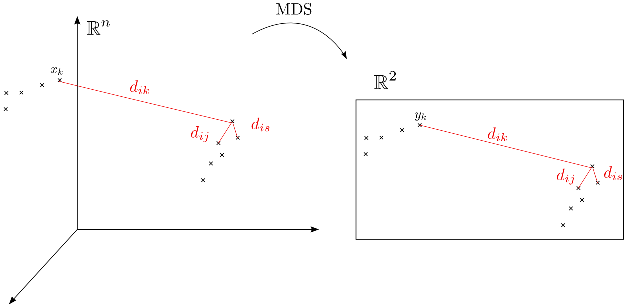
Dimensionality Reduction: PCA

- Matrix data: $X = (x_1 \ldots x_m) \in \mathbb{R}^{n \times m}$
- Problem: find projection $P : \mathbb{R}^n \to \mathbb{R}^3$ with:
- $\text{err}(P, X) = \sum_k ||x_k - P(x_k)||^2$ minimum
- $\text{var}(P(X)) = \sum_k ||P(x_k)||^2$ maximum
- "maximum" eigenvectors of the covariance matrix XX^t
- SVD of X



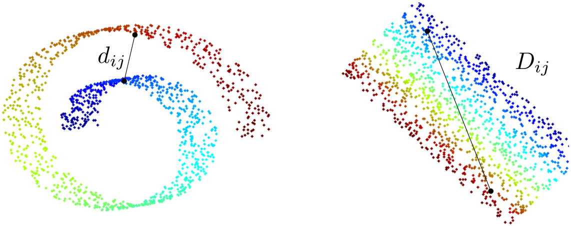
Dimensionality Reduction: MDS

- Matrix data: $X = (x_1 \ldots x_m) \in \mathbb{R}^{n \times m}$
- Problem: find a $Y = (y_1 \ldots y_m) \in \mathbb{R}^{2 \times m}$ with:
- $\text{err}(Y, X) = \sum_k (d_{ij} - ||y_i - y_j||)^2$ is minimum
- $d_{ij} = ||x_i - x_j||$.



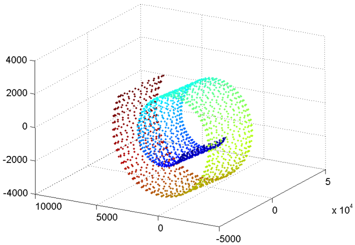
Isomap main ideas:

- Construct neighborhood graph
- Construct geodesic distances
- Use the MDS framework

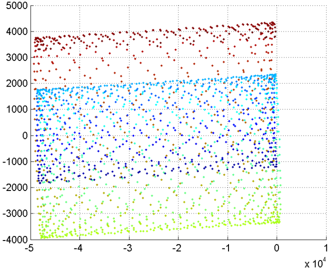


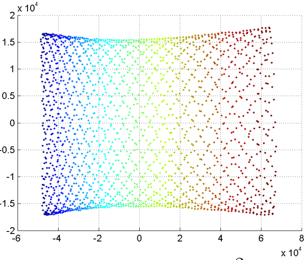
$$D_{ij} > d_{ij}$$

Swiss Roll Dataset $\mathbb{R}^3$

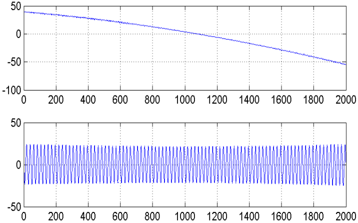


PCA projection $\mathbb{R}^2$





Isomap projection $\mathbb{R}^2$



Isomap projection: Eigenvectors

Curvature of Curves

- For a curve $r : I \longrightarrow \mathbb{R}^n$ with arc-length parametrization:

$$s(a,t) = \int_a^t \|r'(x)\|\, dx$$

the curvature is defined as

$$k(s) = \|r''(s)\|$$

- For a curve r with an arbitrary parametrization we have

$$\boxed{K^2 = \frac{\|r''\|^2 \|r'\|^2 - \langle r'', r'\rangle^2}{(\|r'\|^2)^3}}$$

Curvature for Surfaces

- Gaussian Curvature: $\boxed{K_{\mathcal{M}} := \frac{\langle R(X,Y)Y, X\rangle}{\|X\|^2\|Y\|^2 - \langle X, Y\rangle^2}}$

- Curvature Tensor: $R(X,Y)Z := \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z$

- Connection: $\nabla_{\partial_i}\partial_j = \sum_{k=1}^{n} \Gamma_{ij}^k \partial_k$

- Christoffel symbols (expressed with metric tensor)

$$\Gamma_{ij}^k = \frac{1}{2}\sum_{l=1}^{m}\left(\frac{\partial g_{jl}}{\partial x_i} + \frac{\partial g_{il}}{\partial x_j} + \frac{\partial g_{ij}}{\partial x_l}\right) g^{lk}$$

- Gaussian Curvature (expressed with Christoffel symbols)

$$\boxed{K = -\frac{1}{E}\left(\frac{\partial}{\partial u}\Gamma_{12}^2 - \frac{\partial}{\partial v}\Gamma_{11}^2 + \Gamma_{12}^1\Gamma_{11}^2 - \Gamma_{11}^1\Gamma_{12}^2 + \Gamma_{12}^2\Gamma_{12}^2 - \Gamma_{11}^2\Gamma_{22}^2\right)}$$

Curvature Distortion and Convolution

- Effect of the convolution map T on the $\mathcal{M}$ (and the dataset X).

- $\mathcal{M}_T = \{T(x), x \in \mathcal{M}\} \quad T(x) = x * h, \quad h = (h_1, \ldots, h_m)$

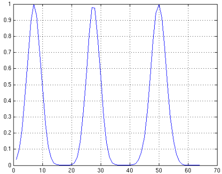
$$T = \begin{pmatrix} h_1 & 0 & \ldots & 0 \\ h_2 & h_1 & \ldots & 0 \\ h_3 & h_2 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ h_m & h_{m-1} & \ldots & h_1 \\ 0 & h_m & \ldots & h_2 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & h_m \end{pmatrix}$$

- Curves: $$\boxed{K_T^2 = \frac{\|Tr''\|^2 \|Tr'\|^2 - \langle Tr'', Tr' \rangle^2}{(\|Tr'\|^2)^3}}$$

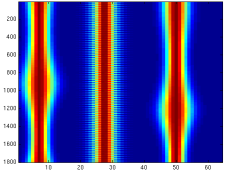
Low dimensional parametrization of scale modulated signals

$$X = \left\{ f_{\alpha^t} = \sum_{i=1}^{3} e^{-\alpha_i(t)(\cdot - b_i)^2}, \alpha \in \Omega \right\}$$

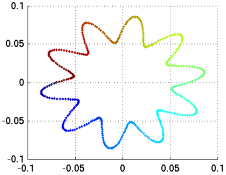
$$\Omega = \left\{ \alpha^t = (\alpha_1(t), \alpha_2(t), \alpha_3(t)), \quad t \in [t_0, t_1] \right\}$$



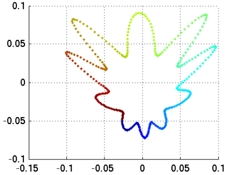
f_α



$X = \{f_\alpha, \alpha \in \Omega\}$



$\Omega \subset \mathbb{R}^3$



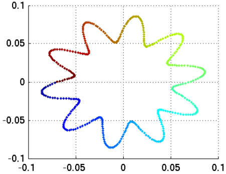
$X \subset \mathbb{R}^{64}$

Low dimensional parametrization of scale modulated signals

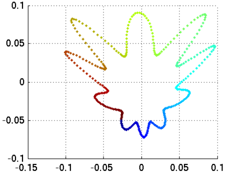
Multiresolution $V_j \longrightarrow V_{j+1} \longrightarrow V_{j+2} \quad \cdots$

$\qquad\qquad\qquad\qquad\searrow W_{j+1} \qquad \searrow W_{j+2} \quad \cdots$

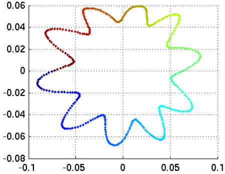
$$V_8 \oplus W_8 \oplus W_{16} \oplus W_{32} = V_{64}$$



$\Omega \subset \mathbb{R}^3$



$X \subset \mathbb{R}^{64}$



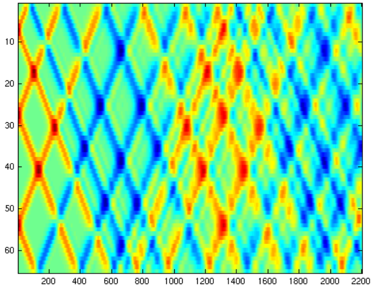
$T(X) \subset \mathbb{R}^{64}$

Example 1b: Curvature Distortion Evolution

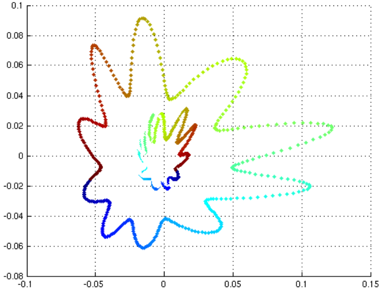
Manifold Evolution under a PDE

$$\text{Wave Equation} \qquad \frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \qquad \text{(WE)}$$

$$U_t = \left\{ u_\alpha(t,x),\ u_\alpha \text{ solution of (WE) with initial condition } f_\alpha,\ \alpha \in \Omega_0 \right\}$$



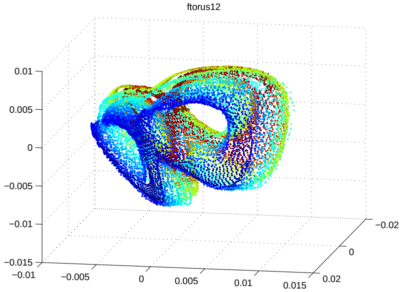
$$\left\{ u_{\alpha_0}(t,x), (t,x) \in [t_0, t_1] \times [x_0, x_1] \right\}$$

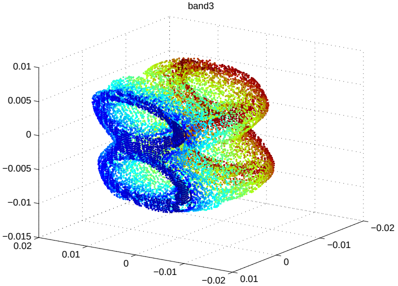


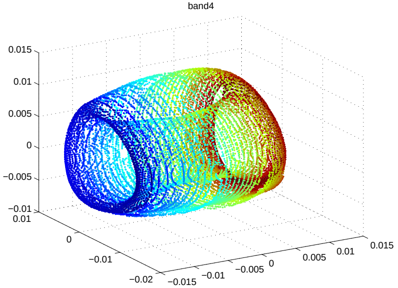
$X_0 \quad X_t$

Example 2: Topological Distortion

Torus Example: (genus 1 and genus 2)







CW-Complex

$$\emptyset \subset \mathcal{M}_0 \subset \cdots \subset \mathcal{M}_n := \mathcal{M}$$

- d-Cell σ: homeomorphic copy of B^d a closed unit ball in $\mathbb{R}^d$.
  $B^d := \{x \in \mathbb{R}^d : |x| \leq 1\}$, $S^{d-1} := \{x \in \mathbb{R}^d : |x| = 1\}$.
  Attaching Map $f : \dot{\sigma} \rightarrow M_{i-1}$ with $\dot{\sigma}$ homeomorphic to S^{d-1}.
- $\mathcal{M}_i := \mathcal{M}_{i-1} \cup_f \sigma$

Simplicial Complex

- Finite set of vertices V and set of subsets K of V with
  1- $V \subset K$
  2- $\alpha \in K$ and $\beta \subset \alpha$, then $\beta \in K$.

Main Theorem of Morse Theory

- Let $f : \mathcal{M} \to \mathbb{R}$ is a smooth function, and $p \in \mathcal{M}$ is a non-degenerate critical point of f with index τ, and $f(p) = q$. If $f^{-1}([q-\epsilon, q+\epsilon])$ is compact and contains no critical points besides p, then $M^{p-\epsilon}$ is homotopy equivalent to $M^{p+\epsilon}$ with a τ-cell attached.

Main Theorem of Discrete Morse Theory

- Suppose K is a simplicial complex with a discrete Morse function. Then K is homotopy equivalent to a CW complex with exactly one cell of dimension p for each critical simplex of dimension p.