MAT327 Term Project
A novel method to derive Geodesics and Parallel Transport on Stiefel Manifold

Yongxin Xi
yxi@princeton.edu

Abstract

This project focuses on the general Stiefel manifold St(n,p) and proposes a method to derive formulas for the geodesics and parallel transport on this manifold. Examples are then illustrated.

1. Introduction

The motivation of this project comes from optimization problems on the Stiefel manifolds. Because studying and understanding geometric facts of this manifold, especially the geodesics and parallel transport, can help obtain a reasonable framework for such problems. As a Riemannian manifold, the Stiefel manifold can be embedded into the Euclidean space, such that the tangent space and the normal space at each point on the manifold are well defined. In order to obtain the simplest formulas, I use extrinsic coordinates instead of the coordinate chart to describe the manifold and the metric. Though unusual, it is computationally useful while still offers correct solution.

The outline of the project: First define the Stiefel manifold, and work out the tangent space and normal space at a point on the manifold. Then we use 1) the property that the acceleration vector of geodesic curve being in the normal space and 2) the local uniqueness theorem for geodesics, to derive the formula for geodesics. Then we take a close look at the parallel transport of tangent vector along a piecewise regular curve. A differential equation regarding the parallel transport is obtained yet no closed form solution provided.

At last, numerical examples are given to illustrate and justify these ideas.

2. Stiefel manifold

Definition: Stiefel manifold $St(n,p)$ is the set of ordered p-tuples orthonormal vectors in $\mathbf{R}^n$, i.e.

$$\forall Y \in St(n,p): \qquad Y^TY = I_p \tag{1}$$

In other words, $St(n,p)$ is a set of n-by-p (p less than or equal to n) matrices with orthonormality condition. It may be embedded in the np-dimensional Euclidean space of n-by-p matrices.

Dimension: From $Y^TY = I_p$ we have $\frac{p(p+1)}{2}$ nonlinear constraints, therefore the dimension of $St(n,p)$ is:

$$Dim(St(n,p)) = np - \frac{p(p+1)}{2} \tag{2}$$

This means that we could characterize the manifold by this number of coordinates. However, it would be very difficult to get a simple parametrization for $St(n,p)$.

3. Tangent and normal space

Let $Y(t)$ be a parametrized by arc length curve on the manifold.

3.1. Tangent space

Claim: At point $Y(0)$, vector $T(0)$ is a tangent vector if and only if

$$T(0)^TY(0) + Y(0)^TT(0) = 0 \tag{3}$$

So the tangent space $\boldsymbol{T}_{Y(0)}St(n,p)$ is well defined.

Proof:
Only if part is easy. Let $T(0)$ be the tangent vector of curve $Y(t)$ at point $Y(0)$, i.e. $T(0) = Y'(0)$. Since differentiating $Y(t)^TY(t) = I$ gives $Y'(t)^TY(t) + Y(t)^TY'(t) = 0$, at $t = 0$ we have $T(0)^TY(0) + Y(0)^TT(0) = 0$.
For the if part, if $T(0)^TY(0) + Y(0)^TT(0) = 0$ yet $T(0)$ is not a tangent vector at $Y(0)$, then it means this equation is not enough to define a tangent vector, in other words, there must be some other constraints that $T(0)$ have to satisfy. Since $T(0)^TY(0) + Y(0)^TT(0) = 0$ introduces $\frac{p(p+1)}{2}$ linear constraints on $T(0)$, then the total number of constraints is greater than $\frac{p(p+1)}{2}$, which says that the dimension of the tangent space is less than $np - \frac{p(p+1)}{2}$. However a common fact about d-dimensional manifolds is that the tangent plane at a point is a d-dimensional vector space with origin at the point of tangency. So this contradicts with (2).

3.2. Normal space

Definition: The normal space at point Y(0) is defined to be the orthogonal complement of the tangent space in the np-dimensional Euclidean space. Since orthogonality depends upon the definition of an inner product, we choose the standard inner product

$$g_e(\Delta_1, \Delta_2) = tr\Delta_1^T \Delta_2 \tag{4}$$

in the np-dimensional Euclidean space, which is also the Frobenius inner product for n-by-p matrices. So the normal space at the point Y(0) consists of all matrices N such that

$$tr\Delta^T N = 0, \forall \Delta \in T_{Y(0)} \tag{5}$$

Claim: The normal space at point $Y(0)$ is

$$\boldsymbol{N}_{Y(0)} = Y(0)S, S \in sym^{pxp} \tag{6}$$

where sym stands for symmetric matrices.
Proof:
First note that $Dim\boldsymbol{N}_{Y(0)} = np - Dim\boldsymbol{T}_{Y(0)} = np - (np - \frac{p(p+1)}{2}) = \frac{p(p+1)}{2}$.
$\forall T(0) \in \boldsymbol{T}_{Y(0)}$ and $\forall S \in sym^{pxp}$,

$$\begin{aligned} tr(T(0)^T Y(0)S) &= tr(-Y(0)^T T(0)S) \\ &= tr(-Y(0)^T T(0)S)^T \\ &= tr(-S(T(0)^T Y(0))) \\ &= tr(-T(0)^T Y(0)S) \end{aligned}$$

So, $tr(T(0)^T Y(0)S) = 0$. Since $S \in sym^{pxp}$ has dimension $DimS = \frac{p(p+1)}{2} = Dim\boldsymbol{N}_{Y(0)}$, so $\boldsymbol{N}_{Y(0)} = Y(0)S$, where $S \in sym^{pxp}$.

3.3. Projection

Now we are interested in knowing the projection of a vector $Z \in \mathbf{R^{nxp}}$ onto the tangent space and normal space of point Y. It can be shown that

$$\pi_{N_Y}(Z) = Ysym(Y^T Z) \tag{7}$$

$$\pi_{T_Y}(Z) = Yskew(Y^T Z) + (I - YY^T)Z \tag{8}$$

Where $sym(A) = \frac{1}{2}(A+A^T)$, and $skew(A) = \frac{1}{2}(A-A^T)$.
Proof: Let $\pi_{N_Y}(Z) = YS$, for some symmetric matrix S, and denote $\pi_{T_Y}(Z)$ as Δ. Then,

$$\begin{aligned} Z &= \Delta + YS \\ Z^T &= \Delta^T + SY^T \end{aligned}$$

Pre and post multiply Y^T and Y respectively, we then have

$$\begin{aligned} Y^T Z &= Y^T \Delta + S \\ Z^T Y &= \Delta^T Y + S \end{aligned}$$

Adding the two equations leads to $2S = Y^T Z + Z^T Y$, so $S = \frac{1}{2}(Y^T Z + Z^T Y) = sym(Y^T Z)$. So,

$$\begin{aligned} \pi_{T_Y}(Z) &= Z - \pi_{N_Y}(Z) \\ &= Z - Y * \frac{1}{2}(Y^T Z + Z^T Y) \\ &= Y * \frac{1}{2}(Y^T Z - Z^T Y) + (I - YY^T)Z \\ &= Yskew(Y^T Z) + (I - YY^T)Z \end{aligned}$$

4. Embedded geodesics

Definition: Let $Y(t)$ be a nonconstant parametrized curve on Stiefel manifold $St(n,p)$. Then $Y(t)$ is called a parametrized geodesic if $Y'(t)$ is a parallel vector field along $Y(t)$, i.e. the covariant derivative $\frac{DY'(t)}{dt} = 0$.

Intuitively, a geodesic is a generalization of a 'straight line' to the 'curved spaces'. It gives the shortest path on the manifold between two points. To construct a geodesic, therefore, requires that the acceleration vector $Y''(t)$ being in the normal space of $Y(t)$ at every point along the curve.

Claim: A geodesic Y(t) on Stiefel manifold satisfies the differential equation

$$\ddot{Y}(t) + Y(t)(\dot{Y}(t)^T \dot{Y}(t)) = 0 \tag{9}$$

Proof:

$$\begin{aligned} Y(t)^T Y(t) &= I \\ => Y'(t)^T Y(t) + Y(t)^T Y'(t) &= 0 \\ => Y''(t)^T Y(t) + 2Y'(t)^T Y'(t) + Y(t)^T Y''(t) &= 0 \end{aligned}$$

Y''(t) is in the normal space of Y(t), so from (6) let $Y''(t) = -Y(t)S(t)$, where $S(t) \in sym^{pxp}$. Then the equation becomes:

$$\begin{aligned} & Y^T Y(-S) + 2Y'^T Y' + (-YS)^T Y = 0 \\ & => -S + 2Y'^T Y' - S^T = 0 \\ & => S = Y'^T Y' \end{aligned}$$

$$\begin{aligned} Y''(t) &= -Y(t)S(t) \\ &= -Y(t)Y'(t)^T Y'(t) \end{aligned}$$

Therefore $\ddot{Y}(t) + Y(t)(\dot{Y}(t)^T \dot{Y}(t)) = 0$.

Claim: The geodesic Y(t) with the Y(0) and Y'(0) at t=0 is given by

$$Y(t) = (Y(0), Y'(0)) \exp t \begin{pmatrix} A & -S(0) \\ I & A \end{pmatrix} I_{2p,p} \exp(-At) \tag{10}$$

Where $I_{2p,p} = \begin{pmatrix} \mathbf{I_p} \\ \mathbf{0_p} \end{pmatrix}$, and $S(0) = Y'(0)^T Y'(0)$.
Proof: Let $A = Y^T \dot{Y}$, $S = \dot{Y}^T \dot{Y}$. Then

$$\begin{aligned} \dot{A} &= \dot{Y}^T \dot{Y} + Y^T \ddot{Y} \\ &= \dot{Y}^T \dot{Y} + Y^T (-Y \dot{Y}^T \dot{Y}) \\ &= \dot{Y}^T \dot{Y} - \dot{Y}^T \dot{Y} \\ &= 0 \end{aligned}$$

So $A(t) = A$ is a constant matrix.

$$\begin{aligned} \dot{S} &= \ddot{Y}^T \dot{Y} + \dot{Y}^T \ddot{Y} \\ &= (-Y \dot{Y}^T \dot{Y})^T \dot{Y} + \dot{Y}^T (-Y \dot{Y}^T \dot{Y}) \\ &= -\dot{Y}^T \dot{Y} Y^T \dot{Y} + Y^T \dot{Y} \dot{Y}^T \dot{Y} \\ &= -SA + AS \\ &= [A, S] \end{aligned}$$

Therefore, $\dot{Y}^T \dot{Y} = S(t) = e^{At} S(0) e^{-At}$.
Now, it is easy to check that the following equation holds:

$$\frac{d}{dt}(Y e^{At}, \dot{Y} e^{At}) = (Y e^{At}, \dot{Y} e^{At}) \begin{pmatrix} A & -S(0) \\ I & A \end{pmatrix} \quad (11)$$

This immediately lead to (10).
From the uniqueness theorem for geodesics, this formula will provide a unique geodesic on the Stiefel manifold given the initial point $Y(0)$ and its tangent vector $Y'(0)$.

5. Parallel transport

Definition: A vector field of an open set $U \subset St(n,p)$ of the Stiefel manifold is a correspondence w that assigns to each $Y \in U$ a vector $w(Y) \in T_Y St(n,p)$.
Definition: A vector filed w along a parametrized curve $Y(t) : I- > St(n,p)$ is said to be parallel if $\frac{Dw}{dt} = 0$ for every $t \in I$.
Property: Let $Y : I- > St(n,p)$ be a parametrized curve in $St(n,p)$, and let $w_0 \in T_{Y(0)} St(n,p)$. Then there exists a unique parallel vector field $w(t)$ along $Y(t)$, with $w(0) = w_0$. And $w(t)$ is called the parallel transport of w_0 along the curve at point $Y(t)$.

To derive the equation for the parallel transport $w(t)$ given curve $Y(t)$, $Y(0)$ and w_0, we look at the derivative of $w(t)$ to obtain a differential equation out of it. Since $\frac{Dw}{dt} = 0$, the derivative of $w(t)$ should be constrained in the normal space of $Y(t)$. So to transport the vector $w(t)$ along the curve to $w(t+\delta)$, where δ is infinitesimal, we can decompose $w(t)$ into the projections onto the tangent space and normal space of point $Y(t+\delta)$, and remove the normal part of it. In this way, $w(t+\delta)$ belongs to the tangent space of $Y(t+\delta)$, and $\frac{Dw}{dt} = 0$ is guaranteed. Following this idea,

$$\begin{aligned} w'(t) &= \lim_{\delta \to 0} \frac{w(t+\delta) - w(t)}{\delta} \\ &= \lim_{\delta \to 0} \frac{-\pi_{N_{Y(t)+\delta Y'(t))}}}{\delta} \\ &= \lim_{\delta \to 0} \frac{-(Y(t) + \delta Y'(t)) sym((Y(t) + \delta Y'(t))^T w(t))}{\delta} \\ &= -\lim_{\delta \to 0} \frac{(Y(t) + \delta Y'(t))(\delta Y'(t)^T w(t) + \delta w(t)^T Y'(t))}{2\delta} \\ &= -\frac{1}{2} Y(t)(Y'(t)^T w(t) + w(t)^T Y'(t)) \end{aligned}$$

Therefore,

$$w'(t) = -\frac{1}{2} Y(t)(Y'(t)^T w(t) + w(t)^T Y'(t)) \quad (12)$$

defines the parallel transport of w_0 from $Y(0)$ along $Y(t)$.

6. Examples

6.1. Sphere

Let us first consider a simple Stiefel manifold St(3,1), which is equivalent to a unit sphere. Given $Y(0) = (0,1,0)$, and $Y'(0) = (0,0,1)$, the geodesic is shown in Fig.1 as blue. The other three red circles are the geodesics given the same $Y(0)$ but different $Y'(0)$.

The parallel transport of a tangent vector on $St(3,1)$ can be obtained numerically from (12). For example, let $Y(t) = (\frac{\sqrt{3}}{2} sin(\sqrt{2/3}t), \frac{\sqrt{3}}{2} cos(\sqrt{2/3}t), \frac{1}{2})$, so $Y(0) = (0, \frac{\sqrt{3}}{2}, \frac{1}{2})$. Let $w_0 = (-1,0,0)$ be the initial tangent vector to be transported. The reason to choose this particular curve, is that the cone tangent to the sphere along $Y(t)$, when cut open, is isometric to a plane with an angle π in polar coordinates. It is then very easy to verify if the parallel transport derived is correct. Fig.2 shows the vector field $w(t)$.

6.2. St(5,2)

The strength of this novel method of obtaining geodesics and parallel transports is that it can be applied to any dimension of the Stiefel manifold. Now consider the $St(5,2)$, on which lie all the 5 by 2 orthonormal matrices. We can again obtain the geodesic from (10). It is very hard to plot the geodesic on $St(5,2)$, so instead 10 points on the geodesic with equal distance will be shown by matrices.

Similarly, we can work out the parallel transport by numerically integrating (12), and will be omitted here.