MATHEMATICS FOR COMPUTER VISION

WEEK 6

MANIFOLDS

Dr Fabio Cuzzolin
MSc in Computer Vision
Oxford Brookes University
Year 2013-14
OUTLINE OF WEEK 6

- topics: manifolds and dimensionality reduction

- Topological and differential manifolds
- Riemannian metrics
- Geodesics
- Dimensionality reduction
- Laplacian embeddings
  - Locally Linear Embedding
- Geodesic embeddings
  - ISOMAP
TOPOLOGICAL MANIFOLDS

Week 6 – Manifolds
A BRIEF HISTORY OF SPACE(S)

- Euclidean space is the most common space we deal with
- Cartesian coordinates, vectors...
- for a long time, only space/geometry people considered possible (Euclid’s elements)
- then, non-Euclidean geometries were introduced in 1800s (Lobachevsky, Riemann, Gauss)
- Gauss in particular, defined abstract spaces that need not be immersed in a Euclidean one
- you can defined them on their own
EXAMPLE OF MANIFOLD: THE CIRCLE

- simplest example: the circle $x^2 + y^2 = 1$
- Not Euclidean: it cannot be stretched to become the real line, but..
- .. the four pieces in color (extrema excluded) can
- in each piece, a point on the circle has coordinates in $\mathbb{R}$
  - e.g., points on the top (yellow) arc have coordinate $x \in [-1,1]$
- the manifold circle can be divided into pieces, each with its coordinate chart
- these pieces can overlap, in general
- each piece can be stretched to become the Euclidean space
- can be “covered” by an atlas of charts
NOTION OF (TOPOLOGICAL) MANIFOLD

- general manifold: a space (a set of points in which each point has neighbourhood, which meet certain properties) which is “locally isomorphic” to the Euclidean space
- this means that every point has a neighbourhood that can be mapped (continuously and with continuous inverse) to a ball in $\mathbb{R}^n$
  \[ B_r(x) = \{ y \in \mathbb{R}^n : d(x, y) < r \} \]
- map is called homeomorphism (right)
- If all neighbourhoods are mapped to balls of the same dimension $n$, the manifold has dimension $n$
OTHER EXAMPLES

- more examples:
  - they can have holes, self-intersections..
  - Moebius strip
  - Whitehead manifold
  - triple torus
ELEMENTS OF A MANIFOLD

- the ingredients of a (topological) manifold are:
  - an atlas of coordinate charts
  - a change of coordinates between charts, in the regions where they overlap

- various types of manifolds:
  - topological manifold: the type we saw
  - with boundary (have edges)
    - sheet of paper is 2-dim manifold, with 1-dim boundary
  - differentiable manifold
    - they generalise surfaces
  - Riemannian manifold -> manifold with a metric
  - many others

- we have a look only at these two types
DIFFERENTIABLE AND RIEMANNIAN MANIFOLDS

Week 6 – Manifolds
FROM SURFACES TO MANIFOLDS

- we have studied surfaces in Week 5 (Geometry)
- surfaces are embedded in the Euclidean space
- we know that when they are smooth, they admit tangent vectors
- at each point they have a local curvature
- question: can we get rid of the surrounding 3D Euclidean space, and define “tangent”, “curvature” intrinsically on a space? Yes!
- add additional structure to a manifold (which is defined intrinsically, without reference to a surrounding space)
- surfaces are just a special case of differentiable manifold
DIFFERENTIABLE MANIFOLDS

- manifold in which the change of coordinate between two charts is differentiable, as a function
  - on such a space, one can define differentiable functions $f$
    - have to be differentiable w.r.t. any coordinate chart that include the point considered
  - can then define a directional derivative along a curve $\gamma$
  - what is a tangent vector in $m$ then? A class of curves passing through $m$ with the same directional derivative in that point
  - very abstract definition! Remember we cannot assume there is a space out there a “vector” can live in! But it works
  - set of all tangent vectors in a point: tangent space
  - You can do calculus on differentiable manifolds! We don’t ;}
RIEMANNIAN MANIFOLD

- generalises the notion of Euclidean distance
- curved spaces have a natural distance attached
- formally: a Riemannian manifold is a smooth manifold, equipped with an inner product on its tangent space at each point
  - (you can measures angles between tangent vectors)
    \[ g_p : T_p M \times T_p M \rightarrow \mathbb{R} \]
- notation:
  \[ (u, v) \quad g_p(u, v) \]
- Riemannian manifolds are metric space (Week 5)
- distance function is the length of the geodesic (shortest path) connecting any two points p
GEODESICS AND CURVATURE

- Example of geodesic distance: great circle distance on a sphere (e.g. the Earth)

- Riemannian manifolds have curvature, just like surfaces

- Gauss' *Theorema Egregium*: the curvature of a manifold can be computed just by measuring distances along paths on it

- No need to an embedding 3D space! Curvature can also be computed intrinsically, just like tangent vectors
EXAMPLES AND VISION APPLICATIONS

- A number of entities of interest in vision live on a Riemannian manifold.
- The space of all rigid motions in the three-dimensional space is a Riemannian manifold: $SE(3)$.

- Common families of probability distributions live on a Riemannian manifold:
  - Exponential families, work of Amari, Fisher metric.
- Same for common types of dynamical models:
  - Auto-regressive models AR(2).
- Some dimensionality reduction techniques are based on geodesic distances (e.g. ISOMAP, next section).
APPLICATION: GENERAL RELATIVITY

- coolest application of Riemannian geometry
- Einstein’s general relativity theory
- it is a general theory of gravitation
- idea: gravity of masses bends the space-time
- GR space-time is a Riemannian manifold
NONLINEAR DIMENSIONALITY REDUCTION

Week 6 – Manifolds
NONLINEAR DIMENSIONALITY REDUCTION

- in many machine learning and vision problem, data live in a very high-dimensional space
  - e.g., images are vectors of as many real numbers as there are pixels, videos even worse!
  - algorithms have very high time complexity, curse of dimensionality

- in many cases, though, the data really live in a lower-dimensional “surface”, embedded in this higher-dim space

- how to find the lower-dim space the data actually live in?

- linear methods: PCA is one way of doing that
  - also ICA (independent component analysis)

- nonlinear dimensionality reduction: model the fact that the input data varies in a non-linear way
  - i.e., they live on a curved manifold
EXAMPLES

- images of people’s faces are a very small region of all possible images
- flowers (above)
- hand digits (right)
TECHNIQUES

- lots of techniques for nonlinear dimensionality reduction
- some are global methods: the mapping is based on the **global** structure of the data (e.g. distances between all pairs of points)
  - ISOMAP
- some other are local: mapping is based on the **local structure of the data** (e.g. distances within a neighbourhood of each point)
  - Locally Linear Embedding
  - Laplacian eigenmaps
- list goes on and on
  - Kohonen maps (neural network)
  - nonlinear PCA
  - Kernel PCA
  - Diffusion maps
  - Multidimensional scaling
MULTIDIMENSIONAL SCALING (MDS)

- global method, pretty old idea
- suppose we have a set of objects $x_i, i = 1, ..., N$
- we do not have a vectorial representation for them ..
  - cannot represent them as points of a Cartesian space
- .. but we their pairwise similarities/dissimilarities $d_{ij}$
- then we can come up with a vector representation!
- ..assuming the $d$ is a Euclidean distance
- algorithm:
  - given the matrix of dissimilarities $D = [d_{ij}, i, j = 1, ..., N]$
  - Build the centered inner product matrix $B = HDH$ where $H = I - \frac{1}{N} 11^T$
  - apply SVD and find the largest $k$ eigenvalues of $B$
  - they form a matrix $N$ by $k$
  - the corresponding eigenvectors form the coordinates of the datapoints
  - i.e., the coordinates of $x_i$ form the $i$-th row of such matrix
ISOMAP

- it is a also global method
- what is the dissimilarity matrix $D$ is not Euclidean? We need to assume that datapoints live on a curved manifold
- the mapping to a lower-dimensional space depends on the (approximate) geodesic distance between all pairs of points
- this distance is measured along a path on a weighted graph
- algorithm (isomap.stanford.edu)
  1. given a training set of points, build a weighted graph $G$ which approximates the underlying manifold
  2. for each pair of points in the training set, their geodesic distance on the graph $G$ is calculated (for instance by Dijkstra’s algo)
  3. multidimensional scaling (MDS) is applied to the resulting matrix of pairwise distances (affinity matrix)
  4. the output is a lower-dim vector for each data point
EXAMPLES OF ISOMAP REDUCTION

- Swiss roll example

- can improve cluster separation
Spectral methods

- given a dataset of points \( \{X_i, i=1,..,N\} \) ...
- ... they compute an affinity matrix

\[
A(i,j) = d(X_i,X_j)
\]

- apply SVD to this affinity matrix
- this yields a list of eigenvalues and associated eigenvectors
- a number of eigenvectors (say, D) are selected, and used to build a D-dimensional "embedded cloud" of points \( \{Y_i, i=1,..,N\} \)
Laplacian Methods

- the affinity matrix is the Laplacian operator, or some function of it
- **Graph Laplacian**: operator on functions $f$ defined on sets $X$ of points (the nodes of the graph) of the form

$$L[f]_i = \sum_{j \in N(i)} w_{ij} (f_i - f_j)$$

- maps each such function $f$ to another function $L[f]$
- $N(i)$ is the set of neighbors of $X_i$
- $f_i$ is the value of the function $f$ on $X_i$
LAPLACIAN EIGENFUNCTIONS

- Laplacian eigenfunctions/values have nice topological properties
- Eigenvalues are invariant for volume-preserving transformations
- Eigenfunctions form a “base” for all functions on $X$
- Their level sets are related to protrusions and symmetries of the underlying cloud
- E.g. level sets of Laplacian eigenfunctions on an action figure (below)
**Locally Linear Embedding**

- For each data point we compute the weights $W_{ij}$ that best reconstruct $X_i$ from its neighbors:

$$\arg\min_W \sum_i |X_i - \sum_j W_{ij} X_j|^2$$

- Low-dimensional embeddings $Y_i$ are obtained by

$$\arg\min_Y \sum_i |Y_i - \sum_j W_{ij} Y_j|^2$$

- I.e., local neighbors are the same, subject to

$$\frac{1}{N} \sum_{i=1}^{N} Y_i \otimes Y_i = I$$

- The affinity matrix is $M = (I - W)^T(I - W)$

- Optimal embedding $\rightarrow$ bottom $d+1$ eigenvectors (but last one)
LLE ALGORITHM

- a diagram explaining the LLE algorithm
**LAPLACIAN EIGEMAPS**

- local method, uses spectral techniques
- builds a graph from neighbourhood information
- graph is a discrete approximation of the underlying manifold
  - weight of edge can be binary, or a function of the distance between the two points $i$ and $j$
- minimises a cost function based on this graph
- result: close original datapoints are mapped to close points after embedding *(preserving local structure)*
- embedding dimensions are the *eigenvectors of the Laplace-Beltrami operator*

\[ Lf = \lambda Df \]

- where $L = W - D$ is the Laplacian matrix, $W$ is the adjacency matrix of the graph, and $D$ is diagonal with $D_{i,i} = \sum_j w_{ij}$
- seems to produce better results for non-convex manifolds
AN EMPRICAL COMPARISON

- different embedding techniques have different performances
- A Swiss roll example from Cao and Chen, 2012

Fig. 1. Examples of Swiss roll. (a) Sample data. (b) RLEM. (c) Laplacian eigenmaps. (d) Diffusion maps. (e) LLE. (f) Isomap.
SUMMARY OF WEEK 6

- topological manifolds as generalisation of Euclidean space
- atlases and charts
- different types of manifolds
- differentiable manifolds
  - tangent space
- Riemannian manifolds
  - metrics, curvature and geodesics
- dimensionality reduction
  - Multidimensional scaling (MDS)
  - Isometric mapping (ISOMAP)
  - Locally Linear Embedding (LLE)
  - Laplacian Eigenmaps