Geodesics on G and S manifolds (cont)

We now consider more about the geodesics of Grassmann and Stiefel manifolds.

We Know that each

$$Y(t) = Q e^{Xt} I_{n,p} K$$

where $Q \in SO(n)$, $X = \begin{bmatrix} A & -B^T \\ B & 0 \end{bmatrix}$, $I_{n,p} = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \\ & 0 & \end{bmatrix}$

is a Stiefel geodesic inside $Mat_{n\times p}(\mathbb{R})$.

But we would like a formula in terms of $Y(0) = Y$ and $Y'(0) = H$.

We can compute

$$Y'(t) = Q e^{Xt} X I_{n,p} K$$

so

$$Y'(0) = Q X I_{n,p} K$$

or

$$H = \begin{bmatrix} Y & \sim \end{bmatrix} \begin{bmatrix} A & -B^T \\ B & 0 \end{bmatrix} I_{n,p} K$$

Multiplying by Q^T on the left, we see

$$\begin{bmatrix} \longleftarrow \Psi^T \longrightarrow \\ \sim \end{bmatrix} \begin{bmatrix} \uparrow \\ H \\ \downarrow \end{bmatrix} = \begin{bmatrix} A \\ B \end{bmatrix}$$

This lets us compute $\Psi^T H = A$.

We can now observe that since

$$H^T H = -I_{n,k}^T \begin{bmatrix} A & -B^T \\ B & 0 \end{bmatrix} \cancel{Q^T Q} \begin{bmatrix} A & -B^T \\ B & 0 \end{bmatrix} I_{n,k}$$

$$= -I_{n,k}^T \begin{bmatrix} AA - B^T B & \sim \\ \sim & \sim \end{bmatrix} I_{n,k}$$

recall A is skew-symmetric!

$$= A^T A + B^T B,$$

we have a way to solve for $B^T B$ as

$$B^T B = H^T H - A^T A = H^T H - H^T \Psi \Psi^T H$$

$$= H^T (I - \Psi \Psi^T) H.$$

We then can define $C := H^T(I-\Psi\Psi^T)H$.

Now we prove

Theorem. If $\Psi(t) = Q e^{t\begin{bmatrix} A & -B^T \\ B & 0 \end{bmatrix}} I_{n,p}$ with $\Psi(0) = \Psi$ and $\dot{\Psi}(0) = H$, then

$$\Psi(t) = \Psi M(t) + (I - \Psi\Psi^T) H \int_0^t M(t)\, dt$$

where $M(t)$ is the solution to the second order constant coefficient ODE

$$M''(t) - A M'(t) + C M = 0, \quad M(0) = I_p, \quad M'(0) = A$$

and $A = \Psi^T H$ is skew-symmetric, while $C = H^T(I - \Psi\Psi^T)H$ is non-negative definite.

Proof. We start by observing that $M(t) = I_{n,k}^T e^{t\begin{bmatrix} A & -B^T \\ B & 0 \end{bmatrix}} I_{n,k}$. To prove this, we just differentiate:

We'll need a little improvement to a previous lemma:

Lemma. $\frac{d}{dt} e^{tA} = A e^{tA} = e^{tA} A$.

Now we can check our matrices $M(t)$ satisfy the claimed ODE.

We compute

$$M(t) = I_K \; I_{n,k}^T \, e^{t\begin{bmatrix} A & -B^T \\ B & 0 \end{bmatrix}} I_{n,k}$$

$$M'(t) = A \; I_{n,k}^T \, e^{t\begin{bmatrix} A & -B^T \\ B & 0 \end{bmatrix}} I_{n,k}$$

$$M''(t) = \left(-A^TA - B^TB\right) I_{n,k}^T \, e^{t\begin{bmatrix} A & -B \\ B & 0 \end{bmatrix}} I_{n,k}.$$

Thus, if we let $I_{n,k}^T \, e^{t\begin{bmatrix} A & -B \\ B & 0 \end{bmatrix}} I_{n,k} = M(t)$ again, we see that

$$M''(t) - AM'(t) + CM(t)$$

$$= \left[(-A^TA - B^TB) - AA + C \right] M(t)$$

$$= \left[-B^TB + C \right] M(t) = 0, \text{ as claimed.}$$

Farther, plugging in $t=0$ yields

$$M(0) = I_k, \quad M'(0) = A.$$

By uniqueness of solutions to ODEs, we now know that

$$M(t) = I_{n,k}^T \, e^{t\begin{bmatrix} A & -B^T \\ B & 0 \end{bmatrix}} I_{n,k},$$

as desired.

~~We now consider some geometry. Suppose that~~

$$\cancel{Y = [\ ]_{n\times k}}$$ ~~= orthonormal basis for k-dim subspace S~~.

~~Then for any n×k matrix H,~~

$$\cancel{Y^T H = [\ ]_{k\times k}}$$

~~the expression of $\pi_S H$ in the column coordinates of Y.~~

Now we introduce a neat trick.
Suppose we have an $n \times K$ "orthogonal" matrix Ψ, so that $\Psi^T \Psi = I_K$.
Consider the $n \times n$ matrix $\Psi\Psi^T$. This is called the <u>projector</u> onto the column space of Ψ, since it orthogonally projects vectors ~~onto this~~ in $\mathbb{R}^n$ onto this space.

To see this, note that for an $n \times 1$ vector $\vec{v}$,

$$\Psi^T \vec{v} = \text{dot products of } \vec{v} \text{ w/cols of } \Psi$$

$$\Psi\Psi^T \vec{v} = \text{linear combo. of those cols. according to dot products.}$$

Of course, this means that

$$I - \Psi\Psi^T = \text{projector onto } (\text{col space } \Psi)^{\perp}.$$

~~Notice that for any $n \times p$ matrices A and B, this means~~

These two projectors also have the ~~am~~ useful properties:

$$(\Psi\Psi^T)^2 = \Psi\Psi^T$$

$$(I - \Psi\Psi^T)^2 = (I - \Psi\Psi^T)$$

$$\Psi\Psi^T(I - \Psi\Psi^T) = (I - \Psi\Psi^T)\Psi\Psi^T = 0.$$

We now turn back to verifying

$$\Psi(t) = \Psi M(t) + (I - \Psi\Psi^T) H \int_0^t M(t)\,dt.$$

Multiplying on the left by $\Psi\Psi^T$, we see that we must show

$$\Psi\Psi^T\,\Psi(t) = \Psi\Psi^T\Psi\,M(t)$$
$$= \Psi M(t)$$

Now we know $\Psi(t) = Q e^{tX} I_{n,k}$ and $M(t) = I_{n,k}^T e^{tX} I_{n,k}$, so we can write this as

$$(\Psi\Psi^T Q)\, e^{tX} I_{n,k} = (\Psi I_{n,k}^T)\, e^{tX} I_{n,k}$$

But

$$\underset{k\times n}{\Psi^T}\underset{n\times n}{Q} = \Psi^T \begin{bmatrix} \uparrow & \uparrow \\ \Psi & \text{stuff in } \Psi^\perp \\ \downarrow & \downarrow \end{bmatrix} = \begin{bmatrix} I_k & 0 \end{bmatrix} = I_{n,k}^T,$$

as desired.

Multiplying on the left by $I - \Psi\Psi^T$, we get to show

$$(I-\Psi\Psi^T)\,\Psi(t) = (I-\Psi\Psi^T)\,H\int_0^t M(t)\,dt$$

Now $\Psi(t) = Q e^{tX}$, where $Q = \begin{bmatrix} \Psi & \text{stuff orthogonal to } \Psi \end{bmatrix}$, so let's call the right-hand $n\times(n-k)$ block Z. We see that

$$(I-\Psi\Psi^T)\begin{bmatrix} \Psi & Z \end{bmatrix} = \begin{bmatrix} 0 & Z \end{bmatrix},$$

so the lhs obeys

$$(I-\Psi\Psi^T)\,\Psi(t) = \begin{bmatrix} 0 & Z \end{bmatrix} e^{tX} I_{n,k}.$$

We can understand the term $\int_0^t M(t)\,dt$ by integrating the defining differential equation for $M(t)$.

Now we know

$$(I-\Psi\Psi^T)\Psi(0) = (I-\Psi\Psi^T)\Psi = 0$$

and

$$(I-\Psi\Psi^T)H\int_0^0 M(t)\,dt = 0.$$

So we most only show

$$\frac{d}{dt}(I-\Psi\Psi^T)\Psi(t) \stackrel{?}{=} \frac{d}{dt}(I-\Psi\Psi^T)H\int_0^t M(t)\,dt,$$

or

$$(I-\Psi\Psi^T)\Psi'(t) \stackrel{?}{=} (I-\Psi\Psi^T)H\,M(t)$$

Now $\Psi'(t) = QXe^{Xt}I_{n,k}$, so we want

$$(I-\Psi\Psi^T)QXe^{Xt}I_{n,k} \stackrel{?}{=} (I-\Psi\Psi^T)H\,I_{n,k}^T e^{Xt}I_{n,k}.$$

It suffices to show

$$(I-\Psi\Psi^T)QX \stackrel{?}{=} (I-\Psi\Psi^T)H\,I_{n,k}^T$$

~~actually, we need to show this since e^{Xt} is orthogonal and $I_{n,k}$ weakly orthogonal.~~

Now

$$\Psi'(0) = QXI_{n,k} = H$$

so we can substitute this in on the right, leaving us to prove

$$(I-\Psi\Psi^T)QX \stackrel{?}{=} (I-\Psi\Psi^T)QXI_{n,k}I_{n,k}^T.$$

Now the right hand operator $I_{n,k}I_{n,k}^T$ is projection onto the first k columns.

Further, we previously saw

$$(I-\Psi\Psi^T)Q = \begin{array}{cc} \left[\begin{matrix} 0 & Z \end{matrix}\right] \\ \begin{matrix} \scriptstyle k & \scriptstyle n-k \end{matrix} \end{array}.$$

Multiplying by $X = \begin{array}{c} \begin{bmatrix} A & -B^T \\ B & 0 \end{bmatrix} \\ \begin{matrix} \scriptstyle k & \scriptstyle n-k \end{matrix} \end{array}$, we get

$$(I-\Psi\Psi^T)QX = [0\ Z]\begin{bmatrix} A & -B^T \\ B & 0 \end{bmatrix} = \begin{array}{c} \left[\begin{matrix} ZB & 0 \end{matrix}\right] \\ \begin{matrix} \scriptstyle k & \scriptstyle n-k \end{matrix} \end{array},$$

which is clearly invariant under projection to the first k colums. $\square$

(whew!).

That was lengthy (and, frankly, ugly), but in the end we got a prize.

Now how do we solve <u>matrix</u> ODEs like

$$M''(t) - AM'(t) + CM = 0.$$

The solution turns out to be given by solving the "quadratic eigenvalue problem"

$$(\lambda^2 I - A\lambda + C)x = 0$$

Such problems (for $K \times K$ matrices) generally have $2K$ eigenvalues and $2K$ eigenvectors. If $\Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_{2K} \end{bmatrix}$, then if $X = [\quad]_{K \times 2K}$ the matrix of eigenvectors, and Z is a $2K \times K$ matrix chosen so that $XZ = I$, $X\Lambda Z = A$, then

$$M(t) = X e^{\Lambda t} Z.$$

We can check this directly:

$$M'(t) = X\Lambda e^{\Lambda t} Z$$

$$M''(t) = X\Lambda^2 e^{\Lambda t} Z$$

so

$$M''(t) - AM'(t) + CM(t) =$$

$$\left(X\Lambda^2 - AX\Lambda + CX\right) e^{\Lambda t} Z$$

But column-by-column, ~~the~~ ~~Z~~ X must satisfy this matrix equation by construction.