An alternate form for Stiefel Geodesics.
Grassmann Geodesics.

Edelman, Arias, and Smith give the following consequence of the last form:

Theorem. Suppose that Y and H are $n \times k$ matrices so that $Y^T Y = I_k$ and $Y^T H = A$ is skew-symmetric.

Let $K = (I - YY^T)H$, (the normal-to-colspace(Y) component of H) and $QR = K$ be the QR decomposition of K (Q is $n \times k$, R is $k \times k$).

Let $M(t)$ and $N(t)$ be given by

$$\underbrace{\begin{pmatrix} M(t) \\ N(t) \end{pmatrix}}_{2k \times k} = e^{t \overbrace{\begin{bmatrix} A & -R^T \\ R & 0 \end{bmatrix}}^{2k \times 2k}} \cdot \underbrace{\begin{bmatrix} I_k \\ 0 \end{bmatrix}}_{2k \times k}$$

$\begin{bmatrix} I_k \\ 0 \end{bmatrix}$ " projection onto 1st k columns.

Then the Stiefel geodesic from Y in direction H is given by

$$Y(t) = Y M(t) + Q N(t).$$

<u>This</u> is the formula that we'd like to use for computations! To compare Grassmann geodesics, first note that

$$V_k(\mathbb{R}^n) \longrightarrow G_k(\mathbb{R}^n)$$

is also a Riemmannian submersion. We will use this to work out the tangent space to $G_k(\mathbb{R}^n)$.

Now the tangent space to $\mathrm{Mat}_{n\times k}(\mathbb{R})$ can be orthogonally decomposed into

$$M = \underbrace{(YY^T)M}_{\substack{\text{vectors in} \\ \text{colspace } Y}} + \underbrace{(I - YY^T)M}_{\substack{\text{vectors in} \\ (\text{colspace } Y)^\perp}}$$

We know that

$$T_Y V_k(\mathbb{R}^n) = \left\{ M \,\middle|\, Y^T M = A,\ A \text{ skew-symmetric} \right\}.$$

Further, we know that inside $V_k(\mathbb{R}^n) \subset Mat_{n \times k}$, elements of $G_k(\mathbb{R}^n)$ are represented by equivalence classes

$$[Y] = \left\{ Y Q_k \,\middle|\, Q_k \in O(k) \right\},$$

so the vertical space is

$$Vert_Y = \left\{ YA \,\middle|\, A \text{ is skew-symmetric} \right\}.$$

The horizontal space is then given in the following form. Suppose $Y_\perp$ is an $n \times (n-k)$ matrix so that $(Y\ Y_\perp) \in O(n)$. (That is, any basis for $\text{colspace}(Y)^\perp$.)

Then we have

$$\mathrm{Horiz}_Y = \{ Y_\perp B \mid B \text{ is } n\text{-}k \times k \}.$$

This tells us how to see the tangent space to the Grassmannian: it's <u>this</u> $k(n-k)$ dimensional space, or equivalently the space of $n \times k$ vectors Δ so that

$$Y^T \Delta = 0$$

This means they don't rotate the colums of Y within colspace Y.

~~Proposition. The invariant metric on $G_k(\mathbb{R}^n)$ is the same as a constant multiple of the subspace quotient space metric on $G_k(\mathbb{R}^n)$.~~

What's the metric on this space?

For this ~~proposition~~ question to make sense, we have to recall our constructions:

$$G_k(\mathbb{R}^n) = O(n) / O(k) \times O(n-k)$$

$$[Q] = \left\{ Q \begin{pmatrix} Q_k & 0 \\ 0 & Q_{n-k} \end{pmatrix} \middle| \begin{array}{c} Q \in O(n),\ Q_k \in O(k), \\ Q_{n-k} \in O(n-k) \end{array} \right\}$$

$$\mathrm{Horiz}_Q = T_{[Q]} G_k(\mathbb{R}^n) = \left\{ Q \begin{pmatrix} 0 & -B^T \\ B & 0 \end{pmatrix} \middle| B \text{ is } (n-k) \times k \right\}.$$

We can take various metrics on this tangent space:

1) Restriction of $O(n)$ metric (in $n \times n$ representation)

$$g_c(\Delta_1, \Delta_2) = \frac{1}{2} \operatorname{tr} \left(Q \begin{pmatrix} 0 & -B_1^T \\ B_1 & 0 \end{pmatrix} \right)^T \left(Q \begin{pmatrix} 0 & -B_2^T \\ B_2^T & 0 \end{pmatrix} \right)$$

$$= \frac{1}{2} \operatorname{tr} \left(\begin{pmatrix} 0 & B_1^T \\ -B_1 & 0 \end{pmatrix} Q^T Q \begin{pmatrix} 0 & -B_2^T \\ B_2 & 0 \end{pmatrix} \right)$$

$$= \frac{1}{2} \operatorname{tr} \begin{pmatrix} B_1^T B_2 & 0 \\ 0 & B_2^T B_1 \end{pmatrix} = \operatorname{tr} B_1^T B_2$$

(using $\operatorname{tr} AB = \operatorname{tr} BA$)

2) Restriction of Stiefel metric (in $n \times k$ representation).

Using the orthogonal decomposition of tangents to $Mat_{n\times k}$ as

$$\Delta = YA + Y_\perp B,$$

we see that

$$g_{\text{Stiefel}}(\Delta_1, \Delta_2) = \tfrac{1}{2}\operatorname{tr} A_1^T A_2 + \operatorname{tr} B_1^T B_2$$

$$= \operatorname{tr}\left(\Delta_1^T (I - \tfrac{1}{2} Y Y^T)\Delta_2\right)$$

$$= \operatorname{tr}\left((A_1^T Y^T + B_1^T Y_\perp^T)(YA_2 + Y_\perp B_2 - \tfrac{1}{2} YA_2)\right)$$

$$= \tfrac{1}{2}\operatorname{tr} A_1^T A_2 + \operatorname{tr} B_1^T B_2 .$$

So if $\Delta_1, \Delta_2 \in \text{Horiz}_Y$, or $Y^T\Delta_i = 0$, we have

$$g_{\text{Stiefel}}(\Delta_1, \Delta_2) = \operatorname{tr} B_1^T B_2 .$$

3) Restriction of Euclidean metric on $Mat_{n\times k}$.

In fact, if Δ_1, Δ_2 are actually $Y_\perp B_1$, $Y_\perp B_2$, then

$$
\begin{aligned}
g_e(\Delta_1, \Delta_2) &= \operatorname{tr} \Delta_1^T \Delta_2 \\
&= \operatorname{tr}\left(B_1^T Y_\perp^T Y_\perp B_2 \right) \\
&= \operatorname{tr} B_1^T B_2 .
\end{aligned}
$$

So what about the Grassmann <u>geodesics</u>? First, recall ~~that~~ the

~~Theorem~~

~~Singular Value Decomposition Theorem.~~
~~Any $n\times K$ matrix H can be written~~

$$\cancel{H = U \Sigma V^T}$$

~~where $U \in O(n)$, $V \in O(K)$ and $\Sigma = ($~~

Compact Singular Value Decomposition.

Any $n \times K$ matrix H can be written

$$H = U \Sigma V^T$$

where U is $n \times K$, $U^T U = I_K$, Σ is $K \times K$ and diagonal, and V is $K \times K$ orthogonal.

Now we have

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

$$\Psi(t) = \underset{n \times 2K}{\begin{pmatrix} \Psi V & U \end{pmatrix}} \underset{2K \times K}{\begin{pmatrix} \cos \Sigma t \\ \sin \Sigma t \end{pmatrix}} \underset{K \times K}{V^T}.$$

This gives us an "$n \times K$" characterization of geodesics on the Grassmannian.

The proof is an exercise.

Let's look at a consequence:
the (constant) speed of the geodesic $\Psi(t)$ is given by $\|H\|_F$, which is

$$
\begin{aligned}
\|H\|_F^2 &= \operatorname{tr}(H^T H) \\
&= \operatorname{tr}\left((U\Sigma V)^T (U\Sigma V)\right) \\
&= \operatorname{tr}(V^T \Sigma^T U^T U \Sigma V) \\
&= \operatorname{tr}(\Sigma \Sigma^T) = \sum \sigma_i^2
\end{aligned}
$$

where the σ_i are the singular values of H.
Thus the distance between $\Psi(0)$ and $\Psi(t)$ is just

$$d(\Psi(0), \Psi(t)) = t\left(\sum \sigma_i^2\right)^{1/2}.$$