Manifolds, Riemannian Geometry Lie Groups and Harmonic Analysis, With Applications CIS610, Spring 2011

Jean Gallier
Department of Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104, USA
e-mail: jean@cis.upenn.edu

January 13, 2011

Chapter 1

Introduction: Problems, Questions, Motivations

Let M be some "space" of data. Often, the space M has some geometric and topological structure. Sometimes, there is a way of multiplying the objects in M which makes M into a group.

Problems we often want to solve:

(1) *Interpolate data*: Given $a, b \in M$, compute

$$(1-\lambda)a + \lambda b, \quad \lambda \in [0,1].$$

(2) Compute the *mean* of a finite set of data (a *sample*) $a = \{a_1, \ldots, a_n\}$,

$$\overline{a} = \frac{a_1 + \cdots + a_n}{n}.$$

(3) Compute the (*sample*) *variance* of a finite set of data $a = \{a_1, \dots, a_n\}$,

$$\operatorname{var}(a) = \frac{\sum_{i=1}^{n}(a_i - \overline{a})^2}{n-1}.$$

(4) Given two samples, $a = \{a_1, \dots, a_n\}$ and $b = \{b_1, \dots, b_n\}$, find their (*sample*) *covariance*,

$$\operatorname{cov}(a, b) = \frac{\sum_{i=1}^{n}(a_i - \overline{a})(b_i - \overline{b})}{n-1}$$

The covariance of a and b measures how a and b varies from the mean with respect to each other. Note, $\operatorname{cov}(a, a) = \operatorname{var}(a)$. If $\operatorname{cov}(a, b) = 0$ we say that a and b are *uncorrelated*.

(5) Let X be an $n \times d$ matrix with entries in M. We denote the ith row of X by X_i, with $1 \leq i \leq n$. The jth column is denoted C_j $(1 \leq j \leq d)$. (It is sometimes called a *feature vector*, but this terminology is far from being universally accepted. In fact, many people in computer vision call the data points, X_i, feature vectors!)

Perform *PCA analysis* of the data set, $X_1, \ldots, X_n$, with $X_i \in M^d$.

The purpose of *principal components analysis* (PCA) is to identify patterns in data and understand the *variance-covariance* structure of the data.

This is useful for

(a) Data reduction: Often much of the variability of the data can be accounted for by a smaller number of *principal components*.

(b) Interpretation: PCA can show relationships that were not previously suspected.

(6) Assume M is some kind of geometric space.
For $a, b \in M$, find a *shortest path* from a to b.

If M is a vector space, there are good methods for solving these problems.

However, if M is a "curved space" it may be very difficult, even impossible, to solve these problems.

If fact, it is not clear that the notion of mean makes any sense if M is not a vector space!

For example, in medical imaging, DTI (*Diffusion Tensor Imaging*) produces a 3D **s**ymmetric, **p**ositive **d**efinite matrix, at each voxel of an imaging volume. In brain imaging, this method is used to track the white matter fibres, which demonstrate higher diffusivity of water in the direction of the fibre.

One would hope to produce statistical atlases from diffusion tensor images and to understand the anatomical variability caused by a disease.

Unfortunately, the space of $n \times n$ *symmetric, positive definite matrices*, $\mathbf{SPD}(n)$, is not a vector space. Consequently, standard linear statistical methods do not apply.

Recall that a matrix, A, is in $\mathbf{SPD}(n)$ iff it is symmetric and if its eigenvalues are all strictly positive. The second condition is equivalent to

$$x^\top A x > 0 \qquad \text{for all} \quad x \neq 0 \quad (x \in \mathbb{R}^n).$$

For example, it is easy to show that a matrix

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

is positive definite iff $ac - b^2 > 0$ and $a, c > 0$.

So, $\mathbf{SPD}(2)$ can be viewed as a certain open region in $\mathbb{R}^3$. It is a convex cone, so (convex) interpolation and the notion of mean make sense. The space $\mathbf{SPD}(n)$ is also a convex cone.

Actually, this may not be the best way to represent $\mathbf{SPD}(n)$. It turns out that $\mathbf{SPD}(n)$ is the *homogeneous space* $\mathbf{SL}(n)/\mathbf{SO}(n)$ (in fact, a *symmetric space*). It is a Riemannian manifold of nonpositive sectional curvature.

Since $\mathbf{SPD}(3)$ is convex, the mean of m matrices in $\mathbf{SPD}(3)$ belongs to $\mathbf{SPD}(3)$.

However, the mean of matrices in $\mathbf{SPD}(3)$ with the same determinant can be a matrix with a greater determinant, which is undesirable. Worse, PCA is invalid because it does not preserve positive-definiteness.

Some of these problems can be alleviated using the exponential map and its inverse (log). Indeed, if $\mathbf{S}(n)$ denotes the vector space of all symmetric $n \times n$ matrices, then the *exponential map*,

$$\exp\colon \mathbf{S}(n) \to \mathbf{SPD}(n),$$

is a bijection!

This fact is the basis of the approach of Arsigny, Fillard, Pennec and Ayache.

An other example is given by the group of rotations, $\mathbf{SO}(3)$. These are the 3×3 (orthogonal) matrices, R, such that $RR^\top = R^\top R = I$ and $\det(R) = 1$.

This is a curved space and if $R_1, R_2 \in \mathbf{SO}(3)$, in general, $(1-\lambda)R_1 + \lambda R_2$ is *not* a rotation matrix!

It is also possible to interpolate, again using the exponential map,

$$\exp\colon \mathfrak{so}(3) \to \mathbf{SO}(3),$$

where $\mathfrak{so}(3)$ is the vector space of 3×3 skew-symmetric matrices. Another interpolation method uses the quaternions.

However, computing the mean is a harder problem.

All this raises a number of questions:

(1) What is a "good" notion of space, M, so that

(a) At least, "locally," such a space "looks" Euclidean.

(b) We can *promote calculus* on $\mathbb{R}^n$ to functions, $f\colon M \to \mathbb{R}$. In particular, we can

(c) define *smooth* functions on M.

(d) define the derivative (differential) of functions on M.

(e) solve differential equations (ODE's) on M.

(f) integrate (nice) functions on M.

(g) We can define the distance between two points.

(h) We can find *shortest paths* (perhaps only under certain conditions).

(2) What does it mean for a space to be *curved*? How do we define the notion of *curvature*?

(3) What is the effect of curvature (whatever it is)?

(4) What is the *Laplacian* on a manifold? For example, what is the Laplacian on the sphere $S^n \subseteq \mathbb{R}^{n+1}$, with

$$S^n = \{(x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_1^2 + \cdots + x_{n+1}^2 = 1\}.$$

What are the eigenfunctions of the Laplacian on the sphere?
(the nonzero functions, f, such that

$$\Delta f = \lambda f, \qquad \text{for some} \quad \lambda \in \mathbb{R})$$

They turn out to be the *spherical harmonics*.

The concept of a *manifold* seems to be a very good candidate for the notion of space we are seeking.

Another problem where 2-dimensional manifolds come up is the *reconstruction of a surface from a 3D mesh*.

Typically, the surface is *not* known ahead of time, so the standard definition of a manifold is useless.

What is needed is the notion of constructing a manifold given some set of *gluing data*.

This idea goes back to the development of the notion of algebraic variety in the 1940's (André Weil) and is a well-known technique for constructing fibre bundles.

This idea was first applied to computer science by Hughes and Grimm (1995) who introduced *proto-manifolds*.

The above methods were too restrictive and had some technical problems.

Marcelo Siqueira discovered much more general methods and Gallier, Siqueira and Xu rectified the technical problems and developed practical methods.

Proto-manifolds have been replaced by *pseudo-manifolds*.

A very useful thing about manifolds is that for every point, $p \in M$, there is a kind of "linear approximation," the *tangent space*, T_pM, to M at p. Near p, T_pM approximates M.

Given a smooth map, $f\colon M \to N$, for every $p \in M$, there is a *linear* map (the *tangent map*),

$$df_p\colon T_pM \to T_{f(p)}N,$$

which can be viewed as a "linear approximation" of f (near p).

Furthermore, if every tangent space, T_pM, has an *inner product* (a way to define orthogonality and distances), then M is called a *Riemannian manifold* and there is a map,

$$\exp\colon T_pM \to M,$$

defined at least near 0 (in T_pM).

In a Riemannian manifold, we can also define *covariant derivatives* and various notions of *curvature*.

When a manifold also has a group structure (so that multiplication and inversion are smooth), we get a very interesting structure called a *Lie group*.

Even if a manifold, M, is not a Lie group, there may be an *action*, $\cdot\colon G \times M \to M$, of a Lie group, G, on M and under certain conditions, M, can be viewed as a "quotient" G/K, where K is a subgroup of G.

When $M \cong G/K$, as above, certain notions on G can be transported to M. We say that M is a *homogeneous space*.

Let us now begin to familiarize ourselves with Lie groups and manifolds by looking at many concrete examples. We begin with groups of matrices (*matrix groups*).