The principal angles and the gap

Taslaman, Leo

2014

MIMS EPrint: **2014.9**

ISSN 1749-9097

THE PRINCIPAL ANGLES AND THE GAP

LEO TASLAMAN

ABSTRACT. In this note we provide proofs for some known results on the principal angles and the gap between two subspaces of $\mathbb{C}^n$. Both the principal angles and the gap are introduced with respect to an arbitrary positive definite inner product. We show that the principal angles between two subspaces $\mathscr{U}$ and $\mathscr{V}$ are unique and prove that the largest one, $\theta_{\max}$, satisfies

$$\theta_{\max} = \max_{\substack{u\in\mathscr{U}\\ \|u\|=1}} \min_{\substack{v\in\mathscr{V}\\ \|v\|=1}} \measuredangle(u,v) \quad \text{and} \quad \sin\theta_{\max} = \operatorname{gap}(\mathscr{U},\mathscr{V})$$

when $\dim\mathscr{U} = \dim\mathscr{V}$.

Keywords: principal angles, canonical angles, gap, canonical correlations.

1. CHARACTERIZATIONS OF THE SINGULAR VALUES

The following two results are needed for the next section.

Theorem 1.1. *The singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p$ of a matrix A can be characterized recursively as follows:*

$$\begin{aligned} \sigma_i &= \max\{\, |x^*Ay| \,:\, \|x\|_2 = \|y\|_2 = 1, \\ &\qquad x^*x_j = y^*y_j = 0,\ j = 1,2,\ldots,i-1\} \\ &= |x_i^*Ax_i|, \end{aligned}$$

where x_i and y_i are maximizing vectors (in fact singular vectors).

Proof. We have $|x^*Ay| \leq \|x\|_2\|A\|_2\|y\|_2 = \sigma_1$ and $x_1^*Ay_1 = \sigma_1$ where x_1 and y_1 can be any left and right first singular vectors, respectively. Hence the result is true for $i = 1$. Set $B = \sum_{k=i}^{p} \sigma_k x_k y_k^*$. For any vectors x and y in the ith set above, we have $|x^*Ay| = |x^*By| \leq \|x\|_2\|B\|_2\|y\|_2 = \sigma_i$. Since $x_i^*Ay_i = \sigma_i$ for any ith left and right singular vectors x_i and y_i, the result follows. $\square$

SCHOOL OF MATHEMATICS, THE UNIVERSITY OF MANCHESTER, MANCHESTER, M13 9PL, UK

E-mail address: `leo.taslaman@manchester.co.uk`.

Date: February 28, 2014.

Theorem 1.2. *The smallest singular value of a $p\times p$ matrix $A=U\Sigma V^*$, is given by*

$$\sigma_p = \min_{\|x\|_2=1}\max_{\|y\|_2=1}|x^*Ay|.$$

Proof. Let f denote the right hand side above. We have

$$f \le \max_{\|y\|_2=1}|x_p^*Ay| = \sigma_p,$$

for any pth left singular vector x_p. For any x we can pick y such that $\|y\|_2=1$ and $V^*y=U^*x$. Hence

$$f \ge \min_{\|x\|_2=1}|(U^*x)^*\Sigma U^*x| = \min_{\|w\|_2=1}|w^*\Sigma w| = \sigma_p.$$

□

2. The principal angles

For the remainder of this note, let $\langle\cdot,\cdot\rangle$ denote an arbitrary positive definite inner product on $\mathbb{C}^n$ and $\|\cdot\|$ the induced norm. Suppose, without loss of generality, that the inner product is defined by a symmetric positive definite matrix M, so $\langle u,v\rangle = u^*Mv$, for any u and v.

We define the angle between two nonzero vectors u and v as

$$\measuredangle(u,v) = \arccos\left(\frac{|\langle u,v\rangle|}{\|u\|\,\|v\|}\right),$$

and the angle between a nonzero vector u and a nonzero subspace $\mathscr{V}$ as

$$\measuredangle(u,\mathscr{V}) = \min_{\substack{v\in\mathscr{V}\\ \|v\|=1}}\measuredangle(u,v).$$

Now, consider two subspaces $\mathscr{U},\mathscr{V}\subseteq\mathbb{C}^n$. If $p=\dim\mathscr{U}\le\dim\mathscr{V}=q$, then there are p principal angles (or canonical angles)

$$0 \le \theta_1(\mathscr{U},\mathscr{V}) \le \theta_2(\mathscr{U},\mathscr{V}) \le \cdots \le \theta_p(\mathscr{U},\mathscr{V}) \le \pi/2$$

between $\mathscr{U}$ and $\mathscr{V}$. We shall with $\theta_{\max}(\mathscr{U},\mathscr{V})$ refer to $\theta_p(\mathscr{U},\mathscr{V})$. The principal angles are defined recursively by

$$\begin{aligned}\theta_i(\mathscr{U},\mathscr{V}) &= \min\{\measuredangle(u,v) : u\in\mathscr{U},\ v\in\mathscr{V}, \|u\|=\|v\|=1,\\ &\qquad \langle u,u_j\rangle = \langle v,v_j\rangle = 0,\ j=1,2,\ldots,i-1\}\\ &= \measuredangle(u_i,v_i),\end{aligned}$$

where u_i and v_i are minimizing vectors, known as principal vectors. Note that $u_1,u_2,\ldots,u_p$ and $v_1,v_2,\ldots,v_p$ are M-orthonormal bases for $\mathscr{U}$ and $\mathscr{V}$ respectively. The principal vectors are not unique but we shall see that principal angles are. We note that

$$\theta_i(\mathscr{U},\mathscr{V}) = \theta_i(\mathscr{V},\mathscr{U}) \tag{1}$$

for any i.

The next result is due to Björck and Golub [1].

Theorem 2.1. *Let* $U = [u_1, u_2, \dots, u_p]$ *and* $V = [v_1, v_2, \dots, v_p]$. *If* $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p$ *are the singular values of* U^*MV, *then*

$$\theta_i(\mathscr{U}, \mathscr{V}) = \arccos(\sigma_i).$$

We will make use of the following fact in the proof.

Fact 2.2. *Let* $f : X \to \mathbb{R}$ *such that* $f[X]$ *is a closed interval of the real line. For any decreasing function* $g : f[X] \to \mathbb{R}$ *it holds that*

$$g\left(\max_{x \in X} f(x)\right) = \min_{x \in X} g(f(x)) \quad \text{and} \quad g\left(\min_{x \in X} f(x)\right) = \max_{x \in X} g(f(x)).$$

Proof of Theorem 2.1. Cosine is a decreasing function on $[0, \pi/2]$. Hence, by using Fact 2.2, we see that taking cosine on both sides of the definition of $\theta_i(\mathscr{U}, \mathscr{V})$ yields

$$\begin{aligned}
\cos(\theta_i(\mathscr{U}, \mathscr{V})) &= \max\{\, |\langle u, v\rangle| \,:\, u \in \mathscr{U},\, v \in \mathscr{V}, \|u\| = \|v\| = 1, \\
&\qquad \langle u, u_j\rangle = \langle v, v_j\rangle = 0,\ j = 1, 2, \dots, i-1\} \\
&= |\langle u_i, v_i\rangle|.
\end{aligned}$$

By defining $u = Ux$, $v = Vy$ and $u_i = Ux_i$, $v_i = Vy_i$ we get

$$\begin{aligned}
\cos(\theta_i(\mathscr{U}, \mathscr{V})) &= \max\{\, |x^*(U^*MV)y| \,:\, \|x\|_2 = \|y\|_2 = 1, \\
&\qquad x^* x_i = y^* y_i = 0,\ j = 1, 2, \dots, i-1\} \\
&= |x_i^*(U^*MV)y_i|.
\end{aligned}$$

The result now follows from Theorem 1.1. □

To see that the principal angles are unique, consider any two matrices, U' and V', whose columns are M-orthonormal bases for $\mathscr{U}$ and $\mathscr{V}$, respectively. Then $U' = UQ_1$ and $V' = VQ_2$ where $Q_1^* Q_1 = I_p$ and $Q_2^* Q_2 = I_q$. Thus $U'^* MV' = Q_1^*(U^*MV)Q_2$ and the singular values are identical to those of U^*MV. This shows that principal angles are independent of choice of principal vectors, and the uniqueness follows from the uniqueness of the singular values.

The next result is stated without proof in [2, p. 249].

Corollary 2.3. *If* $p = q$, *then the largest principal angle is given by*

$$\theta_{\max}(\mathscr{U}, \mathscr{V}) = \max_{\substack{u \in \mathscr{U} \\ \|u\| = 1}} \min_{\substack{v \in \mathscr{V} \\ \|v\| = 1}} \measuredangle(u, v).$$

Proof. Using Theorem 1.2 we get

$$\begin{aligned}\cos(\theta_{\max}(\mathcal{U},\mathcal{V})) &= \sigma_p(U^*MV) = \min_{\|x\|_2=1}\max_{\|y\|_2=1}|x^*U^*MVy| \\ &= \min_{\substack{u\in\mathcal{U}\\ \|u\|=1}}\max_{\substack{v\in\mathcal{V}\\ \|v\|=1}}|u^*Mv| = \min_{\substack{u\in\mathcal{U}\\ \|u\|=1}}\max_{\substack{v\in\mathcal{V}\\ \|v\|=1}}|\langle u,v\rangle|.\end{aligned}$$

Since arccos is a decreasing function on $[0,1]$, Fact 2.2 implies that

$$\begin{aligned}\theta_{\max}(\mathcal{U},\mathcal{V}) &= \arccos\left(\min_{\substack{u\in\mathcal{U}\\ \|u\|=1}}\max_{\substack{v\in\mathcal{V}\\ \|v\|=1}}|\langle u,v\rangle|\right) = \max_{\substack{u\in\mathcal{U}\\ \|u\|=1}}\arccos\left(\max_{\substack{v\in\mathcal{V}\\ \|v\|=1}}|\langle u,v\rangle|\right) \\ &= \max_{\substack{u\in\mathcal{U}\\ \|u\|=1}}\min_{\substack{v\in\mathcal{V}\\ \|v\|=1}}\arccos(|\langle u,v\rangle|) = \max_{\substack{u\in\mathcal{U}\\ \|u\|=1}}\min_{\substack{v\in\mathcal{V}\\ \|v\|=1}}\measuredangle(u,v).\end{aligned}$$

□

We will use the next lemma to relate the largest principal angle to the gap. It is a special case of Theorem 2.2 in [2].

Lemma 2.4. *Let the columns of V and $V_\perp$ be M-orthonormal bases of $\mathcal{V}$ and $\mathcal{V}_\perp$ respectively. If $\|u\|=1$, then $\sin\measuredangle(u,\mathcal{V}) = \|V_\perp^* Mu\|_2$.*

Proof. Define

$$\begin{bmatrix} V^* \\ V_\perp^* \end{bmatrix} Mu = \begin{bmatrix} C \\ S \end{bmatrix}$$

Since VV^*M and $V_\perp V_\perp^* M$ are M-orthogonal projectors onto $\mathcal{V}$ and $\mathcal{V}_\perp$, respectively, we have

$$C^*C + S^*S = u^*M(VV^*M + V_\perp V_\perp^* M)u = u^*Mu = 1.$$

Now, C^*C and S^*S are scalars and equals the squares of 2-norms of V^*Mv and $V_\perp^*Mv$, respectively. By Theorem 2.1, $\cos^2\theta = C^*C$, where θ is the principal angle between $\operatorname{span}\{u\}$ and $\mathcal{V}$. It follows that $\sin\theta = (S^*S)^{1/2}$. □

3. The gap

In this section we introduce the gap between subspaces in $\mathbb{C}^n$ and relate it to the largest principal angle. Recall that all norms are with respect to the M-inner product. The following definitions can be found in [3, p. 7 and p. 197]:

$$\operatorname{dist}(u,\mathcal{V}) = \min_{v\in\mathcal{V}}\|u-v\|,$$

$$\delta(\mathcal{U},\mathcal{V}) = \begin{cases} 0 & \text{if } \mathcal{U}=0, \\ \max\limits_{\substack{u\in\mathcal{U}\\ \|u\|=1}}\operatorname{dist}(u,\mathcal{V}) & \text{otherwise,}\end{cases}$$

and

$$\operatorname{gap}(\mathcal{U},\mathcal{V}) = \max(\delta(\mathcal{U},\mathcal{V}),\delta(\mathcal{V},\mathcal{U})).$$

Note that we in general have $\delta(\mathcal{U},\mathcal{V}) \neq \delta(\mathcal{V},\mathcal{U})$, so gap $\not\equiv \delta$. If $\dim\mathcal{U} = \dim\mathcal{V}$, however, we will see that it always holds that $\operatorname{gap}(\mathcal{U},\mathcal{V}) = \delta(\mathcal{U},\mathcal{V})$.

The next lemma can be found in [2, Theorem 2.3].

Lemma 3.1. *If* $\|u\| = 1$*, then*

$$\sin\measuredangle(u,\mathcal{V}) = \min_{v\in\mathcal{V}} \|u - v\| = \operatorname{dist}(u,\mathcal{V}).$$

Proof. Let V and $V_\perp$ be as in Lemma 2.4. Write $V^*Mu = \widehat{u}$ and $V_\perp^*Mu = \widehat{u}_\perp$, and note that $V^*MVx = x$ and $V_\perp^*MVx = 0$. We have

$$\begin{bmatrix} V^* \\ V_\perp^* \end{bmatrix} M(u - Vx) = \begin{bmatrix} \widehat{u} - x \\ \widehat{u}_\perp \end{bmatrix}.$$

Further,

$$\|u - Vx\| = \left\| \begin{bmatrix} V^* \\ V_\perp^* \end{bmatrix} M(u - Vx) \right\|_2 = \left\| \begin{bmatrix} \widehat{u} - x \\ \widehat{u}_\perp \end{bmatrix} \right\|_2$$

which is minimized for $x = \widehat{u}$, with minimum $\|\widehat{u}_\perp\|_2 = \|V_\perp^*Mu\|_2$. By Lemma 2.4, $\|V_\perp^*Mu\|_2 = \sin\measuredangle(u,\mathcal{V})$. □

Now, if $\dim\mathcal{U} = \dim\mathcal{V} > 0$, then we have

$$\begin{aligned}
\max_{\substack{u\in\mathcal{U}\\ \|u\|=1}} \operatorname{dist}(u,\mathcal{V}) &= \max_{\substack{u\in\mathcal{U}\\ \|u\|=1}} \sin\measuredangle(u,\mathcal{V}) \\
&= \sin \max_{\substack{u\in\mathcal{U}\\ \|u\|=1}} \measuredangle(u,\mathcal{V}) \\
&= \sin \max_{\substack{u\in\mathcal{U}\\ \|u\|=1}} \min_{\substack{v\in\mathcal{V}\\ \|v\|=1}} \measuredangle(u,v) \\
&= \sin\theta_{\max}(\mathcal{U},\mathcal{V}),
\end{aligned}$$

where the first equality follows from Lemma 3.1; the second from the fact that sine is an increasing function on $[0,\pi/2]$; the third from the definition of the angle between a vector and a subspace; and the fourth from Corollary 2.3. Since $\theta_{\max}(\mathcal{U},\mathcal{V}) = \theta_{\max}(\mathcal{V},\mathcal{U})$, we have proved the following theorem.

Theorem 3.2. *If* $\dim\mathcal{U} = \dim\mathcal{V}$*, then*

$$\operatorname{gap}(\mathcal{U},\mathcal{V}) = \delta(\mathcal{U},\mathcal{V}) = \sin\theta_{\max}(\mathcal{U},\mathcal{V}).$$

References

[1] Å. Björck and G. H. Golub. Numerical Methods for Computing Angles Between Linear Subspaces, Mathematics of Computation, Vol. 27, No. 123 (Jul., 1973), pp. 579-594.

[2] G. W. Stewart. *Matrix Algorithms Volume II: Eigensystems.* SIAM, Philadelphia, PA, USA, 2001.

[3] T. Kato. *Pertubation Theorey for Linear Operators.* Springer-Verlag, Berlin Heidelberg, Germany, 1995. Reprint of the 1980 Edition.