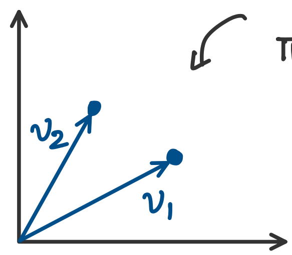
Riemannian Geometry

Monday, April 11, 2022 4:04 PM

Let us first start with the basics of a vector space V

as the name suggests its a space comprising of vectors v_i

For simplicity think of these vectors being embedded in a Euclidean space $\mathbb{R}^2$ (i.e. space of dimension 2)



This is $\mathbb{R}^2$ where we can draw two vectors $v_1, v_2 \in \mathbb{R}^2$

Fig: vector space $\mathbb{R}^2$

Vector space is defined using two key properties:

1) Any scalar multiple of a vector $v \in V$ is still a part of V

i.e. if $v \in V$ then $\alpha v \in V \quad \forall \alpha \in \mathbb{R}$

eg: Let $v = (1,2) \in \mathbb{R}^2$ and $\alpha = 3 \in \mathbb{R}$

then $\alpha v = (3,6)$ which is also a part of $\mathbb{R}^2$

2) if any two vectors $v_1, v_2 \in V$ then $\alpha_1 v_1 + \alpha_2 v_2 \in V$

eg: $v_1 = [1 \ 2] \quad v_2 = [2 \ 4]$

$\alpha_1 = 1, \ \alpha_2 = 3$

$$\Rightarrow \alpha_1 v_1 + \alpha_2 v_2 = [1 \ 2] + 3[2 \ 4]$$
$$= [1 \ 2] + [6 \ 12]$$
$$= [7 \ 14] \in \mathbb{R}^2$$

In machine learning, most of the data are assumed to be vectors often belonging to higher dimensions but its always a good idea to verify them!

One particularly useful result of considering vector space is that one can use an inner product structure to compare different vectors - this is key in pattern matching often done in machine learning

eg: inner product in Euclidean vector spaces.

Let $v_1, v_2 \in \mathbb{R}^d$ a vector spaces of d-dimensions, then an inner product

$$\langle v_1, v_2 \rangle = v_1^T v_2$$ where v_1, v_2 are represented as row vectors

(in ML, the dimensions d are generally referred to as features)

One can use the inner product to measure lengths of vectors in a vector space namely $\|v\| = \sqrt{\langle v, v \rangle}$ where $\|\cdot\|$ represents norm.

We can then make use of the norm to compute distances using:

$$d(v, w) = \|v - w\|$$ where $d(\cdot,\cdot)$ is a distance

(Note that we can compute $v - w = v + (-1)w$ using the vector space definition (2) and the norm as defined above)

eg: To illustrate that this is an idea familiar to you, lets consider an example in two dimensional Euclidean space $\mathbb{R}^2$

Lets denote $v = [v_1 \;\; v_2]$ $w = [w_1 \;\; w_2]$ be two vectors in $\mathbb{R}^2$

$$\Rightarrow v - w = [v_1 - w_1 \;\; v_2 - w_2]$$

$$\Rightarrow \|v - w\| = \left([v_1 - w_1 \;\; v_2 - w_2]^T [v_1 - w_1 \;\; v_2 - w_2] \right)^{1/2}$$

$$= \left(\begin{bmatrix} v_1 - w_1 \\ v_2 - w_2 \end{bmatrix} [v_1 - w_1 \;\; v_2 - w_2] \right)^{1/2}$$

$$= \left((v_1 - w_1)^2 + (v_2 - w_2)^2 \right)^{1/2}$$

does this remind you of a formula you know?

Now, lets revisit one of the properties of the vectorspace : property 2

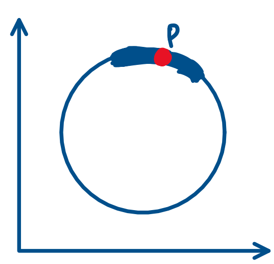
This property says the you could generate a new vector by linearly combining two vectors that belong to the vector space you're interested in

This assumption is too strong if we don't know that data we considered can be generated in the same way.

in the case that the data is not part of a linear space, we use the term manifold to imply that there's a inherent non-linear structure to the data

without making it extremly complex, a manifold can be thought of as a "space" where at any given point p, it looks like a Euclidean vector space but further away from it, its highly "curved".

To illustrate, consider a circle drawn in $\mathbb{R}^2$



- Circle is an example of a manifold
- around p (the thick blue region) the space is fairly "flat" thus like a vector space
- but far from p, we see that the space has curvature.

An alternate definition of a manifold that should be extremely useful for our purposes is : its a set of points that are constrained to follow a non-linear equation.

eg: The circle in $\mathbb{R}^2$ is defined as points (x, y) such that

$$x^2 + y^2 = 1$$

So, now how do we compare two points on a manifold? what's inner-product? what is a distance?

we will define these below using Riemannian geometry.

We have described earlier that inner products are useful to compare two vectors. For manifolds a definition of inner product comes from a tangent space.

Consider a point p on a manifold M. Suppose C be a curve passing through p.
we can then define velocity v of the curve C at point p as the time derivative.
This velocity vector can be represented in the local coordinates of the manifold at p in a Euclidean space.
(This is because manifold is locally flat and can be approximated as an Euclidean space)

Thus we have a space of velocity vectors for each point p on the manifold called Tangent Space $T_p(M)$

eg: Every Euclidean space is a manifold and we can derive its tangent spaces using the following procedure.

1) Start by representing a curve as a time snapshots of points in Euclidean this would be $\gamma(t) = vt + p$ at $p \in \mathbb{R}^d$

2) Take the time derivate of the curve at $t=0$

$$\dot{\gamma}(0) = v$$

3) since the curve γ we considered is in $\mathbb{R}^n$, the vector $v \in \mathbb{R}^n$ then $T_p(\mathbb{R}^n) = \mathbb{R}^n$

eg: Let's consider a slightly complex example in a circle embedded in $\mathbb{R}^2$
we again start with a curve γ on the circle (S^1) as a collection

we again start with a curve γ on the circle (S^1) as a collection of points

since the circle is in $\mathbb{R}^2$ we can parameterize the curve based on time (i.e. a path traveled between $t=0$ and $t=1$)

$$\Rightarrow \quad \gamma(t) = [\gamma_1(t) \quad \gamma_2(t)] \in \mathbb{R}^2$$

Now using the constraint : $\gamma_1^2(t) + \gamma_2^2(t) = 1$

taking the derivative : $\gamma_1(t)\dot{\gamma}_1(t) + \gamma_2(t)\dot{\gamma}_2(t) = 0$

writing it using inner product : $\langle \gamma(t), \dot{\gamma}(t) \rangle = 0$

The instantaneous velocity vector of $\gamma(t)$ is $v = \dot{\gamma}(t)$ is now also constrained by the above equation.

Namely the tangent space for a point on the circle is the orthogonal vector in $\mathbb{R}^2$

Hopefully, the above discussion has convinced you that, for a curved manifold, we can attach a tangent space that allow us to do vector calculus and distance computation.

The tangent space of a manifold allows us to define an inner product using Riemannian metric thus the name Riemannian geometry

The definition of Riemannian metric is as follows:

given a vector space V, we define a function $\phi : V \times V \to \mathbb{R}$ (meaning that it takes two vectors in V and spits out a real number)

Such that

1. $\phi(\alpha v_1 + \beta v_2, w) = \alpha\phi(v_1, w) + \beta\phi(v_2, w)$

2. $\phi(v, \alpha w_1 + \beta w_2) = \alpha\phi(v, w_1) + \beta\phi(v, w_2)$

The above formulation might look innocus but it simply suggest that the metric decomposes over the arguments to the function.

Two key results follow:

a) ϕ is symmetric i.e $\phi(v,w) = \phi(w,v)$

b) it becomes an inner product if it is positive definite

i.e. $\phi(v,v) \geq 0$ and

$\phi(v,v) = 0 \iff v = 0$

eg: Euclidean space $\mathbb{R}^n$ with tangent space $T_p(M) = \mathbb{R}^n$

the Riemannian metric is something we have seen before

$$\phi(v_1, v_2) = v_1^T v_2$$

eg: Consider the space $\mathbb{R}^2_+ = \{ p = (p_1, p_2) \in \mathbb{R}^2 \mid p_2 > 0 \}$ also called as the upper-half plane

This space has a very different Riemannian metric (called hyperbolic) and it changes from point to point.

The tangent space is again the entire $\mathbb{R}^2_+$ i.e. $T_p(M) = \mathbb{R}^2_+$

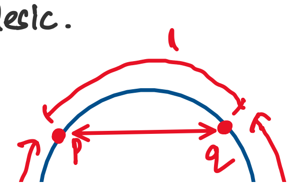
The Riemannian metric $\phi(v_1, v_2) = \frac{1}{p_2^2} v_1^T v_2$

contrast this to the Euclidean given above

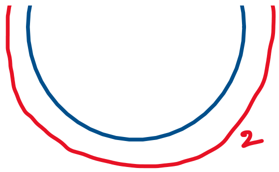
As we did before in vector spaces, we can use the inner product structure to define and compute a distance.

For manifolds, however, the distances are computed using length of a path along the manifold.

As you might imagine there could be multiple paths connecting two points but we are interested in the shortest one of those called a geodesic.



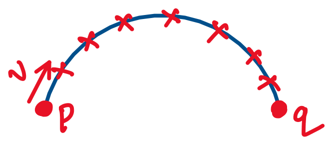
– There are two paths connecting points p, q along the manifold and the path 1 would



be a geodesic

- Note however, the line connecting p, q is not a geodesic although it appears to be computing a distance

let us know shift our focus onto how do we measure the length of the geodesic.



we can consider the path between two points to be composed of set of points (drawn as × above)

if we make the spacing between the points on the path infintesimally small, then we can approximate length of the infinitesimal vector using the norm on the tangent space

in the figure above, we can compute the spacing by measuring length of vector v

Now, the length of the path connecting p and q is simply a sum (integration) over all such infinitesimally small vectors.

To write this into a formula, consider a parameterized path $\gamma(t)$ as before.

Then we have, infinitesimal vector $v = \frac{d}{dt}\gamma(t)$

$$len(v) = \left[\phi\left(\frac{d\gamma}{dt}, \frac{d\gamma}{dt} \right)_t \right]^{1/2}$$

To compute the path length, we need to integrate $len(v)$ over the time

$$L(\gamma) = \int_{t=0}^{1} len(v)(t)\, dt$$

The geodesic is the path γ that minimizes the length $L(\gamma)$.

we define the geodesic distance

$$\boxed{d(p,q) = \min_{\gamma} L(\gamma) \qquad \begin{array}{l} \gamma(0) = p \\ \gamma(1) = q \end{array}}$$

As you can see from the above equation, the geodesic is an optimization problem!

But for particular manifolds of interest, we can obtain closed form solutions.

What to do next?

- Hopefully you got a taste of what does Riemannian geometry mean and why is it important to data analysis?

- you should try to play around with different manifolds in geomstats and visualize geodesics!

- Can you think of where can we use the geodesics we defined above?

  hint : Think of non-linear dimensionality reduction methods such as isomap