Journal of
Mathematical
Inequalities

Volume 15, Number 1 (2021), 143–150 doi:10.7153/jmi-2021-15-13

TWO LOWER BOUNDS ABOUT SINGULAR SUBSPACES

CHUNGUANG REN

(*Communicated by M. Niezgoda*)

Abstract. Cai and Zhang establish two lower bounds for $\sin\Theta$ distances with spectral and Frobenius norms (Cai, T. T. and Zhang, A., Rate-optimal perturbation bounds for singular subspaces with applications to high-dimensional statistics, The Annals of Statistics, Vol. 46, No. 1 (2018) 60–89). We provide two lower bounds under any unitarily invariant norm. It turns out that our estimation is better in some sense.

1. Introduction

We begin with some notations in this section. For $a,b \in \mathbb{R}$, $a \wedge b := \min(a,b)$. Let $\mathbb{O}_{p,r} = \{V \in \mathbb{C}^{p\times r} : V^H V = I_r\}$ be the set of all $p \times r$ orthonormal columns and write $\mathbb{O}_p$ for the set of p-dimensional unitary matrices. Here, V^H denotes the conjugate transpose of V.

For $X, Z \in \mathbb{R}^{p_1\times p_2}$, the matrix X denotes a true and unobserved matrix, and Z is a small perturbation matrix, so $\hat{X} := X + Z$ can represent an observed matrix. Assume rank $(X) \geqslant r$, and there exists a significant gap between $\sigma_r(X)$ and $\sigma_{r+1}(X)$. Furthermore, the SVD of X can be given as follows,

$$X = [U\ U_\perp] \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma_\perp \end{bmatrix} [V\ V_\perp]^T, \tag{1}$$

where $U \in \mathbb{O}_{p_1,r}$ with $[U\ U_\perp] \in \mathbb{O}_{p_1}$; $V \in \mathbb{O}_{p_2,r}$ with $[V\ V_\perp] \in \mathbb{O}_{p_2}$; $\Sigma = \mathrm{diag}\,\{\sigma_1(X), \cdots, \sigma_r(X)\} \in \mathbb{R}^{r\times r}$ with the singular values $\sigma_1(X) \geqslant \cdots \geqslant \sigma_r(X)$ and $\Sigma_\perp \in \mathbb{R}^{(p_1-r)\times(p_2-r)}$.

Clearly, the SVD of $\hat{X}$ can be also given as follows,

$$\hat{X} = [\hat{U}\ \hat{U}_\perp] \begin{bmatrix} \hat{\Sigma} & 0 \\ 0 & \hat{\Sigma}_\perp \end{bmatrix} [\hat{V}\ \hat{V}_\perp]^T, \tag{2}$$

while the quantities $\hat{U}$, $\hat{U}_\perp$, $\hat{\Sigma}$, $\hat{\Sigma}_\perp$, $\hat{V}$ and $\hat{V}_\perp$ are defined analogously.

For a matrix $A \in \mathbb{C}^{p_1\times p_2}$, we denote $\mathbb{P}_A \in \mathbb{C}^{p_1\times p_1}$ the orthogonal projection operator onto the column space of A. The perturbation Z can be decomposed into four blocks

$$Z = Z_{11} + Z_{12} + Z_{21} + Z_{22}, \tag{3}$$

Mathematics subject classification (2010): 15A18, 15A42, 34D15.

Keywords and phrases: Lower bounds, unitarily invariant norm, $\sin\Theta$ distance, singular value decomposition.

where

$$Z_{11} = \mathbb{P}_U Z \mathbb{P}_V,\ Z_{12} = \mathbb{P}_U Z \mathbb{P}_{V_\perp},\ Z_{21} = \mathbb{P}_{U_\perp} Z \mathbb{P}_V,\ Z_{22} = \mathbb{P}_{U_\perp} Z \mathbb{P}_{V_\perp}.$$

Two matrix norms will be used in the paper: $\|A\|_2 = \sigma_1(A)$ stands for the spectral norm; $\|A\|_F = \sqrt{\sum_{i=1}^{p_1\wedge p_2} \sigma_i^2(A)}$ the Frobenius norm.

To reflect the perturbation effective, we should select one kind of quantities. The $\sin\Theta$ distances are well-known and classical, so we select them in this paper. Suppose the singular values of $V^H\hat{V}$ are $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r \geqslant 0$. Then

$$\Theta(V,\ \hat{V}) = \mathrm{diag}\{\cos^{-1}(\sigma_1),\ \cos^{-1}(\sigma_2),\ \cdots,\ \cos^{-1}(\sigma_r)\}$$

is said to be principle angles. A quantitative measure of distance between the column spaces of V and $\hat{V}$ is $\|\sin\Theta(V,\hat{V})\|$ under any unitarily invariant norm. Clearly, the established spectral and Frobenius $\sin\Theta$ distances in Cai and Zhang ([1]) are two special kinds of unitarily invariant $\|\sin\Theta(V,\hat{V})\|$ distances.

To state Cai and Zhang's results, we introduce

$$\mathscr{F} := \mathscr{F}_{r,\alpha,\beta,z_{12},z_{21}} = \big\{(X,Z):\ \text{with } \hat{X},U,V,Z_{ij} \text{ in (1)–(3)},$$

$$\sigma_r(U^H\hat{X}V) \geqslant \alpha,\ \|U_\perp^H\hat{X}V_\perp\|_2 \leqslant \beta,\ \|Z_{12}\|_2 \leqslant z_{12},\ \|Z_{21}\|_2 \leqslant z_{21}\big\} \text{ and}$$

$$\mathscr{G} := \mathscr{G}_{\alpha,\beta,z_{12},z_{21},\tilde{z}_{12},\tilde{z}_{21}} = \big\{(X,Z) \in \mathscr{F} : \|Z_{12}\|_F \leqslant \tilde{z}_{12}, \|Z_{21}\|_F \leqslant \tilde{z}_{21}\big\}.$$

THEOREM 1. ([1]) *Let* $r \leqslant \frac{1}{2}(p_1 \wedge p_2)$ *and* $\tilde{V} \in \mathbb{O}_{p_2\times r}$ *be any estimator of* V *based on* $\hat{X}$.

(i). If $\alpha^2 > \beta^2 + z_{12}^2 + z_{21}^2$,

$$\inf_{\tilde{V}} \sup_{(X,Z)\in\mathscr{F}} \|\sin\Theta(V,\hat{V})\|_2 \geqslant \frac{1}{8\sqrt{10}} \left(\frac{\alpha z_{12} + \beta z_{21}}{\alpha^2 - \beta^2 - z_{12}^2 \wedge z_{21}^2} \wedge 1\right).$$

(ii). With $\alpha^2 > \beta^2 + z_{12}^2 + z_{21}^2$, $\tilde{z}_{12} \leqslant \sqrt{r} z_{12}$ *and* $\tilde{z}_{21} \leqslant \sqrt{r} z_{21}$,

$$\inf_{\tilde{V}} \sup_{(X,Z)\in\mathscr{G}} \|\sin\Theta(V,\hat{V})\|_F \geqslant \frac{1}{8\sqrt{10}} \left(\frac{\alpha \tilde{z}_{12} + \beta \tilde{z}_{21}}{\alpha^2 - \beta^2 - z_{12}^2 \wedge z_{21}^2} \wedge \sqrt{r}\right).$$

THEOREM 2. ([1]) *Let* $r \leqslant \frac{1}{2}(p_1 \wedge p_2)$ *and* $\tilde{V} \in \mathbb{O}_{p_2\times r}$ *be any estimator of* V *based on* $\hat{X}$. *If* $\alpha^2 \leqslant \beta^2 + z_{12}^2 \wedge z_{21}^2$,

$$\inf_{\tilde{V}} \sup_{(X,Z)\in\mathscr{F}} \|\sin\Theta(V,\hat{V})\|_2 \geqslant \frac{1}{2\sqrt{2}}.$$

DEFINITION 1. ([3]) A norm $\|\cdot\|$ is said to be unitarily invariant norm, if $\|UAV^H\| = \|A\|$ holds for each unitary U, V. It is normalized if $\|A\| = \|A\|_2$ holds for rank (A)=1.

Both spectral and Frobenius norms are examples of unitarily invariant norms. In the following discussions, we always use $\|\cdot\|$ to represent a unitarily invariant norm.

To establish our lower bounds, we introduce

$$\mathscr{H} := \mathscr{H}_{\alpha,\beta,z_{12},z_{21},\tilde{z}_{12},\tilde{z}_{21}} = \left\{(X,Z) \in \mathscr{F} : \|Z_{12}\| \leqslant \tilde{z}_{12}, \|Z_{21}\| \leqslant \tilde{z}_{21}\right\}.$$

When $\|\cdot\| = \|\cdot\|_2$ and $\|\cdot\| = \|\cdot\|_F$ respectively, $\mathscr{H}$ reduces to $\mathscr{F}$ and $\mathscr{G}$ respectively. Furthermore, $\|I_r\|_2 = 1$ and $\|I_r\|_F = \sqrt{r}$.

Our first result establishes an lower bound under any unitarily invariant norm when $\alpha^2 > \beta^2 + z_{12}^2 + z_{21}^2$. We allow both X and Z taking complex values, i.e. $X, Z \in \mathbb{C}^{p_1 \times p_2}$.

THEOREM 3. *Let $\alpha^2 > \beta^2 + z_{12}^2 + z_{21}^2$ and $r \leqslant \frac{p_1 \wedge p_2}{2}$. Then for all estimate $\tilde{V} \in \mathbb{O}_{p_2 \times r}$ of V based on $\hat{X}$, we have*

$$\inf_{\tilde{V}} \sup_{(X,Z)\in\mathscr{H}} \|\sin\Theta(V,\tilde{V})\| \geqslant \frac{1}{\sqrt{10+2\sqrt{5}}} \left(\frac{\alpha z_{12}\|I_r\|}{\alpha^2 - \beta^2 - z_{12}^2} \wedge \|I_r\| \right).$$

REMARK 1. If $z_{12} \leqslant z_{21}$, $\alpha z_{12} \geqslant \beta z_{21}$, $\tilde{z}_{12} \leqslant z_{12}\|I_r\|$ and $\tilde{z}_{21} \leqslant z_{21}\|I_r\|$, Theorem 3 implies

$$\inf_{\tilde{V}} \sup_{(X,\tilde{Z})\in\mathscr{H}} \|\sin\Theta(V,\hat{V})\| \geqslant \frac{1}{2\sqrt{10+2\sqrt{5}}} \left(\frac{\alpha\tilde{z}_{12} + \beta\tilde{z}_{21}}{\alpha^2 - \beta^2 - z_{12}^2 \wedge z_{21}^2} \wedge \|I_r\| \right).$$

For example, when X and Z are symmetrical, $z_{12} = z_{21}$ and $\tilde{z}_{12} = \tilde{z}_{21}$. Then it satisfies the assumption of Remark 1 thanks to $\alpha > \beta$.

In fact, $\frac{1}{2\sqrt{10+2\sqrt{5}}} \Big/ \frac{1}{8\sqrt{10}} \approx 3.3$. Hence, the estimation of Theorem 3 is better than Cai and Zhang's in some situations.

The following result extends Theorem 2 ([1]) from spectral norm to any unitarily invariant norm.

THEOREM 4. *Let $\alpha^2 \leqslant \beta^2 + z_{12}^2 \wedge z_{21}^2$ and $r \leqslant \frac{p_1 \wedge p_2}{2}$. Then for all estimate $\tilde{V} \in \mathbb{O}_{p_2 \times r}$ of V based on $\hat{X}$, we have*

$$\inf_{\tilde{V}} \sup_{(X,Z)\in\mathscr{H}} \|\sin\Theta(V,\hat{V})\| \geqslant \frac{1}{2\sqrt{2}}\|I_r\|.$$

2. Proof of Theorem 3

Firstly, we present a proposition which is needed in our later discussions.

PROPOSITION 1. *Suppose 2-by-2 matrix A satisfies*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad a,b,c,d \geqslant 0.$$

Let $V = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}$ *be the right singular vectors of* A. *If* $a^2 + c^2 > d^2 + b^2$, *we have*

$$|v_{21}| \geqslant \frac{2}{\sqrt{10+2\sqrt{5}}} \left(\frac{ab+cd}{a^2+c^2-b^2-d^2} \wedge 1 \right).$$

Proof. By solving the two eigenvalues of $A^H A$, one finds that the maximal eigenvalue of $A^H A$ is

$$\begin{aligned} \lambda_1 &= \frac{a^2+b^2+c^2+d^2+\sqrt{(a^2+b^2+c^2+d^2)^2-4(ad-bc)^2}}{2} \\ &= \frac{a^2+b^2+c^2+d^2+\sqrt{(a^2+c^2-b^2-d^2)^2+4(ab+cd)^2}}{2}. \end{aligned} \tag{4}$$

By the definition of singular vectors, $(\lambda_1 I_2 - A^H A) \begin{bmatrix} v_{11} \\ v_{21} \end{bmatrix} = 0$. Thus,

$$-(ab+cd)v_{11} + (\lambda_1 - b^2 - d^2)v_{21} = 0.$$

By $v_{11}^2 + v_{21}^2 = 1$ and the above identity, one finds that

$$|v_{21}| = \frac{ab+cd}{\sqrt{(\lambda_1-b^2-d^2)^2+(ab+cd)^2}}. \tag{5}$$

The assumption $a^2+c^2 > b^2+d^2$ with (4) and (5) implies

$$\begin{aligned} |v_{21}| &= \frac{ab+cd}{\sqrt{\left[\frac{a^2+b^2+c^2+d^2+\sqrt{(a^2+c^2-b^2-d^2)^2+4(ab+cd)^2}}{2} - b^2 - d^2\right]^2 + (ab+cd)^2}} \\ &= \frac{ab+cd}{\sqrt{\left[\frac{a^2+c^2-b^2-d^2+\sqrt{(a^2+c^2-b^2-d^2)^2+4(ab+cd)^2}}{2}\right]^2 + (ab+cd)^2}} \\ &\geqslant \frac{2}{\sqrt{10+2\sqrt{5}}} \left(\frac{ab+cd}{a^2+c^2-b^2-d^2} \wedge 1 \right). \quad \square \end{aligned}$$

The following two lemmas are also needed in the proof of Theorem 3.

LEMMA 1. ([4]) *Let* $A = (a_{ij}) \in \mathbb{C}^{p_1 \times p_2}$. *Then*

$$\max_{1 \leqslant i \leqslant p_1, 1 \leqslant j \leqslant p_2} |a_{ij}| \leqslant \|A\|_2.$$

LEMMA 2. ([5]) *Let* $\|\cdot\|$ *be a unitarily invariant norm,* $V,\hat{V}\in\mathbb{O}_{p,r}$ *with* $[V\ V_\perp]\in\mathbb{O}_p$ *and* $[\hat{V}\ \hat{V}_\perp]\in\mathbb{O}_p$*. Then*

$$\|\sin\Theta(V,\hat{V})\| = \|V_\perp^H\hat{V}\| = \|V^H\hat{V}_\perp\|.$$

For $V_1,V_2,V_3\in\mathbb{O}_{p,r}$,

$$\|\sin\Theta(V_2,V_3)\| \leqslant \|\sin\Theta(V_1,V_2)\| + \|\sin\Theta(V_1,V_3)\|.$$

We use Proposition 1 and Lemma 1–2 to prove Theorem 3.

The SVD of $\begin{bmatrix}\alpha & z_{12}\\ 0 & \beta\end{bmatrix}$ can be given as

$$\begin{bmatrix}\alpha & z_{12}\\ 0 & \beta\end{bmatrix} = \begin{bmatrix}u_{11} & u_{12}\\ u_{21} & u_{22}\end{bmatrix}\begin{bmatrix}\sigma_1 & 0\\ 0 & \sigma_2\end{bmatrix}\begin{bmatrix}v_{11} & v_{12}\\ v_{21} & v_{22}\end{bmatrix}^H. \tag{6}$$

Note that $\alpha^2>\beta^2+z_{12}^2+z_{21}^2$ implies $\alpha^2>\beta^2+z_{12}^2$. Then by Proposition 1,

$$|v_{21}| \geqslant \frac{2}{\sqrt{10+2\sqrt{5}}}\left(\frac{\alpha z_{12}}{\alpha^2-\beta^2-z_{12}^2}\wedge 1\right). \tag{7}$$

From (6), we know that

$$\begin{bmatrix}\alpha & 0\\ 0 & 0\end{bmatrix} + \begin{bmatrix}0 & z_{12}\\ 0 & \beta\end{bmatrix} = \begin{bmatrix}\sigma_1 u_{11}v_{11} & \sigma_1 u_{11}v_{21}\\ \sigma_1 u_{21}v_{11} & \sigma_1 u_{21}v_{21}\end{bmatrix} + \begin{bmatrix}\sigma_2 u_{12}v_{12} & \sigma_2 u_{12}v_{22}\\ \sigma_2 u_{22}v_{12} & \sigma_2 u_{22}v_{22}\end{bmatrix}. \tag{8}$$

We find that the constructions of $(X_1,Z_1),(X_2,Z_2)\in\mathscr{H}$ which given in Reference [2] can also be use for our proof. Hence, we apply the constructions to complete the proof.

Choose

$$X_1 := \begin{bmatrix}\sigma_1 u_{11}v_{11}I_r & \sigma_1 u_{11}v_{21}I_r & 0\\ \sigma_1 u_{21}v_{11}I_r & \sigma_1 u_{21}v_{21}I_r & 0\\ 0 & 0 & 0\end{bmatrix},\quad Z_1 := \begin{bmatrix}\sigma_2 u_{12}v_{12}I_r & \sigma_2 u_{12}v_{22}I_r & 0\\ \sigma_2 u_{22}v_{12}I_r & \sigma_2 u_{22}v_{22}I_r & 0\\ 0 & 0 & 0\end{bmatrix},$$

$$X_2 := \begin{matrix}r\\ r\\ p_1-2r\end{matrix}\begin{bmatrix}\alpha I_r & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix} \text{ and } Z_2 := \begin{matrix}r\\ r\\ p_1-2r\end{matrix}\begin{bmatrix}0 & z_{12}I_r & 0\\ 0 & \beta I_r & 0\\ 0 & 0 & 0\end{bmatrix}.$$

Then

$$X_1+Z_1 = X_2+Z_2 = \hat{X}_1 = \hat{X}_2 = \begin{bmatrix}\alpha I_r & z_{12}I_r & 0\\ 0 & \beta I_r & 0\\ 0 & 0 & 0\end{bmatrix}.$$

Denote $U_1=[u_{11}I_r\ u_{21}I_r\ 0]^T$, $V_1=[v_{11}I_r\ v_{21}I_r\ 0]^T$,

$$U_{1\perp} = \begin{bmatrix}u_{12}I_r & 0\\ u_{22}I_r & 0\\ 0 & I_{p_1-2r}\end{bmatrix} \text{ and } V_{1\perp} = \begin{bmatrix}v_{12}I_r & 0\\ v_{22}I_r & 0\\ 0 & I_{p_2-2r}\end{bmatrix}.$$

Hence,

$$\hat{X}_1 = \hat{X}_2 = \begin{bmatrix} \alpha I_r & z_{12} I_r & 0 \\ 0 & \beta I_r & 0 \\ 0 & 0 & 0 \end{bmatrix} = [U_1\ U_{1\perp}] \begin{bmatrix} \sigma_1 I_r & 0 & 0 \\ 0 & \sigma_2 I_r & 0 \\ 0 & 0 & 0 \end{bmatrix} [V_1\ V_{1\perp}]^H, \tag{9}$$

$$X_1 = \sigma_1 U_1 V_1^H \text{ and } Z_1 = U_{1\perp} \begin{bmatrix} \sigma_2 I_r & 0 \\ 0 & 0 \end{bmatrix} V_{1\perp}^H \tag{10}$$

thanks to (6) and (8).

From (9) and (10), one knows

$$U_1^H \hat{X}_1 V_1 = \sigma_1 I_r, \ \ U_{1\perp}^H \hat{X}_1 V_{1\perp} = \begin{bmatrix} \sigma_2 I_r & 0 \\ 0 & 0 \end{bmatrix}, \tag{11}$$

$$Z_{1_{12}} = \mathbb{P}_{U_1} Z_1 \mathbb{P}_{V_{1\perp}} = 0 \text{ and } Z_{1_{21}} = \mathbb{P}_{U_{1\perp}} Z_1 \mathbb{P}_{V_1} = 0. \tag{12}$$

By (11) and (12), $\sigma_r(U_1^H \hat{X}_1 V_1) = \sigma_1$, $\|U_{1\perp}^H \hat{X}_1 V_{1\perp}\|_2 = \sigma_2$ and $\|Z_{1_{12}}\| = \|Z_{1_{21}}\| = 0$. Note that $\sigma_1 \geqslant \alpha$ thanks to (6) and Lemma 1. From (6), $\alpha\beta = \sigma_1\sigma_2$. Moreover,

$$\sigma_1 \geqslant \alpha \text{ and } \sigma_2 \leqslant \beta.$$

These conclude $(X_1, Z_1) \in \mathscr{H}$.

On the other hand, denote $U_2 = [I_r\ 0\ \ 0]^T$, $V_2 = [I_r\ 0\ \ 0]^T$,

$$U_{2\perp} = \begin{bmatrix} 0 & 0 \\ I_r & 0 \\ 0 & I_{p_1-2r} \end{bmatrix} \text{ and } V_{2\perp} = \begin{bmatrix} 0 & 0 \\ I_r & 0 \\ 0 & I_{p_2-2r} \end{bmatrix},$$

one finds that the SVD can be given as follows,

$$X_2 = \begin{bmatrix} \alpha I_r & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = [U\ U_\perp] \begin{bmatrix} \alpha I_r & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} [V\ V_\perp]^H.$$

Then

$$U_2^H \hat{X}_2 V_2 = \alpha I_r, \ \ U_{2\perp}^H \hat{X}_2 V_{2\perp} = \begin{bmatrix} \beta I_r & 0 \\ 0 & 0 \end{bmatrix}, \tag{13}$$

$$Z_{2_{12}} = \mathbb{P}_{U_2} Z_2 \mathbb{P}_{V_{2\perp}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & z_{12} I_r & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } Z_{2_{21}} = \mathbb{P}_{U_{2\perp}} Z_2 \mathbb{P}_{V_2} = 0 \tag{14}$$

thanks to (9).

From (13)–(14), $\sigma_r(U_2^H \hat{X}_2 V_2) = \alpha$, $\|U_{2\perp}^H \hat{X}_2 V_{2\perp}\|_2 = \beta$, $\|Z_{2_{12}}\| = z_{12}\|I_r\|$ and $\|Z_{2_{21}}\| = 0$. Hence, $(X_2, Z_2) \in \mathscr{H}$.

Note that $\tilde{V} \in \mathbb{O}_{p_2 \times r}$ is an estimator of V based on $\hat{X}$, then

$$\sup_{(X,Z)\in\mathscr{H}} \|\sin\Theta(\tilde{V},V)\| \geqslant \max\left\{\|\sin\Theta(\tilde{V},V_1)\|, \|\sin\Theta(\tilde{V},V_2)\|\right\}$$
$$\geqslant \frac{1}{2}\left[\|\sin\Theta(\tilde{V},V_1)\| + \|\sin\Theta(\tilde{V},V_2)\|\right] \geqslant \frac{1}{2}\|\sin\Theta(V_1,V_2)\|.$$

This with Lemma 2 tells

$$\sup_{(X,Z)\in\mathscr{H}} \|\sin\Theta(V,\tilde{V})\| \geqslant \frac{1}{2}\|V_{2\perp}^H V_1\| = \frac{1}{2}\|v_{21}I_r\| = \frac{1}{2}|v_{21}|\cdot\|I_r\|.$$

The above inequality with (7) shows

$$\sup_{(X,Z)\in\mathscr{H}} \|\sin\Theta(V,\tilde{V})\| \geqslant \frac{1}{\sqrt{10+2\sqrt{5}}}\left(\frac{\alpha z_{12}\|I_r\|}{\alpha^2-\beta^2-z_{12}^2} \wedge \|I_r\|\right).$$

The proof is done. $\square$

3. Proof of Theorem 4

We introduce a useful lemma, which plays a key role for proof of Theorem 4.

LEMMA 3. ([2]) *Let* $V = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}$ *be right singular vectors of* $A = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$ *with* $a,b,d > 0$ *and* $a^2 \leqslant b^2 + d^2$,

$$|v_{21}| \geqslant \frac{1}{\sqrt{2}}.$$

Now, we can give the proof of Theorem 4.

Consider the following SVD of $\begin{bmatrix} \alpha & z_{12} \\ 0 & \beta \end{bmatrix}$,

$$\begin{bmatrix} \alpha & z_{12} \\ 0 & \beta \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}^H.$$

Note that $\alpha^2 \leqslant \beta^2 + z_{12}^2 \wedge z_{21}^2$ implies $\alpha^2 \leqslant \beta^2 + z_{12}^2$. Then by Lemma 3,

$$|v_{21}| \geqslant \frac{1}{\sqrt{2}}. \tag{15}$$

In fact, the constructions of $(X_1,Z_1),(X_2,Z_2) \in \mathscr{H}$ in the proof of Theorem 3 can also be use for following discussions. We borrow the conclusions about $(X_1,Z_1),(X_2,Z_2)$ in the proof of Theorem 3 for completeness.

Note that $\tilde{V} \in \mathbb{O}_{p_2 \times r}$ is an estimator of V based on $\hat{X}$, then

$$\sup_{(X,Z)\in\mathscr{H}} \|\sin\Theta(\tilde{V},V)\| \geqslant \max\left\{\|\sin\Theta(\tilde{V},V_1)\|, \|\sin\Theta(\tilde{V},V_2)\|\right\}$$

$$\geqslant \frac{1}{2}\left[\|\sin\Theta(\tilde{V},V_1)\| + \|\sin\Theta(\tilde{V},V_2)\|\right] \geqslant \frac{1}{2}\|\sin\Theta(V_1,V_2)\|.$$

This with Lemma 2 and (15) implies

$$\sup_{(X,Z)\in\mathscr{H}} \|\sin\Theta(V,\tilde{V})\| \geqslant \frac{1}{2}\|V_{2\perp}^H V_1\| = \frac{1}{2}\|v_{21} I_r\| = \frac{1}{2}|v_{21}| \cdot \|I_r\| \geqslant \frac{1}{2\sqrt{2}}\|I_r\|.$$

The proof is done. □

Acknowledgements. The author Chunguang Ren would like to thank the referee for the valuable and kind comments. The author Chunguang Ren is grateful to his advisor, Professor Youming Liu, for his helpful guidance. This paper is supported by the National Natural Science Foundation of China (No. 11771030) and the Beijing Natural Science Foundation (No. 1172001).

REFERENCES

[1] T. T. CAI AND A. R. ZHANG, *Rate-optimal perturbation bounds for singular subspaces with applications to high-dimensional statistics*, The Annals of Statistics. 46 (2018), 60–89.

[2] T. T. CAI AND A. R. ZHANG, *Supplement to Rate-optimal perturbation bounds for singular subspaces with applications to high-dimensional statistics*, 2018. doi:10.1214/17-AOS1541SUPP.

[3] R. A. HORN AND C. R. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

[4] R. A. HORN AND C. R. JOHNSON, *Matrix analysis* (Second Edition), Cambridge University Press, Cambridge, 2012.

[5] Y. M. LIU AND C. G. REN, *An optimal perturbation bound*, Mathematical Methods in the Applied Sciences 42 (2019), 3791–3798.

(Received September 22, 2019)

Chunguang Ren
College of Applied Sciences
Beijing University of Technology
Beijing, 100124, P.R. China
e-mail: renchunguang@emails.bjut.edu.cn