Applied Mathematics and Computation 143 (2003) 559–563
APPLIED MATHEMATICS AND COMPUTATION
www.elsevier.com/locate/amc

A note on variational representation for singular values of matrix

Zhi-Hao Cao *, Li-Hong Feng

Department of Mathematics and Laboratory of Mathematics for Nonlinear Science, Fudan University, Shanghai 200433, China

Abstract

In this note we address the variational property of singular values of matrix and point out that a theorem in [Matrix Computations, John Hopkins University Press, Baltimore, MD, 1989,1993,1996] is incomplete.

Keywords: Singular value; Symmetric matrix; Saddle point problem

Let $A \in \mathscr{R}^{m\times n}$ and let $\sigma_i(A)$ denote the ith largest singular value of A. The following theorem is given in [2] (Theorem 8.3.1 in eds. 1989 and 1993, and Theorem 8.6.1 in ed. 1996):

Theorem 1. *If $A \in \mathscr{R}^{m\times n}$, then for $k = 1 : \min\{m, n\}$*

$$\sigma_k(A) = \max_{\substack{\dim(S)=k \\ \dim(T)=k}} \min_{\substack{x\in S \\ y\in T}} \frac{y^{\mathrm{T}}Ax}{\|x\|_2\|y\|_2}, \tag{1}$$

$$\sigma_k(A) = \max_{\dim(S)=k} \min_{\substack{x\in S \\ x\neq 0}} \frac{\|Ax\|_2}{\|x\|_2}, \tag{2}$$

where $S \subseteq \mathscr{R}^n$ and $T \subseteq \mathscr{R}^m$ are subspaces.

* Corresponding author.
E-mail address: zcao@fudan.edu.cn (Z.-H. Cao).

0096-3003/02/$ - see front matter
doi:10.1016/S0096-3003(02)00385-5

However, the first part of Theorem 1, i.e. (1), is incomplete, as the following simple example shows.

Example. Let

$$A = \begin{pmatrix} 1 & \\ & 1/2 \end{pmatrix},$$

then we have, obviously, $\sigma_2(A) = 1/2$. On the other hand, let

$$\hat{x} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \quad \hat{y} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

then

$$\min_{\substack{x,y\in\mathscr{R}^2 \\ x,y\neq 0}} \frac{y^{\mathrm{T}}Ax}{\|x\|_2\|y\|_2} \leqslant \frac{\hat{y}^{\mathrm{T}}A\hat{x}}{\|\hat{x}\|_2\|\hat{y}\|_2} = 0,$$

since $\hat{y}^{\mathrm{T}}A\hat{x} = 0$.

In fact, a variational representation for singular values of matrix can be given as follows.

Theorem 2. *If $A \in \mathscr{R}^{m\times n}$, then for $k = 1 : \min\{m, n\}$*

$$\sigma_k(A) = \max_{\dim(S)=k} \min_{\substack{x\in S \\ x\neq 0}} \max_{\substack{y\in\mathscr{R}^m \\ y\neq 0}} \frac{y^{\mathrm{T}}Ax}{\|x\|_2\|y\|_2} \tag{3}$$

$$= \max_{\dim(T)=k} \min_{\substack{y\in T \\ y\neq 0}} \max_{\substack{x\in\mathscr{R}^n \\ x\neq 0}} \frac{y^{\mathrm{T}}Ax}{\|x\|_2\|y\|_2}. \tag{4}$$

Proof. We note that the second part of Theorem 1, i.e. (2), is correct (see also [3]):

$$\sigma_k(A) = \max_{\dim(S)=k} \min_{\substack{x\in S \\ x\neq 0}} \frac{\|Ax\|_2}{\|x\|_2}. \tag{5}$$

First, let us prove (3), we have

$$\begin{aligned}\max_{\dim(S)=k}\min_{\substack{x\in S\\ x\neq 0}}\max_{\substack{y\in\mathscr{R}^m\\ y\neq 0}}\frac{y^{\mathrm T}Ax}{\|x\|_2\|y\|_2} &= \max_{\dim(S)=k}\min_{\substack{x\in S\\ x\neq 0}}\frac{1}{\|x\|_2}\max_{\substack{y\in\mathscr{R}^m\\ y\neq 0}}\frac{(Ax)^{\mathrm T}y}{\|y\|_2}\\ &= \max_{\dim(S)=k}\min_{\substack{x\in S\\ x\neq 0}}\frac{\|(Ax)^{\mathrm T}\|_2}{\|x\|_2} = \max_{\dim(S)=k}\min_{\substack{x\in S\\ x\neq 0}}\frac{\|Ax\|_2}{\|x\|_2}\\ &= \sigma_k(A).\end{aligned}$$

Then, let us prove (4), we have

$$\begin{aligned}\max_{\dim(T)=k}\min_{\substack{y\in T\\ y\neq 0}}\max_{\substack{x\in\mathscr{R}^n\\ x\neq 0}}\frac{y^{\mathrm T}Ax}{\|x\|_2\|y\|_2} &= \max_{\dim(T)=k}\min_{\substack{y\in T\\ x\neq 0}}\frac{1}{\|y\|_2}\max_{\substack{x\in\mathscr{R}^n\\ x\neq 0}}\frac{(y^{\mathrm T}A)x}{\|x\|_2}\\ &= \max_{\dim(T)=k}\min_{\substack{y\in T\\ y\neq 0}}\frac{\|y^{\mathrm T}A\|_2}{\|y\|_2} = \max_{\dim(T)=k}\min_{\substack{y\in T\\ y\neq 0}}\frac{\|A^{\mathrm T}y\|_2}{\|y\|_2}\\ &= \sigma_k(A^{\mathrm T}) = \sigma_k(A). \qquad \square\end{aligned}$$

Without lose of generality, we assume that $m \leqslant n$, then we have the following corollary of Theorem 2.

Corollary 3. *Let $A \in \mathscr{R}^{m\times n}$, $m \leqslant n$, then*

$$\sigma_1 = \max_{\substack{x\in\mathscr{R}^n\\ x\neq 0}}\max_{\substack{y\in\mathscr{R}^m\\ y\neq 0}}\frac{y^{\mathrm T}Ax}{\|x\|_2\|y\|_2}, \quad \sigma_m = \min_{\substack{y\in\mathscr{R}^m\\ y\neq 0}}\max_{\substack{x\in\mathscr{R}^n\\ x\neq 0}}\frac{y^{\mathrm T}Ax}{\|x\|_2\|y\|_2},$$

where the singular values of A are ordered as

$$\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_m.$$

The approximation of saddle point problem, for example, the mixed finite element solution of the Stokes equations describing slow incompressible viscous flow leads to a symmetric indefinite discrete system:

$$\begin{pmatrix} A & B^{\mathrm T}\\ B & 0\end{pmatrix}\begin{pmatrix} u\\ p\end{pmatrix} = \begin{pmatrix} f\\ 0\end{pmatrix} \tag{6}$$

for the pressure component p and velocity component u, where $A \in \mathscr{R}^{n\times n}$ is symmetric positive definite, $B \in \mathscr{R}^{m\times n}$, $m \leqslant n$, is a matrix with full row rank. The matrix form of the so called Babuska–Brezzi condition is as follows (cf. [1])

$$\max_{\substack{u\in\mathscr{R}^n\\ u\neq 0}}\frac{p^{\mathrm T}Bu}{(u^{\mathrm T}Au)^{1/2}} \geqslant \gamma(p^{\mathrm T}M_p p)^{1/2}, \quad \forall p \in \mathscr{R}^m,\ p \neq 0, \tag{7}$$

where (pressure mass matrix) $M_p \in \mathscr{R}^{m\times m}$ is symmetric positive definite, γ is a scalar.

Using Corollary 3 we can easily derive the following result (cf. [1]).

Theorem 4. *The upper bound of the scalars* $\{\gamma\}$ *which satisfy* (7) *is the smallest singular value* $\sigma_{\min}(M_p^{-1/2}BA^{-1/2})$ *of matrix* $M_p^{-1/2}\ BA^{-1/2}$ *and the Babuska–Brezzi condition* (7) *can be expressed as*

$$\frac{p^{\mathrm{T}}(BA^{-1}B^{\mathrm{T}})p}{p^{\mathrm{T}}M_p p} \geqslant \gamma^2, \quad \forall p \in \mathscr{R}^m,\ \ p \neq 0. \tag{8}$$

Proof. For $u \neq 0$ and $p \neq 0$ we have

$$\frac{p^{\mathrm{T}}Bu}{(u^{\mathrm{T}}Au)^{1/2}(p^{\mathrm{T}}M_p p)^{1/2}} = \frac{\hat{p}^{\mathrm{T}}M_p^{-1/2}BA^{-1/2}\hat{u}}{\|\hat{u}\|_2\|\hat{p}\|_2},$$

where $\hat{u} = A^{1/2}u$ and $\hat{p} = M_p^{1/2}p$. Corollary 3 implies

$$\begin{aligned}\min_{\substack{p\in\mathscr{R}^m\\ p\neq 0}}\max_{\substack{u\in\mathscr{R}^n\\ u\neq 0}}\frac{p^{\mathrm{T}}Bu}{(u^{\mathrm{T}}Au)^{1/2}(p^{\mathrm{T}}M_p p)^{1/2}} &= \min_{\substack{\hat{p}\in\mathscr{R}^m\\ \hat{p}\neq 0}}\max_{\substack{\hat{u}\in\mathscr{R}^n\\ \hat{u}\neq 0}}\frac{\widehat{p}^{\mathrm{T}}(M_p^{-1/2}BA^{-1/2})\widehat{u}}{\|\hat{u}\|_2\|\hat{p}\|_2}\\ &= \sigma_{\min}(M_p^{-1/2}BA^{-1/2}).\end{aligned} \tag{9}$$

Comparing (7) and (9) we have

$$\gamma \leqslant \sigma_{\min}(M_p^{-1/2}BA^{-1/2}). \tag{10}$$

Since

$$\sigma^2_{\min}(M_p^{-1/2}BA^{-1/2}) = \lambda_{\min}(M_p^{-1/2}BA^{-1}B^{\mathrm{T}}M_p^{-1/2}),$$

(8) follows from Courant–Fisher Minimax Theorem [2] and (10). □

Remark. The bound condition is defined as [1]

$$\max_{\substack{u\in\mathscr{R}^n\\ u\neq 0}}\frac{p^{\mathrm{T}}Bu}{(u^{\mathrm{T}}Au)^{1/2}} \leqslant \Gamma(p^{\mathrm{T}}M_p p)^{1/2}, \quad \forall p \in \mathscr{R}^m,\ \ p \neq 0. \tag{11}$$

Using Corollary 3 we can analogously deduce that the lower bound of the scalars $\{\Gamma\}$ which satisfy (11) is the largest singular value $\sigma_{\max}(M_p^{-1/2}BA^{-1/2})$ of matrix $M_p^{-1/2}BA^{-1/2}$ and the condition (11) can be expressed as

$$\Gamma^2 \geqslant \frac{p^{\mathrm{T}}(BA^{-1}B^{\mathrm{T}})p}{p^{\mathrm{T}}M_p p}, \quad \forall p \in \mathscr{R}^m,\ \ p \neq 0.$$

Acknowledgements

This work is supported by NSFC Project 10171021, the Foundation of National Key Laboratory of Computational Physics and the Doctoral Point Foundation of China.

References

[1] F. Brezzi, M. Fortin, Mixed and Hybrid Finite Element Methods, Springer-Verlag, New York, 1991.
[2] G.H. Golub, C.F. Van Loan, Matrix Computations, John Hopkins University Press, Baltimore, MD, 1989, 1993, 1996.
[3] G.W. Stewart, Introduction to Matrix Computations, Academic Press, New York, 1973.