Analysis of High-Dimensional Signal Data by Manifold Learning and Convolution Transforms

Mijail Guillemard and Armin Iske
Department of Mathematics - Hamburg University

Sampta 2009

Contents

Manifold Learning and Dimensionality Reduction

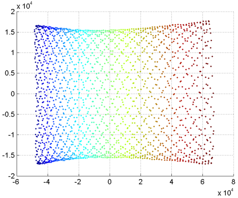
Observed data $X = \{x_1, \ldots, x_m\} \subset \mathcal{M} \subset \mathbb{R}^n$

Hypothesis:

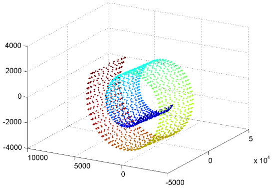
- $Y = \{y_1, \ldots, y_m\} \subset \Omega \subset \mathbb{R}^d$, $d < n$ (Ω domain)
- nonlinear map $\mathcal{A} : \Omega \to \mathbb{R}^n$, $X = \mathcal{A}(Y)$

Task: Recover Y (and Ω)

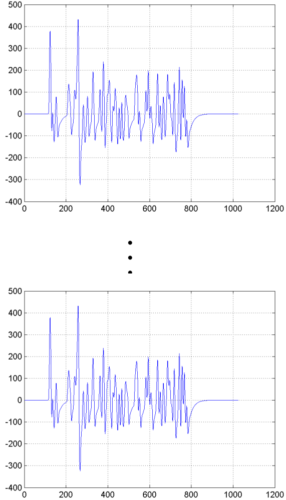
$Y \subset \Omega \subset \mathbb{R}^2$



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



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



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

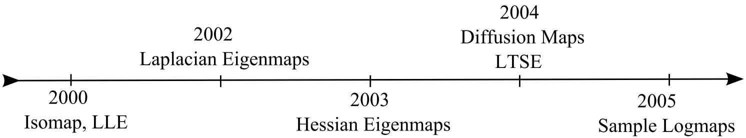
$$\begin{array}{ccc} \Omega \subset \mathbb{R}^d & \xrightarrow{\mathcal{A}} & \mathcal{M} \subset \mathbb{R}^n \\ & & \downarrow T \\ \Omega' \subset \mathbb{R}^d & \xleftarrow{\mathcal{P}} & \mathcal{M}_T \subset \mathbb{R}^n \end{array}$$

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

$\mathcal{P}$ Dimensionality Reduction

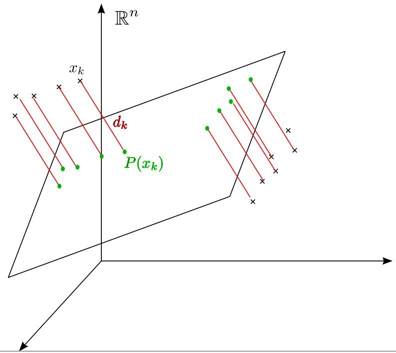
Manifold Learning Techniques

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



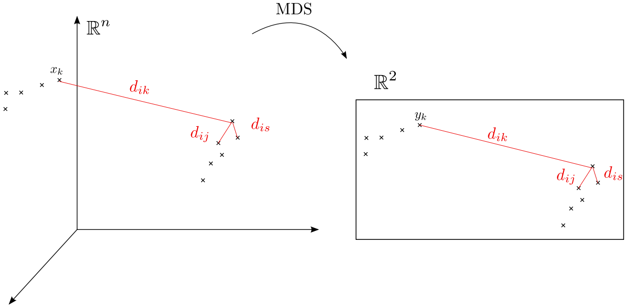
Dimensionality Reduction: PCA

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



Dimensionality Reduction: MDS

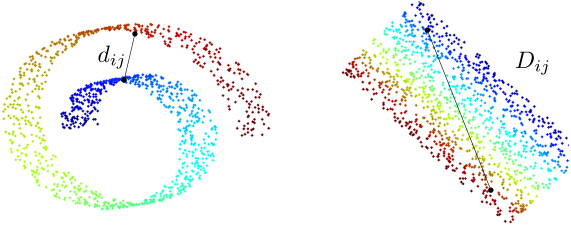
- Matrix data: $X = (x_1 \dots x_m) \in \mathbb{R}^{n \times m}$
- Problem: find a $Y = (y_1 \dots y_m) \in \mathbb{R}^{2 \times m}$ with:
- $\mathrm{err}(Y, X) = \sum_k (d_{ij} - ||y_i - y_j||)^2$ is minimum
- $d_{ij} = ||x_i - x_j||$.



Dimensionality Reduction: Isomap

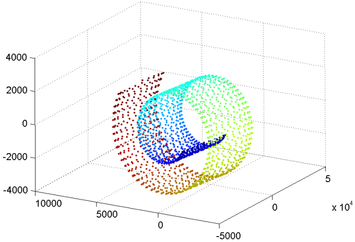
Isomap main ideas:

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

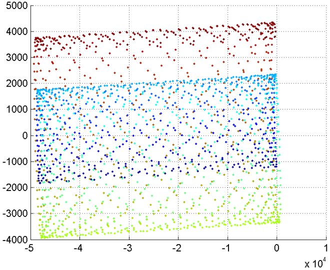


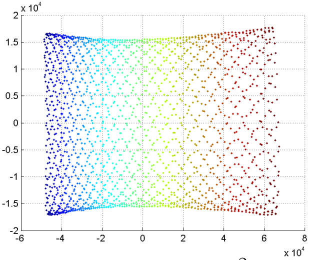
$$D_{ij} > d_{ij}$$

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

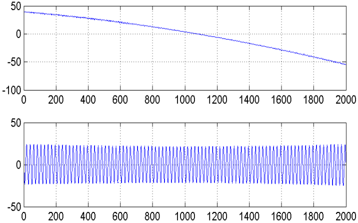


PCA projection $\mathbb{R}^2$





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

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

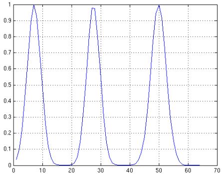
$$T = \begin{pmatrix} h_1 & 0 & \ldots & 0 \\ h_2 & h_1 & \ldots & 0 \\ h_3 & h_2 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ h_m & h_{m-1} & \ldots & h_1 \\ 0 & h_m & \ldots & h_2 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & h_m \end{pmatrix}$$

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

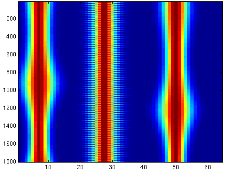
Low dimensional parametrization of scale modulated signals

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

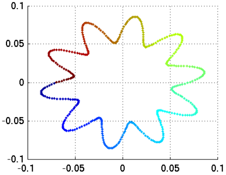
$$\Omega = \left\{ \alpha^t = (\alpha_1(t), \alpha_2(t), \alpha_3(t)), \quad t \in [t_0, t_1] \right\}$$



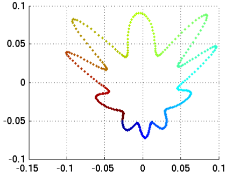
f_α



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



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

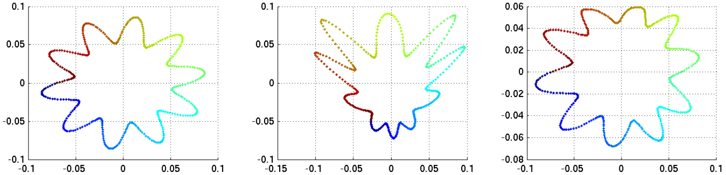


$X \subset \mathbb{R}^{64}$

Low dimensional parametrization of scale modulated signals

Multiresolution
$$\begin{array}{ccccccc} V_j & \longrightarrow & V_{j+1} & \longrightarrow & V_{j+2} & & \cdots \\ & \searrow & & \searrow & & & \\ & & W_{j+1} & & W_{j+2} & & \cdots \end{array}$$

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



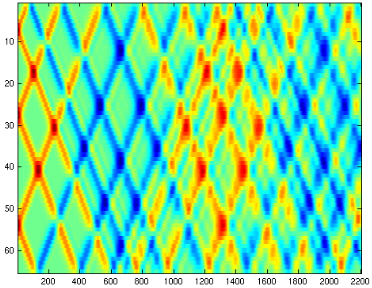
$\Omega \subset \mathbb{R}^3$ $\quad X \subset \mathbb{R}^{64}$ $\quad T(X) \subset \mathbb{R}^{64}$

Example 1b: Curvature Distortion Evolution

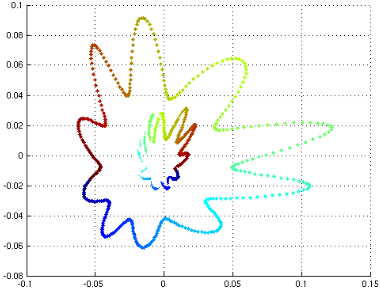
Manifold Evolution under a PDE

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

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



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



$X_0 \quad X_t$

Torus Example: (genus 1 and genus 2)