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 2×2 partioning $r_1 + r_2 = n = c_1 + c_2$, there exist unitary matrices U_1, U_2 ($r_1 \times r_1$ and $r_2 \times r_2$) 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

$$D = \begin{bmatrix} D_1 & D_2 \end{bmatrix} = \left[\begin{array}{ccc|ccc} I & & & O_s^* & & \\ & C & & & S & \\ & & O_c & & & I \\ \hline O_s & & & I & & \\ & S & & & -C & \\ & & I & & & O_c^* \end{array}\right]$$

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

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

~~The~~ Depending on the partition, the matrices of zeros O_c and O_s are rectangular and/or absent. ~~and~~ 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 only in a special case.

Suppose we have subspaces of $\mathbb{C}^n$ given by $E = \left[E_1 \middle| E_2 \right]$ and $F = \left[F_1 \middle| F_2 \right]$, where these are any unitary matrices so that colspace(E_1) and colspace(F_1) give the two subspaces (and colspace(E_2), colspace(F_2) their orthogonal complements).

move to 4

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.

Corollary. When $r_1 = c_1$, $r_2 = c_2$, we have the <u>special</u> CS decomposition

$$\hat{D} = \begin{array}{c c} & \begin{array}{cc} c_1 & c_2 \end{array} \\ \begin{array}{c} r_1 \\ r_2 \\ \\ \end{array} & \left[\begin{array}{c|cc} C & -S & \\ \hline S & C & \\ & & I \end{array} \right] \end{array}$$

Among all CSDs of $Q = \begin{bmatrix} Q_{11} & Q_{21} \\ Q_{12} & Q_{22} \end{bmatrix}$,

$\hat{D}$ is (operator norm) closest to the identity matrix.

→ moved to here.

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

Let $Q = E^*F$, and

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

$$W = EU\hat{D}U^*E^*$$

We claim W maps colspace(E_1) into colspace(F_1).

Consider

$$\underset{n\times k}{W E_1}\,\underset{k\times k}{U_1} = E U \hat{D} U^* E^* E_1 U_1$$

$$= E U \hat{D}\left(U^* \begin{bmatrix} I_k \\ 0 \end{bmatrix} U_1\right)$$

$$= E U \hat{D}\left(\begin{bmatrix} U_1^* & \\ & U_2^* \end{bmatrix}\begin{bmatrix} I_k \\ 0 \end{bmatrix} U_1\right)$$

$$= E U \hat{D}\left(\begin{bmatrix} U_1^* & \\ & U_2^* \end{bmatrix}\begin{bmatrix} U_1 \\ 0 \end{bmatrix}\right)$$

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

$$\cancel{= E U \begin{bmatrix} C \\ S \\ 0 \end{bmatrix}}$$

Now on the other hand, we have

$$U^* E^* F V = \hat{D},$$

so

$$F V = E U \hat{D}$$

So we have

$$WE_1 U_1 = FV\begin{bmatrix} I_k \\ 0 \end{bmatrix}$$

$$= F\begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix}\begin{bmatrix} I_k \\ 0 \end{bmatrix}$$

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

and W takes the basis $E_1 U_1$ for colspace(E_1) to the basis $F_1 V_1$ for colspace(F_1).

In fact, the angles θ_i of which C and S are the cosines and sines are the angles between these basis ~~elements~~ vectors for colspace(E_1) and colspace(E_2), and these are called "the Jordan angles between colspace E_1 and colspace F_1."

Proposition. The direct rotation W is unitary. It is the unitary matrix for which $\|W-I\|$ is smallest among unitary matrices mapping colspace(E_1) to colspace(F_1).

Proof. First, we check

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

Next, we are using the operator norm, so

$$
\begin{aligned}
W-I &= EU\hat{D}U^*E^*-I \\
&= EU(\hat{D}-I)U^*E^*
\end{aligned}
$$

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

But $\|\hat{D} - I\|$ 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 $W(t) = EU \begin{bmatrix} \cos t\theta_1 & & \\ & \ddots & \\ & & \cos t\theta_k \\ \sin t\theta_1 & & \\ & \ddots & \\ & & \sin t\theta_k \\ & 0 & \end{bmatrix}$,

then the ~~span of the~~ colspace ~~first k columns of~~ W follow a geodesic in $Gr_k(\mathbb{C}^n)$ from colspace E to colspace F as t goes from 0 to $\underline{1}$, assuming $U^* Q V = \hat{D}$ and $Q = E^* F$ as usual.

Proof. We start with a useful fact (from EAS, 2.7). The $n \times k$ geodesic equation on the Grassmann manifold is

$$W''(t) + W(t)\Big(W'(t)^{*}\,W'(t)\Big) = 0.$$

Plugging ~~this into~~ our formula for $W(t)$ in and crunching out the derivatives gives the result, as $W(0) = E_1 \mathcal{U}_1$, and $W(1) = E\mathcal{U}\hat{D} \begin{bmatrix} I_k \\ 0 \end{bmatrix} = F_1 V_1$ are the correct endpoints. $\square$.

We conclude with a nice ~~lemma~~ result of James and Wilkinson.

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