Introduction to Manifold Learning

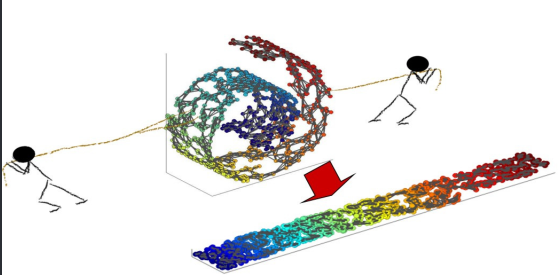
A Geometry View on Machine Learning

Xiaoyu Xue

Outline for Section 1

Manifold Learning

- The data space may not be a Euclidean space, but a nonlinear manifold
- Unfold a manifold, and preserve the geometry structure.
- Euclidean distance $\Longrightarrow$ geodesic distance



Manifold Learning

Find a Euclidean embedding, and then perform traditional learning algorithms in the Euclidean space.

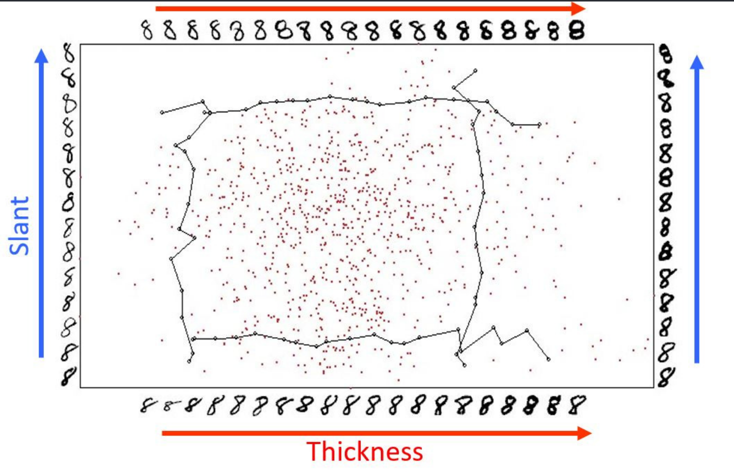
Definition of Manifold Learning

Given data points $\mathbf{x}_1, \ldots, \mathbf{x}_m \in \mathscr{M} \subset \mathbb{R}^n$, try to find a map $f : \mathscr{M} \to \mathbb{R}^d$, $d \ll n$, where $f = (f_1, \ldots, f_n)$, $f_i : \mathscr{M} \to \mathbb{R}$

- The manifold is unknown! We have only samples!
- How to compute the distance on $\mathscr{M}$?
- How to find the mapping function f

Manifold of Face Images

Manifold of Handwritten Digits



Outline for Section 2

PCA: Traditional Dimensionality Reduction Method

Principal Component Analysis using linear projection to project data to some directions which have maximum variances

$$\begin{aligned}\mathbf{p}_{opt} &= \arg\max_{\mathbf{p}} \sum_{i=1}^{m} (y_i - \bar{y})^2 \\ &= \arg\max_{\mathbf{p}} \mathbf{p}^T C \mathbf{p} \\ & \quad s.t. \mathbf{p}^T \mathbf{p} = 1\end{aligned}$$

- If the manifold is linear, PCA can find the optimal result
- PCA can not process nonlinear manifold

MDS and ISOMAP

Multidimensional scaling tries to preserve the Euclidean distances

$$\Delta := \begin{pmatrix} \delta_{1,1} & \delta_{1,2} & \dots & \delta_{1,m} \\ \delta_{2,1} & \delta_{2,2} & \dots & \delta_{2,m} \\ \vdots & \vdots & & \vdots \\ \delta_{m,1} & \delta_{m,2} & \dots & \delta m, m \end{pmatrix}$$

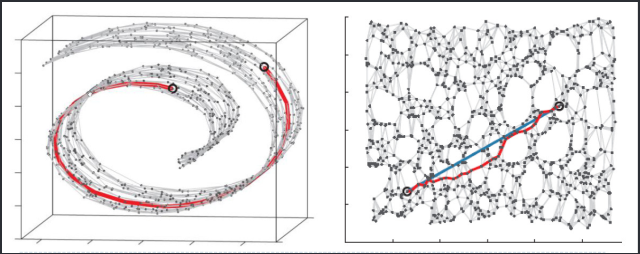
The δ is the Euclidean distance of every two points $\delta_{ij} = ||\mathbf{x}_i - \mathbf{x}_j||$

$$\min_{\mathbf{y}_1,\dots,\mathbf{y}_m} \sum_{i<j} \left(||\mathbf{y}_i - \mathbf{y}_j|| - \delta_{i,j}\right)^2, \quad \dim(\mathbf{y}_i) \ll \dim(\mathbf{x}_i)$$

MDS and ISOMAP

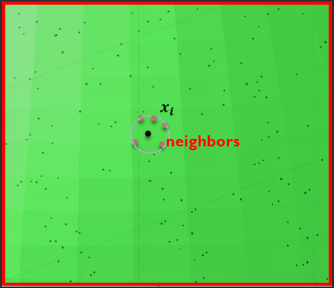
ISOMAP tries to keep the geodesic distances instead of the Euclidean distances.

- How to evaluate the geodesic distances with limited samples?
- Construct the adjacency Graph, and calculate the shortest distances (Dijstra or Floyd algorithm)



Local Linear Embedding

Local Linear Embedding(2000 Science) is another famous manifold learning method. It tries to preserve the local linear relationship.

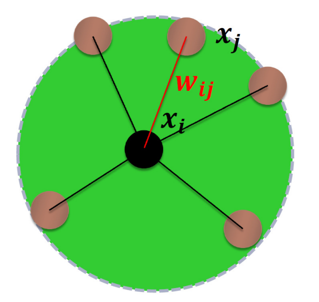


$$\min \epsilon(W) = \min \sum_i \|\mathbf{x}_i - \sum_j W_{ij}\mathbf{x}_j\|$$

$$s.t. \sum_j W_{ij} = 1$$

Local Linear Embedding

Local Linear Embedding(2000 Science) is another famous manifold learning method. It tries to preserve the local linear relationship.



$$\min \Phi(\mathbf{y}) = \min \sum_i ||\mathbf{y}_i - \sum_j W_{ij}\mathbf{y}_j||^2$$

Laplace Eigen Map

In Laplace Eigen Map, a conclusion has been proofed:

$$|f(\mathbf{z}) - f(\mathbf{x})| < \|\nabla f(\mathbf{x})\| \cdot \|\mathbf{z} - \mathbf{x}\| + o(\|\mathbf{z} - \mathbf{x}\|)$$

- If $\mathbf{x}_i$ and $\mathbf{x}_j$ are close to each other and the gradient of map f is small, we can sure that $f(\mathbf{x}_i)$ and $f(\mathbf{x}_j)$ preserve local structure.

Construct Laplace matrix and get object function

$$L\mathbf{y} = \lambda \mathbf{y}$$

Global vs Local

- Global method : ISOMAP
- Local method : LLE, LE
- Global method can keep more informations of data
- But the amount of computation of Global methods is huge

Text blocks

In plain, example, and alert flavour

This text is highlighted.

A plain block

This is a plain block containing some highlighted text.

An example block

This is an example block containing some highlighted text.

An alert block

This is an alert block containing some highlighted text.

Definitions, theorems, and proofs

All integers divide zero

Definition

$\forall a, b \in \mathbb{Z} : a \mid b \iff \exists c \in \mathbb{Z} : a \cdot c = b$

Theorem

$\forall a \in \mathbb{Z} : a \mid 0$

Proof

$\forall a \in \mathbb{Z} : a \cdot 0 = 0$ $\square$

Numerals and Mathematics

Formulae, equations, and expressions

1234567890 1234567890 $\hat{x}, \check{x}, \tilde{a}, \bar{a}, \dot{y}, \ddot{y}$ $\iint f(x, y, z)\, \mathrm{dxdydz}$

$$\frac{1}{1+\frac{1}{2+\frac{1}{3+x}}} + \frac{1}{1+\frac{1}{2+\frac{1}{3+x}}} \qquad F : \begin{vmatrix} F''_{xx} & F''_{xy} & F'_x \\ F''_{yx} & F''_{yy} & F'_y \\ F'_x & F'_y & 0 \end{vmatrix} = 0$$

$$\iint_{\mathbf{x}\in\mathbb{R}^2} \langle \mathbf{x}, \mathbf{y} \rangle\, \mathrm{d}\mathbf{x} \qquad \overline{\overline{a\alpha}^2 + \underline{b\beta} + \overline{\overline{d\delta}}} \qquad]0,1[+ \lceil x \rfloor - \langle x, y \rangle$$

$$\begin{aligned} e^x \quad \approx \quad & 1 + x + x^2/2! + \\ & + x^3/3! + x^4/4! \end{aligned} \qquad \binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$$

Figures

Tables, graphs, and images

Faculty	With TEX	Total	%
Faculty of Informatics	1 716	2 904	59.09
Faculty of Science	786	5 275	14.90
Faculty of Economics and Administration	64	4 591	1.39
Faculty of Arts	69	10 000	0.69
Faculty of Medicine	8	2 014	0.40
Faculty of Law	15	4 824	0.31
Faculty of Education	19	8 219	0.23
Faculty of Social Studies	12	5 599	0.21
Faculty of Sports Studies	3	2 062	0.15

Table: The distribution of theses written using TEX during 2010–15 at MU

Figures

Tables, graphs, and images

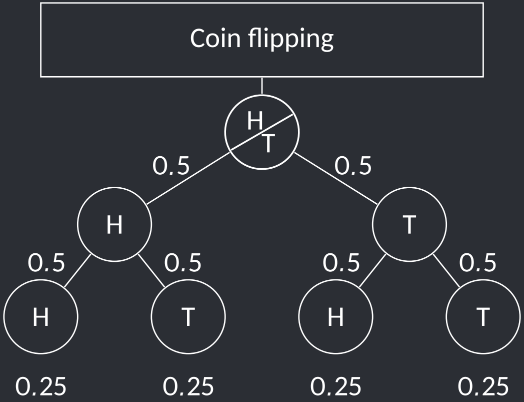


Figure: Tree of probabilities – Flipping a coin[1]

[1]A derivative of a diagram from `texample.net` by cis, CC BY 2.5 licensed

Code listings

An example source code in C

```c
#include <stdio.h>
#include <unistd.h>
#include <sys/types.h>
#include <sys/wait.h>

// This is a comment
int main(int argc, char **argv)
{
	while (--c > 1 && !fork());
	sleep(c = atoi(v[c]));
	printf("%d\n", c);
	wait(0);
	return 0;
}
```

Citations

TeX, LaTeX, and Beamer

TeX is a programming language for the typesetting of documents. It was created by Donald Erwin Knuth in the late 1970s and it is documented in *The TeXbook* [1].

In the early 1980s, Leslie Lamport created the initial version of LaTeX, a high-level language on top of TeX, which is documented in *LaTeX: A Document Preparation System* [2]. There exists a healthy ecosystem of packages that extend the base functionality of LaTeX; *The LaTeX Companion* [3] acts as a guide through the ecosystem.

In 2003, Till Tantau created the initial version of Beamer, a LaTeX package for the creation of presentations. Beamer is documented in the *User's Guide to the Beamer Class* [4].

Bibliography

TeX, LaTeX, and Beamer

[1] Donald E. Knuth. *The TeXbook*. Addison-Wesley, 1984.

[2] Leslie Lamport. *LaTeX: A Document Preparation System*. Addison-Wesley, 1986.

[3] M. Goossens, F. Mittelbach, and A. Samarin. *The LaTeX Companion*. Addison-Wesley, 1994.

[4] Till Tantau. *User's Guide to the Beamer Class Version 3.01*. Available at `http://latex-beamer.sourceforge.net`.

[5] A. Mertz and W. Slough. Edited by B. Beeton and K. Berry. *Beamer by example* In TUGboat, Vol. 26, No. 1., pp. 68-73.

Citations

TeX, LaTeX, and Beamer

TeX is a programming language for the typesetting of documents. It was created by Donald Erwin Knuth in the late 1970s and it is documented in *The TeXbook* [1].

In the early 1980s, Leslie Lamport created the initial version of LaTeX, a high-level language on top of TeX, which is documented in *LaTeX: A Document Preparation System* [2]. There exists a healthy ecosystem of packages that extend the base functionality of LaTeX; *The LaTeX Companion* [3] acts as a guide through the ecosystem.

In 2003, Till Tantau created the initial version of Beamer, a LaTeX package for the creation of presentations. Beamer is documented in the *User's Guide to the Beamer Class* [4].

Bibliography

TeX, LaTeX, and Beamer

[1] Donald E. Knuth. *The TeXbook*. Addison-Wesley, 1984.

[2] Leslie Lamport. *LaTeX: A Document Preparation System*. Addison-Wesley, 1986.

[3] M. Goossens, F. Mittelbach, and A. Samarin. *The LaTeX Companion*. Addison-Wesley, 1994.

[4] Till Tantau. *User's Guide to the Beamer Class Version 3.01.* Available at `http://latex-beamer.sourceforge.net`.

[5] A. Mertz and W. Slough. Edited by B. Beeton and K. Berry. *Beamer by example* In TUGboat, Vol. 26, No. 1., pp. 68-73.