Variational characterization of distance between subspaces

GZ, Xuefei

February 16, 2017

For S_1, S_2, two linear subspaces of $\mathbb{R}^m$ (of dimensions k_1, k_2), with projection matrices P_{S_1}, P_{S_2}, we define the distances between S_1 and S_2 as

$$\mathrm{d}(S_1, S_2) = ||P_{S_1} - P_{S_2}||_2$$

Suppose matrices $W_1 \in \mathbb{R}^{m\times k_1}$ and $Z_1 \in \mathbb{R}^{m\times k_2}$ with orthogonal columns span S_1 and S_2 respectively. If we augment both matrices to form orthonomal basis of $\mathbb{R}^m$: $W = [W_1, W_2]$, $Z = [Z_1, Z_2]$, we have ourselves the following

Lemma:

$$\mathrm{d}(S_1, S_2) = ||W_1' Z_2||_2 = ||Z_1' W_2||_2 = \sqrt{1 - \sigma^2_{min}(W_1' Z_1)}$$

which can be taken as an operational definition of $\mathrm{d}(S_1, S_2)$.

We show here an equivalent variational characterization of $\mathrm{d}(S_1, S_2)$, which can be used to prove the above lemma.

$$\begin{aligned}
\sup_{\substack{x\in S_1^\perp, ||x||=1 \\ y\in S_2, ||y||=1}} < x, y >^2 &= \sup_{||a||=1, ||b||=1} < W_2 a, Z_1 b >^2 \\
&= \sup_{||a||=1} \sup_{||b||=1} ||a' W_2' Z_1 b||_2^2 \\
&= \sup_{||a||=1} ||a' W_2' Z_1||_2^2 \\
&= ||W_2' Z_1||_2^2 \qquad (1)
\end{aligned}$$

This is, if we take a unit vector form $S_1^\perp$ and another unit vector from S_2, (square-root of the cosine of) the smallest angle possible can be taken as the distance between the two subspaces; the dimensions of the two subspaces do not matter.

Similarly,

$$\begin{aligned}
\sigma^2_{min}(W_1' Z_1) &= \inf_{||b||=1} ||W_1' Z_1 b||_2^2 \\
&= \inf_{||b||=1} \left[\sup_{||a||=1} ||a' W_1' Z_1 b||_2^2 \right] \\
&= \inf_{\substack{y\in S_2 \\ ||y||=1}} \sup_{\substack{x\in S_1 \\ ||x||=1}} < x, y >^2 \qquad (2)
\end{aligned}$$

Calculating $\sigma^2_{min}(Z_1' W_1)$ the same way and we can flip the role of x and y.

An equivalent definition of $\mathrm{d}(S_1, S_2)$ can therefore be taken as

$$\mathrm{d}(S_1, S_2) = \sup_{\substack{x\in S_1^\perp, ||x||=1 \\ y\in S_2, ||y||=1}} | < x, y > | = \sqrt{1 - \inf_{\substack{y\in S_2 \\ ||y||=1}} \sup_{\substack{x\in S_1 \\ ||x||=1}} < x, y >^2}$$

Next we give an alternative proof of the last equality in the **lemma**.

First we show that for an arbitrary vector $y \in \mathbb{R}^m$ fixed,

Claim 1

$$\inf_{\substack{x \in S_1 \\ ||x||=1}} (1- < x, y >^2) = \sup_{\substack{z \in S_1^\perp \\ ||z||=1}} < z, y >^2 \qquad (3)$$

This claim is saying that the minimized angle with respect to y from S_1 and $S_1^\perp$ are complementary.

Actually more is true:

Claim 2

There exists a pair of optimizers x^*, z^* of the two sides of (3) that spans a plane containing y, i.e., x^*, z^*, y, and the origin 0 are coplanar.

The proof is geometric

Proof of claim 2:

Suppose z^* attains maximum of the RHS of (3), define

$$x^* = -\frac{< z^*, y >}{\sqrt{1- < z^*, y >^2}} z^* + \frac{1}{\sqrt{1- < z^*, y >^2}} y$$

then $x^* \perp z^*$, and $< x^*, y >^2 = 1- < z^*, y >^2$.

Furtunately $x^* \in S_1$: if we write

$$\begin{aligned}
x^* &= P_{S_1^\perp} x^* + P_{S_1} x^* \\
&= P_{z^*} x^* + (P_{S_1^\perp} x^* - P_{z^*} x^*) + P_{S_1} x^* \\
&= \Big(-\frac{< z^*, y >}{\sqrt{1- < z^*, y >^2}} + \frac{1}{\sqrt{1- < z^*, y >^2}} < z^*, y > \Big) z^* + (P_{S_1^\perp} x^* - P_{z^*} x^*) + P_{S_1} x^* \\
&= (P_{S_1^\perp} - P_{z^*}) x^* + P_{S_1} x^* \\
&= (P_{S_1^\perp} - P_{z^*}) y + P_{S_1} x^*
\end{aligned}$$

Suppose $(P_{S_1^\perp} - P_{z^*}) y \neq 0$, we can write

$$y = \alpha_1 z^* + \alpha_2 z_2 + \beta x \quad \text{where} \;\; \alpha_2 \neq 0, \; z^* \perp z_2 \in S_1^\perp, \; x \in S_1$$

Then

$$< z_*, y >^2 = \alpha_1^2$$

but

$$< \frac{\alpha_1 z^* + \alpha_2 z_2}{||\alpha_1 z^* + \alpha_2 z_2||}, y >^2 = \frac{(\alpha_1^2 + \alpha_2^2)^2}{||\alpha_1 z^* + \alpha_2 z_2||^2} = \alpha_1^2 + \alpha_2^2 > \alpha_1^2$$

contradicting the assumption that z^* maximizes the RHS.

Hence $(P_{S_1^\perp} - P_{z^*}) y = 0$ and we know that $x^* = P_{S_1} x^*$, i.e., $x^* \in S_1$.

Therefore

$$\begin{aligned}
\sup_{\substack{x \in S_1 \\ ||z||=1}} < x, y >^2 &\geq < x^*, y >^2 \\
&= 1 - < z^*, y >^2 \\
&= 1 - \sup_{\substack{z \in S_1^\perp \\ ||z||=1}} < z, y >^2 \\
&= \inf_{\substack{z \in S_1^\perp \\ ||z||=1}} 1 - < z, y >^2
\end{aligned}$$

By **claim 1**, equalities hold throughout, claim 2 is shown.

Proof of claim 1:

$$\begin{aligned}
\inf_{\substack{x \in S_1 \\ ||x||=1}} (1 - < x, y >^2) &= \inf_{||a||=1} [1 - (y' Z_1 a)^2] \\
&= 1 - ||y' Z_1||_2^2 \\
&= ||y' Z_2||_2^2 \\
&= \sup_{||b||=1} (y' Z_2 b)^2 \\
&= \sup_{\substack{z \in S_1^\perp \\ ||z||=1}} < z, y >^2
\end{aligned}$$

where the third line uses the fact that $||y'Z||^2 = ||y'Z_1, y'Z_2||^2$, y is unit vector, and that Z is orthonormal. Claim 1 is shown.

Taking supremum on both sides over $y \in S_2$,

$$\begin{aligned}
||W_2' Z_1||_2^2 &= \sup_{\substack{y \in S_2 \\ |||y||=1}} \sup_{\substack{x \in S_1^\perp \\ ||x||=1}} < x, y >^2 && \text{by (1)} \\
&= \sup_{\substack{y \in S_2 \\ |||y||=1}} \inf_{\substack{x \in S_1 \\ ||x||=1}} (1 - < x, y >^2) && \text{by (3)} \\
&= 1 + \sup_{\substack{y \in S_2 \\ |||y||=1}} (- \sup_{\substack{x \in S_1 \\ ||x||=1}} < x, y >^2) \\
&= 1 - \inf_{\substack{y \in S_2 \\ |||y||=1}} \sup_{\substack{x \in S_1 \\ ||x||=1}} < x, y >^2 \\
&= 1 - \sigma_{min}^2 (W_1' Z_1) && \text{by (2)}
\end{aligned}$$

Lemma is proven.

.

.

.