SRC TR 88-82

TECHNICAL RESEARCH REPORT

On the Norms Used in Computing the Structured Singular Value

by

N. K. Tsing

SYSTEMS RESEARCH CENTER

UNIVERSITY OF MARYLAND

COLLEGE PARK, MARYLAND 20742

On the Norms Used in Computing the Structured Singular Value

Nam-Kiu Tsing

Systems Research Center and Electrical Engineering Department
University of Maryland, College Park, Maryland 20742

Abstract

Different norms are considered to replace the Euclidean norm in an algorithm given by Fan and Tits (IEEE Trans. Automat. Contr., AC-33, pp.284-289, 1988) which is used for the computation of the structured singular value of any matrix. It is shown that the ℓ_1-norm is the best possible norm in a certain sense.

Recently, there has been a considerable amount of interest in the study of the structured singular value, the concept of which was originated by Doyle [1] and is used as a tool for the analysis and synthesis of feedback systems with structured uncertainties (e.g., see [2,3] and their references).

Let M be an $n \times n$ complex matrix, and $\mathcal{K} = (k_1, \ldots, k_m)$ an m-tuple of positive integers which satisfies $\sum_{i=1}^{m} k_i = n$. For $i = 1, \ldots, m$, denote the ith-block-projection matrix by P_i = block diag $(O_{k_1}, \ldots, O_{k_{i-1}}, I_{k_i}, O_{k_{i+1}}, \ldots, O_{k_m})$, where O_k and I_k are the zero matrix and identity matrix, respectively, of order $k \times k$ for any positive integer k. Then the structured singular value of M with respect to the block-structure $\mathcal{K}$ is the nonnegative scalar

$$\mu(M) = \max_{x \in \partial B} \{\|Mx\| : \|P_i x\| \, \|Mx\| = \|P_i M x\|, i = 1, \ldots, m\} ,$$

where $\|\cdot\|$ denotes the Euclidean (ℓ_2 -) norm in $\mathbb{C}^n$ and ∂B the corresponding unit sphere. One major issue in the study of $\mu(M)$ is the computation of it. In [2], the authors devise an algorithm [2, Algorithm 1], which we shall explain immediately, to compute $\mu(M)$. They first define, for any real number α, the hermitian matrices

$$A_i(\alpha) = \alpha P_i - M^H P_i M, \text{ for } i = 1, \ldots, m,$$

and the m-form numerical range associated with $A_1(\alpha), \ldots, A_m(\alpha)$

$$W(\alpha) = \left\{(v_1, \ldots, v_m) \in \mathbb{R}^m : \exists x \in \partial B \text{ such that } v_i = x^H A_i(\alpha) x \text{ for all } i\right\}.$$

A function $c(\cdot) : \mathbb{R} \to \mathbb{R}$, which depends on M and $\mathcal{K}$, is then defined by

$$c(\alpha) = \min \{\|v\| : v \in W(\alpha)\},$$

where $\|\cdot\|$ denotes again the Euclidean norm (on $\mathbb{R}^m$ this time). Then they show [2, Corollary 1 and Proposition 1] that, for any matrix M with structure $\mathcal{K}$, $c(\cdot)$ satisfies

$$c(\cdot) \text{ is continuous,} \tag{1}$$

$$c(\mu^2) = 0, \; c(\alpha) > 0 \text{ for all } \alpha > \mu^2, \tag{2}$$

and

$$c(\alpha + s) \leq c(\alpha) + s \text{ for all } s \geq 0 \text{ and real } \alpha, \tag{3}$$

where $\mu = \mu(M)$. The initial step in Algorithm 1 of [2] is to set $\alpha_0 = \bar{\sigma}^2(M)$, where $\bar{\sigma}(M)$ is the largest singular value of M. The iteration step is to set $\alpha_{k+1} = \alpha_k - c(\alpha_k)$ for $k = 0, 1, 2, \ldots$. The authors show in [2, Theorem 2] that, since $c(\cdot)$ satisfies (1) through (3), Algorithm 1 will generate a monotonic decreasing sequence $\{\alpha_k\}$ with $\lim_{k\to\infty} \alpha_k = \mu^2$. The structured singular value μ can thus be obtained. They also remark in the footnote that the Euclidean norm in the definition of $c(\cdot)$ can be replaced by the ℓ_1 -norm to get the strongest version of Proposition 1. The purpose of the present note is to elaborate this remark.

Suppose we consider any norm $N(\cdot)$ on $\mathbb{R}^m$ instead of the Euclidean norm. Similar to the case of $c(\cdot)$, we may define a function $c_N : \mathbb{R} \to \mathbb{R}$ (which also depends on M and $\mathcal{K}$) by

$$c_N(\alpha) = \min\{N(v) : v \in W(\alpha)\}.$$

It is not hard to see that $c_N(\cdot)$ always satisfies (1) and (2) ; and if in addition $c_N(\cdot)$ satisfies (3) then the function $c(\cdot)$ in Algorithm 1 can be replaced by $c_N(\cdot)$. Let

$$\mathcal{N} = \{N(\cdot) : N(\cdot) \text{ is a norm on } \mathbb{R}^m \text{and } c_N(\cdot) \text{ satisfies}$$
$$\text{condition (3) for all matrices } M \text{ with structure } \mathcal{K} \} .$$

Then Algorithm 1 will work with $c(\cdot)$ being replaced by any $c_N(\cdot)$ where $N(\cdot) \in \mathcal{N}$. In view of the iteration step of the algorithm, we may want to choose a norm $N_0(\cdot) \in \mathcal{N}$ such that

$$c_{N_0}(\alpha) \geq c_N(\alpha) \text{ for all } N(\cdot) \in \mathcal{N} \text{ and } \alpha > \mu^2,$$

so that the resulting algorithm has the fastest convergent rate and is thus the most efficient. The following result shows that $\|\cdot\|_1$, i.e., the ℓ_1-norm on $\mathbb{R}^m$ defined by

$$\|v\|_1 = \sum_{i=1}^{m} |\, v_i \,| \text{ for all } v = (v_1, \ldots, v_m) \in \mathbb{R}^m,$$

will give such a "best possible" norm.

Theorem. Let $\mathcal{K} = (k_1, \ldots, k_m)$ be a given block structure. Then

(a) $\| \cdot \|_1 \in \mathcal{N}$, and

(b) for any $N(\cdot) \in \mathcal{N}$, $\|v\|_1 \geq N(v)$ for all $v \in \mathbb{R}^m$,

so that

$$c_1(\alpha) := \min\{\|v\|_1 \ : \ v \in W(\alpha)\} \ \geq \ c_N(\alpha)$$

for any real α and complex matrix M with block structure $\mathcal{K}$.

Proof. (a) The proof of the fact that $c_1(\cdot)$ satisfies (3) is similar to (and simpler than) that of $c(\cdot)$ given in the proof of Proposition 1 in [2]. Hence $\| \cdot \|_1 \in \mathcal{N}$.

(b) We prove by contradiction. Suppose $N(\cdot) \in \mathcal{N}$ but $N(v) > \|v\|_1$ for some $v \in \mathbb{R}^m$. Let $\{e_1, \ldots, e_m\}$ be the standard basis for $\mathbb{R}^m$. Without loss of generality we may assume

$$N(e_1) \ = \ 1 + q \ > \ \|e_1\|_1 \ = \ 1;$$

for if $N(e_i) \leq \|e_i\|_1$ for all $i \ = \ 1, \ldots, m$, then

$$N(v) \ = \ N\left(\sum_{i=1}^{m} \ v_i e_i\right) \leq \sum_{i=1}^{m} \ | \ v_i \ | \ N(e_i) \leq \sum_{i=1}^{m} \ | \ v_i \ | \ \|e_i\|_1 \ = \ \|v\|_1$$

for all $v = (v_1, \ldots, v_m) \in \mathbb{R}^m$, which contradicts the assumption. Let

$$L \ = \ \{v \in \mathbb{R}^m : v_1 \ + \ w_2 v_2 \ + \ \cdots + w_m v_m \ = \ 1\}$$

be a supporting hyperplane to the ball in $\mathbb{R}^m$ with respect to $N(\cdot)$ with radius $(1 + q)$. As e_1 is on L, we have

$$\min_{v \in L} N(v) \ = \ N(e_1) \ = \ 1 + q.$$

Note that some of the w_i's may be zero. However, for any $\epsilon > 0$, we can choose nonzero $w_2', \ldots, w_m' \in \mathbb{R}$ which are arbitrary close to $w_2, \ldots, w_m$ respectively, such that if

$$L' \ = \ \{v \in \mathbb{R}^m : \ v_1 + w_2' v_2 + \cdots + w_m' v_m \ = \ 1\}$$

then

$$\min_{v \in L'} N(v) > N(e_1) - \epsilon.$$

Now choose $\lambda, \beta_1, \ldots, \beta_m \in \mathbb{R}$ such that $\lambda - \beta_1^2 = 1$ and $\lambda - \beta_i^2 = 1/w_i'$ for $i = 2, \ldots, m$. Then, by defining

$$M = \sum_{i=1}^{m} \beta_i P_i \tag{4}$$

(which depends on ϵ, because β_i's depend on ϵ), we have, for any real α,

$$\begin{aligned} A_i(\alpha) = \ & \alpha P_i - \left(\sum_{j=1}^{m} \beta_j P_j \right) P_i \left(\sum_{j=1}^{m} \beta_j P_j \right) \\ = \ & (\alpha - \beta_i^2) P_i \ \text{ for } i = 1, \ldots, m \end{aligned}$$

and

$$\begin{aligned} c_N(\alpha) = \ & \min \left\{ N(v) : v_i = x^H A_i(\alpha) x, \ i = 1, \ldots, m, \ x \in \partial B \right\} \\ = \ & \min \left\{ N(v) : v_i = (\alpha - \beta_i^2) a_i, \ a_i \geq 0, \ i = 1, \ldots, m, \right. \\ & \left. a_1 + \cdots + a_m = 1 \right\}. \end{aligned}$$

Hence

$$\begin{aligned} c_N(\lambda - 1) = \ & \min \left\{ N(v) : v_i \ = \ (\lambda - 1 - \beta_i^2) a_i, \ a_i \geq 0, \ i = 1, \ldots, m, \right. \\ & \left. a_1 + \cdots + a_m = 1 \right\} \\ = \ & \min \left\{ N(v) : v_1 = 0, \ v_i = (\lambda - 1 - \beta_i^2) a_i, \ a_i \geq 0, \right. \\ & \left. i = 2, \ldots, m, \ a_1 \geq 0, \ a_1 + \cdots + a_m = 1 \right\} . \end{aligned}$$

By putting $(a_1, \ldots, a_m) = e_1$, we get

$$c_N(\lambda - 1) \ = \ N(0) \ = \ 0.$$

Also,

$$\begin{aligned}
c_N(\lambda) &= \min\{N(v):\ v_i = (\lambda - \beta_i^2)a_i,\ a_i \geq 0,\ i = 1,\ldots,m, \\
&\qquad\qquad a_1 + \cdots + a_m = 1\} \\
&= \min\{N(v):\ v_1 = a_1 \geq 0,\ v_i = a_i/w_i',\ a_i \geq 0,\ i = 2,\ldots,m, \\
&\qquad\qquad a_1 + \cdots + a_m = 1\} \\
&= \min\{N(v):\ v_1 + w_2'v_2 + \cdots + w_m'v_m = 1,\ v_1 \geq 0, \\
&\qquad\qquad w_i'v_i \geq 0,\ i = 2,\ldots,m\} \\
&\geq \min\{N(v):\ v \in L'\} \\
&> N(e_1) - \epsilon \\
&= (1 + q) - \epsilon \\
&= (1+q) + c_N(\lambda - 1) - \epsilon.
\end{aligned}$$

As $\epsilon > 0$ is arbitrary, we may choose $0 < \epsilon < q$ so that, for the matrix M defined in (4), we have

$$c_N((\lambda - 1) + 1) = c_N(\lambda) > c_N(\lambda - 1) + 1.$$

As a result, $c_N(\cdot)$ does not satisfy (3) for this matrix M, and hence $N(\cdot) \notin \mathcal{N}$. Thus, if $N(\cdot) \in \mathcal{N}$ then we must have

$$\|v\|_1 \geq N(v) \text{ for all } v \in \mathbb{R}^m,$$

and hence

$$\begin{aligned}
c_1(\alpha) &= \min\{\|v\|_1 : v \in W(\alpha)\} \\
&\geq \min\{N(v) : v \in W(\alpha)\} \\
&= c_N(\alpha) \quad \text{for any real } \alpha.
\end{aligned}$$

□

Now the problem remains to devise a method for computing $c_1(\alpha)$, so that the algorithm can be implemented. The computation of $c_1(\alpha)$ (or $c(\alpha)$) for any given M

and $\mathcal{K}$ can be, in general, difficult. However, similar to the case of $c(\alpha)$, there are existing methods for computing the value

$$c_1'(\alpha) := \min\{\|v\|_1 : v \in co\ W(\alpha)\},$$

where $co\ W(\alpha)$ denotes the convex hull of $W(\alpha)$. Let $< \cdot, \cdot >$ denote the usual inner product in $\mathbb{R}^m$. Since the ℓ_∞-norm ($\|\cdot\|_\infty$) is the dual norm of the ℓ_1-norm, we have

$$\begin{aligned} c_1'(\alpha) &= \min_{v \in co W(\alpha)} \|v\|_1 \\ &= \min_{v \in co W(\alpha)} \max_{\substack{a \in \mathbb{R}^m \\ \|a\|_\infty \le 1}} < v, a > . \end{aligned} \tag{5}$$

As $co\ W(\alpha)$ and $\{a \in \mathbb{R}^m : \|a\|_\infty \le 1\}$ are convex sets, and $< \cdot, \cdot >$ is bilinear, the max and min in (5) can be interchanged to yield

$$\begin{aligned} c_1'(\alpha) &= \max_{\substack{a \in \mathbb{R}^m \\ \|a\|_\infty \le 1}} \min_{v \in co W(\alpha)} < v, a > \\ &= \max_{|a_i| \le 1} \min_{x \in \partial B} \sum_{i=1}^m a_i x^H A_i(\alpha) x \\ &= \max_{|a_i| \le 1} \min_{x \in \partial B} x^H \left(\sum_{i=1}^m a_i A_i(\alpha) \right) x \\ &= \max_{|a_i| \le 1} \lambda_{\min} \left(\sum_{i=1}^m a_i A_i(\alpha) \right) \\ &= - \min_{|a_i| \le 1} \lambda_{\max} \left(\sum_{i=1}^m a_i A_i(\alpha) \right), \end{aligned} \tag{6}$$

where $\lambda_{\min}$ and $\lambda_{\max}$ denote respectively the smallest and largest eigenvalue. Several existing algorithms are available for solving the convex problem (6). We refer to a recent paper by Boyd and Yang [4, Section 7] the discussion on the details, including the advantages and disadvantages, of these numerical algorithms. As a result, if $W(\alpha)$ is convex (which is always the case when $m \le 3$; see [5] for example), $c_1(\alpha)$ $(= c_1'(\alpha))$ could be computed with any of these algorithms. If $m \ge 4$, then $W(\alpha)$ may not be convex. In

this case, replacing $c(\cdot)$ by $c_1'(\cdot)$ in Algorithm 1 of [2] will result in a sequence $\{\sqrt{\alpha_k}\}$ with limit

$$\mu' = \inf\{\sqrt{\alpha} : c_1'(\beta) > 0 \text{ for all } \beta \geq \alpha\},$$

which is clearly an upper bound for $\mu(M)$. This is exactly the same situation as the case of $c(\cdot)$ and $c'(\cdot)$ discussed in [2].

Acknowledgement. The author wishes to thank Drs. M.K.H. Fan and A.L. Tits for introducing their interesting paper to him and their helpful comments.

References

[1] J.C. Doyle, "Analysis of feedback systems with structured uncertainties", Proc. IEE-D, vol. 129, no.6, pp.242-250, 1982.

[2] M.K.H. Fan and A.L. Tits, "m-Form numerical range and the computation of the structured singular value", IEEE Trans. Automat. Contr., vol.AC-33, pp.284-289, 1988.

[3] A.K. Packard, "What's new with μ : Structured uncertainty in multivariable control", University of California, Ph.D. Thesis, Berkeley, California, 1988.

[4] S. Boyd and Q.-P. Yang, "Structured and simultaneous Lyapunov functions for system stability problem", Stanford Electronics Laboratories, Stanford University, Tech. Rep. L-104-88-1, 1988.

[5] M.K.H. Fan and A.L. Tits, "On the generalized numerical range", Linear and Multilinear Algebra, vol.21, no.3, pp.313-320, 1987.