Grassmann Distances and Angles

We now give a more general understanding of distances between subspaces. We start with

Theorem. (CS Decomposition Theorem)

For any $n \times n$ unitary matrix $Q$ and any $2 \times 2$ partitioning $r_1 + r_2 = n = c_1 + c_2$, there exist unitary matrices $U_1, U_2$ ($r_1 \times c_1$ and $r_2 \times c_2$) and $V_1, V_2$ ($c_1 \times c_1$ and $c_2 \times c_2$) so that

\[
U^* Q V = \begin{bmatrix}
U_1^* & 0 \\
0 & U_2^*
\end{bmatrix}
\begin{bmatrix}
Q_{11} & Q_{12} \\
Q_{21} & Q_{22}
\end{bmatrix}
\begin{bmatrix}
V_1^* & 0 \\
0 & V_2^*
\end{bmatrix}
= \begin{bmatrix}
D_{11} & D_{12} \\
D_{21} & D_{22}
\end{bmatrix}
\]

where each $D_{ij}$ is real and almost diagonal (each row and column contains at most one nonzero entry).
The structure of the $D_{ij}$ is given by

$$
D = [D_1 \  D_2] = \begin{bmatrix}
I & C & O_s^* \\
O_s & 0 & S \\
O_s & I & I \\
S & I & -C & O_c^*
\end{bmatrix}
$$

where $C = \begin{bmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{bmatrix}$ in decreasing order,

$S = \begin{bmatrix} \sigma_1 & \sigma_2 & 0 \\ 0 & \sigma_3 & \sigma_4 \end{bmatrix}$ in increasing order, and $C^2 + S^2 = I$.

Depending on the partition, the matrices of zeros $O_s$, and $O_c$ are rectangular and/or absent. The $C$ and $S$ blocks have the same dimensions, but the dimensions depend on $Q$ and could be zero.

This is a complicated looking beast, but we will need to apply it in only a special case.
Suppose we have subspaces of \( \mathbb{C}^n \) given by \( E_2 = [E_1 | E_2] \) and \( F = [F_1 | F_2] \), where these are any unitary matrices so that \( \text{colspan}(E_1) \) and \( \text{colspan}(F_1) \) give the two subspaces (and \( \text{colspan}(E_2), \text{colspan}(F_2) \) their orthogonal complements).

We note that the CSD is not unique, as we can permute rows or columns within the blocks. It turns out that the most useful one is the \( D \) which is (Frobenius) closest to \( I_n \).

Corollary. When \( r_1 = c_1, r_2 = c_2 \), we have the special CS decomposition

\[
D = c_2 \begin{bmatrix}
  c_1 & -S \\
  S & c_2
\end{bmatrix} \\
  \begin{bmatrix}
  C & -S \\
  S & C \\
  \end{bmatrix} \\
  \begin{bmatrix}
  c_2 \\
  I
\end{bmatrix}
\]
Among all CSDs of \( Q_x = [Q_{11} Q_{21}] \), \( \hat{D} \) is (operator norm) closest to the identity matrix. 

We now show how to use \( \hat{D} \) to construct a matrix mapping \( \text{colspace}(E_1) \) into \( \text{colspace}(F_1) \). This \( \hat{D} \) is called the direct rotation, and this \( \hat{D} \) is the core of the direct rotation.

Let \( Q = E^*F \), and 

\[
\begin{bmatrix}
    \mathcal{U}^* & 
    \begin{bmatrix}
        E_1 & \mathcal{V}_1 \\
        0 & \mathcal{V}_2
    \end{bmatrix} & \mathcal{V}_2
\end{bmatrix}
\]

\( U^*QV = \hat{D} \) be the "Minimal" CS decomposition of \( Q \) above. Let the direct rotation 

\[
W = EU^*DU^*E^*
\]

We claim \( W \) maps \( \text{colspace}(E_1) \) into \( \text{colspace}(F_1) \).
Consider

\[ W E_L U_L = E \hat{U} \hat{D} u^* E^* E_L U_L \]

\[ \downarrow_{n \times k} \rightarrow_{k \times k} \]

\[ = E \hat{U} \hat{D} (u^*_k \begin{bmatrix} I_k \\ 0 \end{bmatrix} u_L) \]

\[ = E \hat{U} \hat{D} (\begin{bmatrix} u^*_1 & \ldots & u^*_k \end{bmatrix} \begin{bmatrix} I_k \\ 0 \end{bmatrix} u_L) \]

\[ = E \hat{U} \hat{D} \begin{bmatrix} u^*_1 & \ldots & u^*_k \end{bmatrix} \begin{bmatrix} u_L \\ 0 \end{bmatrix} \]

\[ = E \hat{U} \hat{D} \begin{bmatrix} I_k \\ 0 \end{bmatrix} \]

Now on the other hand, we have

\[ u^* E^* F V = \hat{D}, \]

so

\[ F V = E \hat{U} \hat{D} \]
So we have

\[ WE_1 U_1 = FV[I_k] \]

\[ = F \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} [I_k] \]

\[ = F \begin{bmatrix} V_1 \\ 0 \end{bmatrix} = F_1 V_1, \]

and \( W \) takes the basis \( E_1 U_1 \) for \( \text{colspace}(E_1) \) to the basis \( F_1 V_1 \) for \( \text{colspace}(F_1) \).

In fact, the angles \( \Theta_i \) of which \( C \) and \( S \) are the cosines and sines are the angles between these basis vectors for \( \text{colspace}(E_1) \) and \( \text{colspace}(E_2) \), and these are called "the angles between \( E_1 \) and \( \text{colspace} F_1 \)." - Jordan
Proposition. The direct rotation $W$ is unitary. It is the unitary matrix for which $\|W-I\|$ is smallest among unitary matrices mapping $\text{colspace}(E_\perp)$ to $\text{colspace}(F_\perp)$.

Proof. First, we check

$$W^*W = EU^\dagger U^*E^*EU^\dagger U^*E^*$$

$$= EU \begin{bmatrix} c^*s & -s \\ -s & c \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} U^*E^*$$

$$= EU \begin{bmatrix} c^2s^2 & 0 & -s^2 \\ 0 & c^2s^2 & 0 \\ -s^2 & 0 & c^2s^2 \end{bmatrix} U^*E^* = I.$$  

Next, we are using the operator norm, so

$$W-I = EU^\dagger U^*E^*-I$$

$$= EU(\hat{D}-I)U^*E^*$$

implies $\|W-I\| = \|\hat{D}-I\|$. 


But \( \| D - I \| \) is the least among all CSDs of \( E^*F = Q \), and (retracing our steps) we can eventually work out that any unitary \( W' \) with the desired properties comes from some CSD of \( Q \). \( \square \).

This is all from Paige and Saunders, who cite Davis and Kahan. An equivalent, but more geometric proof is given by Edelman et al., who show:

**Proposition.**

If \( W(t) = EU \)

\[
\begin{bmatrix}
\cos \theta_1 & \cos \theta_k \\
\sin \theta_1 & \sin \theta_k \\
0 & 0
\end{bmatrix}
\]

then the first \( k \) columns of \( W \) follow a geodesic in \( \text{Gr}_k(\mathbb{C}^n) \) from colspace \( E \) to colspace \( F \) as \( t \) goes from 0 to 1,

assuming \( U^*QV = \hat{D} \) and \( Q = E^*F \) as usual.
Proof. We start with a useful fact (from EAS, 2.7). The geodesic equation on the Grassmann manifold is

\[ W''(t) + W(t) (w'(t) \times w'(t)) = 0. \]

Plugging into our formula for \( W(t) \) in and crunching out the derivatives gives the result, as \( W(0) = E_{1} U_{1} \) and \( W(1) = E \mathcal{D} \mathcal{D}^* [I_{k}] = F_{2} V_{2} \) are the correct endpoints. □.

We conclude with a nice result of James and Wilkinson.

Proposition. If \( A \) and \( B \) are projectors onto \( k \)-dimensional subspaces of \( \mathbb{C}^n \) then the eigenvalues of \( ABA \) and \( BAB \) are equal to \( \lambda_i = \cos^2 \theta_i \) where the \( \theta_i \) are the principal angles between subspaces.