Systems & Control Letters 17 (1991) 453–464
North-Holland

An algorithm for computing the distance to uncontrollability

L. Elsner and C. He *

Fakultät fur Mathematik, Universität Bielefeld, Postfach 8640, 4800 Bielefeld 1, Germany

Received 25 February 1991
Revised 18 May 1991

Abstract: In this paper, we present an algorithm to compute the distance to uncontrollability. The problem of computing the distance is an optimization problem of minimizing $\sigma(x, y)$ over the complete plane. This new approach is based on finding zero points of grad $\sigma(x, y)$. We obtain the explicit expression of the derivative matrix of grad $\sigma(x, y)$. The Newton's method and the bisection method are applied to approach these zero points. Numerical results show that these methods work well.

Keywords: Controllability; distance to uncontrollability; singular value decomposition; Newton's method.

1. Introduction

One of the fundamental concepts in linear control theory is that of controllability. A pair (A, B) of matrices $A \in R^{n \times n}$, $B \in R^{n \times m}$ is controllable if in the system

$$\dot{x} = Ax + Bu, \tag{1.1}$$

for any initial state x_0, final state x_1 and any $t_1 > 0$, there is a continuous function $u(t)$ such that the solution of (1.1) with $x(0) = x_0$ satisfies $x(t_1) = x_1$. It is well known that (A, B) is controllable iff

$$\operatorname{rank}([A - sI, B]) = n, \forall s \in C. \tag{1.2}$$

In [9], Paige defined the 'distance to uncontrollability' as the spectral norm distance of the pair (A, B) from the set of all uncontrollable pairs:

$$d(A, B) = \min\{ \|[E, F]\| : (A + E, B + F) \text{ uncontrollable} \}, \tag{1.3}$$

where $\| \ \|$ denotes the spectral norm, and $[E, F]$ is the $n \times (n + m)$ matrix formed by the columns of E followed by those of F. It was pointed out by Eising [5,6] that $d(A, B)$ admits the following description

$$d(A, B) = \min_{s \in C} \sigma_n([A - sI, B]) = \min_{s \in C} \sigma(s), \tag{1.4}$$

where $\sigma_n(G)$ denotes the n-th singular value of a $n \times (n + m)$ matrix G. It is clear that the problem of finding the distance to uncontrollability is the problem of minimizing $\sigma(s)$ over the complex plane.

Another characterization of $d(A, B)$ is given by

$$d(A, B) = \min\left\{ \| q^H [A - q^H A q I, B] \| : \| q \| = 1 \right\}, \tag{1.5}$$

where $\| \ \|$ is the Euclidean vector norm [5,12]. There are several algorithms in the literature for calculating $d(A, B)$. They are based on the minimization of $\sigma(s)$. Their main drawback is that they need a good starting point to converge [2,4,5,12]. Here we propose to use Newton's method with damping.

* Supported by the Alexander von Humboldt research foundation.

0167-6911/91/$03.50

This method is known to show convergence also for not so good starting values, a behaviour observed in our examples too.

We are able to use this method, because we can explicitly calculate the first and the second partial derivatives of $\sigma(x, y) = \sigma(x + iy) = \sigma(s)$ using the singular value decomposition (SVD) of $[A - sI, B]$. Let

$$[A - sI, B] = U\Sigma V^{\mathrm{H}} \tag{1.6}$$

be the SVD, where Σ is the $n \times (m + n)$ diagonal matrix with diagonal elements $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_n \geqslant 0$, and U and V are the $n \times n$ resp. $(m + n) \times (m + n)$ unitary matrices, the columns of which are the left resp. right normalized singular vectors of $[A - sI, B]$. If σ_n is a simple singular value then the normalized left singular vector $u_n(s)$ (the n-th column of U), and the normalized right vector $v_n(s)$ (the n-th column of V) are uniquely determined by (1.6) up to a common factor and so

$$f(s) = v_n^{\mathrm{H}}(s)\begin{pmatrix} u_n(s) \\ 0 \end{pmatrix} \tag{1.7}$$

is well defined. This function plays an important role, it is shown that

$$\frac{\partial\sigma(x + iy)}{\partial x} = -\mathrm{Re}\, f(x + iy), \frac{\partial\sigma(x + iy)}{\partial y} = -\mathrm{Im}\, f(x + iy), \tag{1.8}$$

and hence the zero points of $f(s)$ are the critical points of the function $\sigma(s)$. In addition, we have

$$\sigma(s)f(s) = u_n^{\mathrm{H}}(s)(A - sI)u_n(s),$$

as $[A - sI, B]^{\mathrm{H}}u_n(s) = \sigma(s)v_n(s)$. This shows that the critical points satisfy $s = u_n^{\mathrm{H}}(s)Au_n(s)$, and hence lie in the field of values of A.

The paper is organised as follows. In Section 2 we study the function $\sigma(x, y) = \sigma(x + iy)$. It is analytic as a function of the real parameters x and y for all but a finite number of points. We calculate the first and second derivatives of $\sigma(x, y)$ using an SVD. Here we treat a slightly more general case. In Section 3, the connection between zeros of $f(s)$ and local minima of $\sigma(x, y)$ is studied. Two criteria, one analytic and one in matrix terms, are given which guarantee a critical point of $\sigma(x, y)$ to be a local minimum. In Section 4 several possibilities of using Newton's method are outlined. They are the cases of real or complex parameters s. Numerical results and some discussions concerning the case of multiple singular values are given in Section 5 and 6 respectively.

2. The explicit expressions of the first and second derivatives of $\sigma(x, y)$

In this section, the main results are the explicit expressions of the first and second derivatives of $\sigma(x, y)$ given in (2.10) and (2.11). Let us consider a more general case of a complex matrix $G(s) = G_1 + sG_2$ with a real parameter s. It is well known that a singular value of $G(s)$ is analytic if it is simple [13]. In the following theorem, we give explicit expressions of its first and second derivatives. More general results can be found in [10,11].

Theorem 1. *Let $G(s) = G_1 + sG_2$ be an $n \times p$ complex matrix ($n \leq p$) with a real parameter s, and $G(s) = U(s)\Sigma(s)V(s)^{\mathrm{H}}$ be the SVD of $G(s)$ with the last singular value $\sigma_n(s) = \sigma(s)$ being simple, then*

$$\dot{\sigma} = \frac{\mathrm{d}\sigma}{\mathrm{d}s} = \mathrm{Re}\big(v_n^{\mathrm{H}}(s)G_2^{\mathrm{H}}u_n(s)\big) \tag{2.1}$$

and

$$\ddot{\sigma} = \frac{\mathrm{d}^2\sigma}{(\mathrm{d}s)^2} = \mathrm{Re}\big(v_{\mathrm{d}s}^{\mathrm{H}} G_2^{\mathrm{H}} u_n(s) + v_n^{\mathrm{H}}(s) G_2^{\mathrm{H}} u_{\mathrm{d}s}\big) + \mathrm{Re}\left(-\frac{\mathrm{i}}{\sigma_n(s)} \mathrm{Im}\big(v_n^{\mathrm{H}}(s) G_2^{\mathrm{H}} u_n(s)\big) v_n^{\mathrm{H}}(s) G_2^{\mathrm{H}} u_n(s)\right), \tag{2.2}$$

where $u_{\mathrm{d}s}$ *and* $v_{\mathrm{d}s}$ *are given in* (2.3) *and* (2.4).

Before proving Theorem 1, we will prove Lemma 2, which gives formulas for $\dot{u}_n(s)$ and $\dot{v}_n(s)$ in terms of the $u_j(s)$, $v_j(s)$ and $h_u(s) = u_n^{\mathrm{H}}(s)\dot{u}_n(s)$, $h_v(s) = v_n^{\mathrm{H}}(s)\dot{v}_n(s)$.

Lemma 2. *Under the assumptions of Theorem* 1, *the derivatives of* $u_n(s)$ *and* $v_n(s)$ *satisfy*

$$\dot{u}_n(s) = u_{\mathrm{d}s} + h_u(s) u_n(s), \tag{2.3}$$

$$\dot{v}_n(s) = v_{\mathrm{d}s} + h_v(s) v_n(s), \tag{2.4}$$

where $u_{\mathrm{d}s} = -\sum_{j=1}^{n-1} \alpha_j(s) u_j(s)$ *and*

$$v_{\mathrm{d}s} = -\sum_{j=1}^{n-1} \beta_j(s) v_j(s) + \frac{1}{\sigma_n(s)} \sum_{j=n+1}^{p} \big(v_j^{\mathrm{H}}(s) \dot{G}^{\mathrm{H}}(s) u_n(s)\big) v_j(s)$$

with

$$\alpha_j(s) = \frac{\sigma_n(s) u_j^{\mathrm{H}}(s) \dot{G}(s) v_n(s) + \sigma_j(s) v_j^{\mathrm{H}}(s) \dot{G}^{\mathrm{H}}(s) u_n(s)}{\sigma_j^2(s) - \sigma_n^2(s)},$$

$$\beta_j(s) = \frac{\sigma_j(s) u_j^{\mathrm{H}}(s) \dot{G}(s) v_n(s) + \sigma_n(s) v_j^{\mathrm{H}}(s) \dot{G}^{\mathrm{H}}(s) u_n(s)}{\sigma_j^2(s) - \sigma_n^2(s)},$$

and $h_u(s) = u_n^{\mathrm{H}}(s)\dot{u}_n(s)$, $h_v(s) = v_n^{\mathrm{H}}(s)\dot{v}_n(s)$. *The last two functions satisfy*

$$\mathrm{Re}\ h_u(s) = 0, \qquad \mathrm{Re}\ h_v(s) = 0,$$

$$h_u(s) + \bar{h}_v(s) = -\frac{\mathrm{i}}{\sigma_n(s)} \mathrm{Im}\big(v_n^{\mathrm{H}}(s) \dot{G}^{\mathrm{H}}(s) u_n(s)\big). \tag{2.5}$$

Remark. Observe that h_u and h_v contain the derivatives too.

Proof and Lemma 2. In the following proof, we omit the parameter s. So keep in mind that all the mentioned vectors and matrices are functions of s. According to the SVD of G, we have

$$GG^{\mathrm{H}} u_n = \sigma_n^2 u_n.$$

It is well-known that the eigenvalues and the eigenvectors of GG^{H} are analytic with respect to the real parameter s if the eigenvalues are simple [13]. So the derivative of u_n satisfies

$$(GG^{\mathrm{H}} - \sigma_n^2 I)\dot{u}_n = -\dot{G}G^{\mathrm{H}} u_n - G\dot{G}^{\mathrm{H}} u_n + 2\sigma_n \dot{\sigma}_n u_n.$$

Thus from $G = U\Sigma V^{\mathrm{H}}$,

$$(\Sigma\Sigma^{\mathrm{H}} - \sigma_n^2 I) U^{\mathrm{H}} \dot{u}_n = -\sigma_n U^{\mathrm{H}} \dot{G} v_n - \Sigma V^{\mathrm{H}} \dot{G}^{\mathrm{H}} u_n + 2\sigma_n \dot{\sigma}_n e_n,$$

i.e. the first $n-1$ equations in

$$\begin{pmatrix} \sigma_1^2-\sigma_n^2 & & & \\ & \ddots & & \\ & & \sigma_{n-1}^2-\sigma_n^2 & \\ & & & 0 \end{pmatrix} U^{\mathrm{H}}\dot{u}_n = \begin{pmatrix} -\sigma_n u_1^{\mathrm{H}}\dot{G}v_n - \sigma_1 v_1^{H}\dot{G}^{\mathrm{H}}u_n \\ \vdots \\ -\sigma_n u_{n-1}^{\mathrm{H}}\dot{G}v_n - \sigma_{n-1}v_{n-1}^{H}\dot{G}^{\mathrm{H}}u_n \\ 0 \end{pmatrix}.$$

The last equation is just $u_n^{\mathrm{H}}\dot{u}_n = h_u$. Solving for $U^{\mathrm{H}}\dot{u}_n$ gives (2.3) and similarly (2.4) is obtained by applying the same kind of analysis of $G^{\mathrm{H}}Gv_n = \sigma_n^2 v_n$. Now we consider the properties of h_u and h_v. Note that $u_n^{\mathrm{H}}u_n = 1$, so $u_n^{\mathrm{H}}\dot{u}_n + \dot{u}_n^{\mathrm{H}}u_n = 0$, i.e. Re $h_u = 0$ and Re $h_v = 0$. Formula (2.5) follows from the observation that $\sigma_n = v_n^{\mathrm{H}}G^{\mathrm{H}}u_n$ and

$$\dot{\sigma}_n = v_n^{\mathrm{H}}\dot{G}^{\mathrm{H}}u_n + \sigma_n\left(h_u + \bar{h}_v\right). \tag{2.6}$$

As $\dot{\sigma}_n$ is real, and $h_u + \bar{h}_v$ is purely imaginary, we have

$$h_u + \bar{h}_v = -\frac{\mathrm{i}}{\sigma_n}\mathrm{Im}\left(v_n^{\mathrm{H}}\dot{G}^{\mathrm{H}}u_n\right). \qquad \square \tag{2.7}$$

Proof of Theorem 1. From Lemma 2, it is easy to prove the conclusions of Theorem 1. As a direct consequence of (2.6), we obtain (2.1) as $\dot{G} = G_2$. Now we differentiate (2.1). Replacing $\dot{u}_n$ and $\dot{v}_n$ by (2.3) (2.4), we see that the unknown terms h_u and h_v appear only in the form $h_u + \bar{h}_v$ and can be replaced by (2.7). Thus we get (2.2). □

Let us come back to our original problem. We consider first the real case $G(x) = [A - xI,\ B]$ and $\sigma_n(x)$, $u_n(x)$ and $v_n(x)$ are all real. It is easy to see from Lemma 2 that both $h_u(x)$ and $h_v(x)$ vanish. In view of $\dot{G}(x) = -[I,\ 0]$, Theorem 1 gives now Corollary 3.

Corollary 3. *Let x be a real parameter and $[A - xI,\ B] = U(x)\Sigma(x)V^{\mathrm{T}}(x)$ be the* SVD *of $[A - xI,\ B]$ with $\sigma_n(x) = \sigma(x)$ being simple and $f(x) = v_n^{\mathrm{T}}(x)\binom{u_n(x)}{0}$, then the first and second derivatives of $\sigma(x)$ are given by*

$$\frac{\mathrm{d}\sigma}{\mathrm{d}x} = -f(x) \tag{2.8}$$

and

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}x^2} = -\dot{f}(x) = -\left(\dot{v}_n^{\mathrm{T}}(x)\begin{pmatrix} u_n(x) \\ 0 \end{pmatrix} + v_n^{\mathrm{T}}(x)\begin{pmatrix} \dot{u}_n^{\mathrm{T}}(x) \\ 0 \end{pmatrix}\right). \tag{2.9}$$

Here

$$\dot{u}_n(x) = \sum_{j=1}^{n-1}\alpha_j(x)u_j(x), \qquad \dot{v}_n(x) = \sum_{j=1}^{n-1}\beta_j(x)v_j(x) - \frac{1}{\sigma_n(x)}\sum_{j=n+1}^{n+m}\left(v_j^{\mathrm{T}}(x)\begin{pmatrix} u_n(s) \\ 0 \end{pmatrix}\right)v_j(x),$$

where

$$\alpha_j(x) = \frac{\sigma_n(x)\left[u_j^{\mathrm{T}}(x),\ 0\right]v_n(x) + \sigma_j(x)v_j^{\mathrm{T}}(x)\begin{pmatrix} u_n(x) \\ 0 \end{pmatrix}}{\sigma_j^2(x) - \sigma_n^2(x)},$$

$$\beta_j(x) = \frac{\sigma_j(x)\left[u_j^{\mathrm{T}}(x),\ 0\right]v_n(x) + \sigma_n(x)v_j^{\mathrm{T}}(x)\begin{pmatrix} u_n(x) \\ 0 \end{pmatrix}}{\sigma_j^2(x) - \sigma_n^2(x)}.$$

Now we consider the case of the complex parameter $s = x + \mathrm{i}y$. Note that $G(x, y) = [A - (x + \mathrm{i}y)I, B]$, $\partial G(x, y)/\partial x = -[I, 0]$ and $\partial G(x, y)/\partial y = -\mathrm{i}[I, 0]$. Substituting these two partial derivatives to G_2 in (2.1) and (2.2), Theorem 1 gives the formulas of the partial derivatives of $\sigma(x, y)$.

Corollary 4. *Let* $[A - (x + \mathrm{i}y)I, B] = U(x, y)\Sigma(x, y)V(x, y)^{\mathrm{H}}$ *be the* SVD *of* $[A - (x + \mathrm{i}y)I, B]$ *with* $\sigma_n(x, y) = \sigma(x, y)$ *being simple and* $f(x, y) = f(s)$ *is defined in* (1.7), *then we have*

$$\frac{\partial \sigma}{\partial x} = -\mathrm{Re}\, f, \qquad \frac{\partial \sigma}{\partial y} = -\mathrm{Im}\, f, \tag{2.10}$$

$$\frac{\partial^2 \sigma}{\partial x^2} = -\mathrm{Re}\frac{\partial f}{\partial x}, \qquad \frac{\partial^2 \sigma}{\partial x \partial y} = -\mathrm{Re}\frac{\partial f}{\partial y}, \qquad \frac{\partial^2 \sigma}{\partial y^2} = -\mathrm{Im}\frac{\partial f}{\partial y}. \tag{2.11}$$

Here

$$\frac{\partial f}{\partial x} = v_{\mathrm{d}x}^{\mathrm{H}}\begin{pmatrix} u_n \\ 0 \end{pmatrix} + v_n^{\mathrm{H}}\begin{pmatrix} u_{\mathrm{d}x} \\ 0 \end{pmatrix} + \frac{\mathrm{i}}{\sigma_n}(\mathrm{Im}\, f)f, \tag{2.12}$$

where $u_{\mathrm{d}x}$ *and* $v_{\mathrm{d}}x$ *are given in* (2.14) *and* (2.15).

$$\frac{\partial f}{\partial y} = v_{\mathrm{d}y}^{\mathrm{H}}\begin{pmatrix} u_n \\ 0 \end{pmatrix} + v_n^{\mathrm{H}}\begin{pmatrix} u_{\mathrm{d}y} \\ 0 \end{pmatrix} - \frac{\mathrm{i}}{\sigma_n}(\mathrm{Re}\, f)f, \tag{2.13}$$

where $u_{\mathrm{d}y}$ *and* $v_{\mathrm{d}y}$ *are given in* (2.16) *and* (2.17).

$$u_{\mathrm{d}x} = \sum_{j=1}^{n-1} \alpha_{xj} u_j, \tag{2.14}$$

$$v_{\mathrm{d}x} = \sum_{j=1}^{n-1} \beta_{xj} v_j - \frac{1}{\sigma_n} \sum_{j=n+1}^{n+m} \left(v_j^{\mathrm{H}}\begin{pmatrix} u_n \\ 0 \end{pmatrix} \right) v_j, \tag{2.15}$$

where

$$\alpha_{xj} = \frac{\sigma_n\left[u_j^{\mathrm{H}}, 0\right] v_n + \sigma_j v_j^{\mathrm{H}}\begin{pmatrix} u_n \\ 0 \end{pmatrix}}{\sigma_j^2 - \sigma_n^2}, \qquad \beta_{xj} = \frac{\sigma_j\left[u_j^{\mathrm{H}}, 0\right] v_n + \sigma_n v_j^{\mathrm{H}}\begin{pmatrix} u_n \\ 0 \end{pmatrix}}{\sigma_j^2 - \sigma_n^2}.$$

$$u_{\mathrm{d}y} = \mathrm{i} \sum_{j=1}^{n-1} \alpha_{yj} u_j, \tag{2.16}$$

$$v_{\mathrm{d}y} = \mathrm{i} \sum_{j=1}^{n-1} \beta_{yj} v_j + \frac{\mathrm{i}}{\sigma_n} \sum_{j=n+1}^{n+m} \left(v_j^{\mathrm{H}}\begin{pmatrix} u_n \\ 0 \end{pmatrix} \right) v_j, \tag{2.17}$$

where

$$\alpha_{yj} = \frac{\sigma_n\left[u_j^{\mathrm{H}}, 0\right] v_n - \sigma_j v_j^{\mathrm{H}}\begin{pmatrix} u_n \\ 0 \end{pmatrix}}{\sigma_j^2 - \sigma_n^2}, \qquad \beta_{yj} = \frac{\sigma_j\left[u_j^{\mathrm{H}}, 0\right] v_n - \sigma_n v_j^{\mathrm{H}}\begin{pmatrix} u_n \\ 0 \end{pmatrix}}{\sigma_j^2 - \sigma_n^2}.$$

3. The local minimum of $\sigma(x, y)$

From the nice relation between grad $\sigma(x, y)$ and $f(x, y)$ (2.10), we conclude the following result.

Theorem 5. *$s^*=x^*+\mathrm{i}y^*$ is a zero point of $f(s)$ defined in (1.7) iff (x^*, y^*) is a critical point of $\sigma(x, y)=\sigma_n[A-(x+\mathrm{i}y)I, B]$.*

From this theorem, the computation of $d(A, B)$ is equivalent to find the zero points of $f(x, y)$, in which $-(\mathrm{Re}\, f(x, y), \mathrm{Im}\, f(x, y))^{\mathrm{T}}$ will be the gradient of $\sigma(x, y)$. The critical points of $\sigma(x, y)$ are divided into three groups, local minima, local maxima and saddle points.

Let $f(s^*)=0$, $s^*=x^*+\mathrm{i}y^*$. The following is well known.

(a) If $(\partial^2\sigma/\partial x^2)(\partial^2\sigma/\partial y^2)-(\partial^2\sigma/\partial x\partial y)^2>0$ and $\partial^2\sigma/\partial x^2<0$, then (x^*, y^*) is a local maximum of $\sigma(x, y)$.

(b) If $(\partial^2\sigma/\partial x^2)(\partial^2\sigma/\partial y^2)-(\partial^2\sigma/\partial x\partial y)^2>0$ and $\partial^2\sigma/\partial x^2>0$, then (x^*, y^*) is a local minimum of $\sigma(x, y)$.

(c) If $(\partial^2\sigma/\partial x^2)(\partial^2\sigma/\partial y^2)-(\partial^2\sigma/\partial x\partial y)^2<0$, then (x^*, y^*) is a saddle point of $\sigma(x, y)$.

Using the results of Corollary 4, we can now decide to which group the critical points s^* belongs. We can also give a sufficient condition of (x^*, y^*) being a local minimum of $\sigma(x, y)$ in matrix theoretic terms.

Theorem 6. *If $s^*=x^*+\mathrm{i}y^*$ is a zero point of $f(s)$, $u_n^*=u_n(x^*, y^*)$ and*

$$\sigma_{n-1}^2[A-s^*I, B]-\sigma_n^2[A-s^*I, B]>4\left\|u_n^{*\mathrm{H}}(A-s^*I)\right\|^2, \tag{3.1}$$

then (x^*, y^*) is a local minimum of $\sigma(x, y)$.

Proof. Let $s=x+\mathrm{i}y=s^*+\delta$ be a point near to s^* and $q=u_n(x, y)$ be the n-th left singular vector of $[A-sI, B]$ corresponding to $\sigma_n(x, y)$. Then we have

$$\begin{aligned}
&\left\|q^{\mathrm{H}}[A-q^{\mathrm{H}}AqI, B]\right\|^2\\
&\quad=q^{\mathrm{H}}[A-s^*I, B][A-s^*I, B]^{\mathrm{H}}q\\
&\qquad+q^{\mathrm{H}}\left([A-q^{\mathrm{H}}AqI, B][A-q^{\mathrm{H}}AqI, B]-[A-s^*I, B][A-s^*I, B]^{\mathrm{H}}\right)q\\
&\quad=q^{\mathrm{H}}[A-s^*I, B][A-s^*I, B]^{\mathrm{H}}q-|q^{\mathrm{H}}(A-s^*I)q|^2.
\end{aligned}$$

Write $q=au_n^*+bh$ with $h^{\mathrm{H}}u_n^*=0$, $\|h\|=1$, $|a|^2+|b|^2=1$. Thus according to perturbation theory for eigenvectors using the simplicity of σ_n, one has $|b|=\mathrm{O}(\delta)$. Note that

$$\begin{aligned}
q^{\mathrm{H}}[A-s^*I, B][A-s^*I, B]^{\mathrm{H}}q&=|a|^2\sigma_n^2+|b|^2h^{\mathrm{H}}[A-s^*I, B][A-s^*I, B]^{\mathrm{H}}h\\
&=\sigma_n^2+|b|^2\left(h^{\mathrm{H}}[A-s^*I, B][A-s^*I, B]h-\sigma_n^2\right)
\end{aligned}$$

and

$$\begin{aligned}
|q^{\mathrm{H}}(A-s^*I)q|^2&=\left|\bar{b}ah^{\mathrm{H}}(A-s^*I)u_n^*+\bar{a}bu_n^{*\mathrm{H}}(A-s^*I)h\right|^2+\mathrm{O}(\delta^3)\\
&=\left|2\,\mathrm{Re}\,\bar{a}bu_n^{*\mathrm{H}}(A-s^*I)h\right|^2+\mathrm{O}(\delta^3).
\end{aligned}$$

Thus

$$\begin{aligned}
&\left\|q^{\mathrm{H}}[A-q^{\mathrm{H}}AqI, B]\right\|^2-\sigma_n^2(x^*, y^*)\\
&\quad=|b|^2\left(h^{\mathrm{H}}[A-s^*I, B][A-s^*I, B]^{\mathrm{H}}h-\sigma_n^2\right)-\left|2\,\mathrm{Re}\,\bar{a}bu_n^{*H}(A-s^*I)h\right|^2+\mathrm{O}(\delta^3)\\
&\quad\geq|b|^2\left(\sigma_{n-1}^2-\sigma_n^2\right)-4|a|^2|b|^2\left\|u_n^{*\mathrm{H}}(A-s^*I)\right\|^2+\mathrm{O}(\delta^3)\\
&\quad=|b|^2\left(\sigma_{n-1}^2-\sigma_n^2-4\left\|u_n^{*H}(A-s^*I)\right\|^2\right)+\mathrm{O}(\delta^3).
\end{aligned}$$

According to the condition of (3.1), $\| q^{\mathrm{H}}[A - q^{\mathrm{H}}AqI,\ B] \|^2 \geq \sigma_n^2(x^*,\ y^*)$ is always true. Since

$$\sigma_n[A - (x + \mathrm{i}y)I,\ B] = \| q^{\mathrm{H}}[A - (x + \mathrm{i}y)I,\ B] \| \geq \| q^{\mathrm{H}}[A - q^{\mathrm{H}}AqI,\ B] \|,$$

$\sigma_n(x,\ y) \geq \sigma_n(x^*,\ y^*)$. Thus $(x^*,\ y^*)$ is a local minimum of $\sigma_n(x,\ y)$. □

Since

$$\| u_n^{\mathrm{H}}(A - s^*I) \| = \left\| u_n^{\mathrm{H}}[A - s^*I,\ B]\begin{pmatrix} I \\ 0 \end{pmatrix} \right\| \leq \sigma_n(x^*,\ y^*),$$

we get at once the following sufficient condition.

Corollary 7. *If $s^* = x^* + \mathrm{i}y^*$ satisfies $f(s^*) = 0$ and*

$$\sigma_{n-1}[A - s^*I,\ B] > \sqrt{5}\,\sigma_n[A - s^*I,\ B], \tag{3.2}$$

then $(x^*,\ y^*)$ is a local minimum point of $\sigma(x,\ y)$.

4. Newton's algorithm

Because we have obtained the first and second partial derivatives of $\sigma(x,\ y)$ in terms of the SVD of $[A - sI,\ B]$ when $\sigma(x,\ y)$ is simple, Newton's method can be applied to compute the minimum points of $\sigma(x,\ y)$. Generally speaking, the local minima of $\sigma(x,\ y)$ happen when they are simple. More details are discussed in Section 6. As $u_n^{*\mathrm{H}}Au_n^* = s^*$, all minimum points $s^* = x^* + \mathrm{i}y^*$ lie in the field of values of A, and hence

$$\lambda_{\min}\left(\frac{A + A^{\mathrm{T}}}{2}\right) \leq x^* \leq \lambda_{\max}\left(\frac{A + A^{\mathrm{T}}}{2}\right), \qquad \lambda_{\min}\left(\frac{A - A^{\mathrm{T}}}{2\mathrm{i}}\right) \leq y^* \leq \lambda_{\max}\left(\frac{A - A^{\mathrm{T}}}{2\mathrm{i}}\right). \tag{4.1}$$

Here $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ denote the minimal and the maximal eigenvalue of A. Since $\sigma_n[A - s^*I,\ B] = \sigma_n[A - \bar{s}^*I,\ B]$, the search for minimum points can be restricted to

$$0 \leq y^* \leq \lambda_{\max}\left(\frac{A - A^{\mathrm{T}}}{2\mathrm{i}}\right).$$

Theorem 5 also suggests a method to compute $d(A,\ B)$. We need only to find all zeros of $f(s)$, which are the critical points of $\sigma(s)$. Especially in the case of $f(s)$ being a real function of a real parameter s in order to compute the following $d_{\mathrm{r}}(A,\ B)$, the bisection method can be used to find all zeros of $f(x)$. We ought to say that Theorem 6 and Corollary 7 and the criterions of the second partial derivatives give only sufficient conditions to determine which zeros of $f(s)$ are local minima of $\sigma(s)$.

4.1. Real case

We first consider the problem of computing

$$d_{\mathrm{r}}(A,\ B) = \min_{s \in R} \sigma_n[A - sI,\ B].$$

So $f(s)$ will be real. Since $s^* = u_n^{\mathrm{T}}(s^*)Au_n(s^*)$, s^* is in the interval

$$I_{\mathrm{r}} = \left[\lambda_{\min}\left(\frac{A + A^{\mathrm{T}}}{2}\right),\ \lambda_{\max}\left(\frac{A + A^{\mathrm{T}}}{2}\right)\right]. \tag{4.2}$$

Also since $u_n^{\mathrm{T}}(s)Au_n(s) - s = \sigma_n(s)f(s)$, we have

$$f(s) > 0 \quad \text{for } s < \lambda_{\min}\left(\frac{A + A^{\mathrm{T}}}{2}\right), \qquad f(s) < 0 \quad \text{for } s > \lambda_{\max}\left(\frac{A + A^{\mathrm{T}}}{2}\right). \tag{4.3}$$

The following Newton method is suggested to compute the minimum points of $\sigma(s) = \sigma_n(s)$.

Newton's algorithm (real case). Choose $s_0 \in I_r$. For $k = 1, 2, \ldots,$

$$s_{k+1} = s_k - \theta_k \frac{f(s_k)}{\dot{f}(s_k)}$$

where θ_k is such that $\sigma(s_{k+1}) < \sigma(s_k)$.

In our examples a choice $\theta_k \neq 1$ is only neccessary at the beginning steps of the Newton algorithm. After having a good approximation of a local minimum point, we can take $\theta_k = 1$ and hence have the usual Newton algorithm. Also the following bisection method can be used to find the zeros of $f(s)$.

Bisection method. (a) Find an interval $[a, b]$ such that $f(a) * f(b) < 0$.

(b) Let $c = \frac{1}{2}(a + b)$, if $f(c) * f(b) < 0$ then $a = c$ and go back (b) and if $f(a) * f(c) < 0$ then $b = c$ and go back (b). The step (b) is repeated until c is an acceptable zero point of $f(s)$.

4.2. Complex case

In order to compute $d(A, B) = \min_{s \in C} \sigma_n([A - sI, B])$, we have the following Newton algorithm for complex s.

Newton's algorithm (complex case). Choose $\binom{x_0}{y_0}$. For $k = 1, 2, \ldots,$

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ y_k \end{pmatrix} - \theta_k \begin{pmatrix} p_{k1} \\ p_{k2} \end{pmatrix},$$

where

$$\begin{pmatrix} p_{k1} \\ \\ p_{k2} \end{pmatrix} = \begin{pmatrix} \mathrm{Re}\,\dfrac{\partial f}{\partial x} & \mathrm{Re}\,\dfrac{\partial f}{\partial y} \\ \mathrm{Im}\,\dfrac{\partial f}{\partial x} & \mathrm{Im}\,\dfrac{\partial f}{\partial y} \end{pmatrix}^{-1} \begin{pmatrix} \mathrm{Re}\ f(x, y) \\ \\ \mathrm{Im}\ f(x, y) \end{pmatrix},$$

and θ_k such that

$$\sigma(x_k - \theta_k p_{k1}, y_k - \theta_k p_{k2}) = \min_{-1 \leq \theta \leq 1} \sigma(x_k - \theta p_{k1}, y_k - \theta p_{k2}).$$

Computing the minimum of $\sigma(x_k - \theta p_{k1}, y_k - \theta p_{k2})$ in $[-1, 1]$ is as easy as that of $\sigma(x)$ for real x. Let $u_n(\theta)$, $v_n(\theta)$ be the left, right singular vectors of $[A - (x_k - \theta p_{k1} + y_k \mathrm{i} - \theta p_{k2}\mathrm{i})I, B]$ corresponding to $\sigma(x_k - \theta p_{k1}, y_k - \theta p_{k2})$, then $f(\theta) = v_n^{\mathrm{H}}(\theta)\binom{u_n(\theta)}{0}$ is well defined. We introduce

$$g(\theta) = \frac{\mathrm{d}\sigma}{\mathrm{d}\theta}(x_k - \theta p_{k1}, y_k - \theta p_{k2}),$$

then $g(\theta)$ has the following expression by (2.10),

$$g(\theta) = p_{k1}\ \mathrm{Re}\ f(\theta) + p_{k2}\ \mathrm{Im}\ f(\theta).$$

As a direct consequence of Theorem 1, $\dot{g}(\theta)$ is given by

$$\dot{g}(\theta) = p_{k1} \text{ Re } \dot{f}(\theta) + p_{k2} \text{ Im } \dot{f}(\theta).$$

All together we have the following algorithm to calculate θ_k:

Newton Algorithm to compute θ_k. (a) Initial value $\theta_0 = 1$.
(b) Run the following Newton method for $j = 1, 2, \ldots$:

$$\theta_{j+1} = \theta_j - \eta_j \frac{g(\theta_j)}{\dot{g}(\theta_j)},$$

where η_j is chosen such that $\sigma(x_k - \theta_{j+1} p_{k1}, y_k - \theta_{j+1} p_{k2}) < \sigma(x_k - \theta_j p_{k1}, y_k - \theta_j p_{k2})$.

Also the bisection method can be used to find the zeros of $g(\theta)$. Numerical results suggest that this Newton's method with the parameter θ_k enjoys the property of global convergence. Moreover one needs only to compute two or three θ_k's to get a good initial point for the Newton method. It means that after two or three steps θ_k will be near to 1. Thus Newton's method with $\theta_k = 1$ will converge quadratically. Hence θ_k is only calculated in the first three steps, it is automatically taken to be 1 since then. Computing the minimum θ_k takes much work. One needs generally seven or eight SVDs to find a good approximate value to θ_k. However it seems that this step cannot be neglected. It is worthwhile to say that there exists only one zero point of $g(\theta)$ in our examples.

Another way of selecting θ_k is from the following inequality:

$$\| u_n^{\mathrm{H}}(0)[A - (x_k + \mathrm{i} y_k)I, B] \| \ge \| u_n^{\mathrm{H}}(0)[A - (x_k - \theta_k p_{k1} + y_k \mathrm{i} - \theta_k p_{k2} \mathrm{i})I, B] \|,$$

where $\theta_k = -\{\sigma(x_k, y_k)(p_{k1} \text{ Re } f(0) + p_{k2} \text{ Im } f(0))\}/(p_{k1}^2 + p_{k2}^2)$. From this inequality, we have

$$\sigma(x_k, y_k) \ge \sigma(x_k - \theta_k p_{k1}, y_k - \theta_k p_{k2}).$$

But Newton's method with θ_k selected in this way converges to s^* very slowly.

Before we finish this section, we shall discuss how to select a good initial point for the Newton method. In [3], it has been proven that s^* is located within one of the disks in the complex plane whose centers are the eigenvalues of F, where

$$F = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{4.4}$$

with $[C, D]$ a random matrix or having orthogonal rows such that F is square. Our numerical examples show that those disks are small and almost located in the region (4.1). So the eigenvalues of F and A are generally a good choice of the initial points. For our examples, Newton's algorithm with the initial points being the eigenvalues of F without selecting the parameters θ_k (i.e. $\theta_k = 1$), converges to the local minima of $\sigma_n(s)$ within 5 steps.

5. Numerical examples

Two examples presented in [12] are implemented under MATLAB. Using Newton's method and the bisection method, all minimum points of $\sigma(s)$ with s being real are found. So there exists no difficulty to get $d_{\mathrm{r}}(A, B)$. But for $d(A, B)$, though we have known the exact region containing all zero points of $f(s)$, the number of zero points is still a problem. Generally speaking, the Newton's method converges very quickly if a good initial point is selected.

Example 1. Consider

$$A = \begin{pmatrix} 3.28 & -2.44 & -1.54 & -3.20 & -3.34 \\ -1.58 & -1.02 & 3.86 & 4.15 & 3.94 \\ -4.06 & 3.54 & 1.65 & 1.79 & 2.15 \\ -4.15 & 3.96 & 0.84 & -2.70 & -2.70 \\ -1.76 & 0.29 & -1.14 & -1.64 & -2.21 \end{pmatrix}, \qquad B = \begin{pmatrix} -2.80 \\ 2.79 \\ 1.88 \\ -0.48 \\ -1.89 \end{pmatrix}.$$

For this example, we compute $d_r(A, B)$, so $f(s)$ will be real. All minimal points of $\sigma(s)$ are according to (4.2) in the region

$$(-8.5123, 9.7310).$$

The graphs of $f(s)$ and $\sigma(s)$ are shown in Figure 1. One can see that the minima and maxima of $\sigma(s)$ are interlacing. $f(s)$ has seven zero points including four minima. When $s^* = 0.431388$, $\sigma = 0.231910$ is the minimum value. So s^* will minimize $d_r(A, B) = \min_{s \in R} \sigma(s)$ and $d_r(A, B) = 0.231910$. In [12], only two zero points of $f(s)$ are found. Neither of them reaches the value of $d_r(A, B)$. When taking any point in the interval $(-8.5123, 9.7310)$ as an initial point, the Newton algorithm (real case) converges to a minimum point of $\sigma(s)$ within 5 steps. $|s_k - s^*| \le 10^{-6}$, where s_k is the acceptable iterative value of the Newton's method. When the initial values are taken as the eigenvalues of F defined in (4.4), the Newton method converges to the local minimum points within 5 steps without selecting θ_k ($\theta_k = 1$).

Example 2. Consider

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0.1 & 3 & 5 \\ 0 & -1 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 \\ 0.1 \\ 0 \end{pmatrix}.$$

All zero points $s^* = x^* + iy^*$ are in the rectangular region given by

$$-1.851295 \le x \le 3.992519, \qquad -3.074491 \le y \le 3.074491.$$

$f(s)$ has only one real zero point $s_r^* = 1.027337$ and $d_r(A, B) = 0.172460$. The minimum point $s^* = 0.937084 + 0.998571i$ minimizes $\min_{s \in C} \sigma(s)$ and $d(A, B) = 0.039238$. We tried several initial points, the

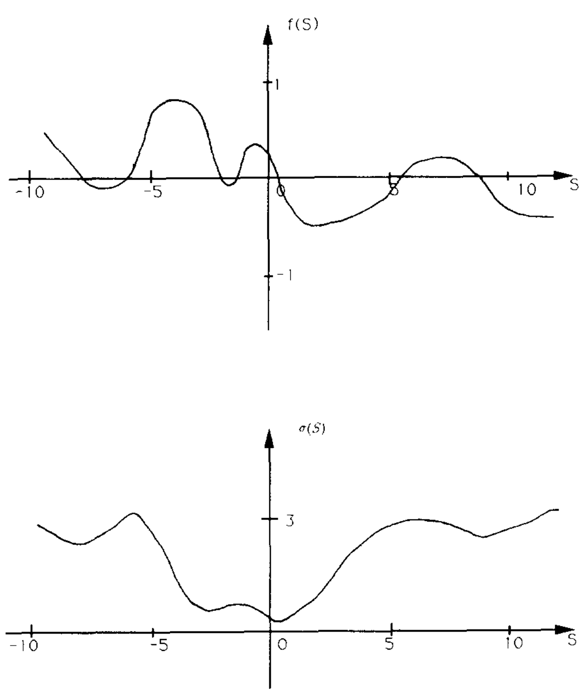

Fig. 1.

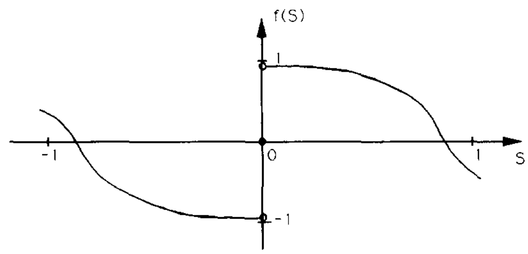

Fig. 2.

Newton method with each of them converges to s^* in 5 steps. Say $s_0 = 1.5 + i$, we found $\theta_0 = 0.09935$, $\theta_1 = 0.564158$ and $\theta_2 = 1.00126$. So from step 3 on, $\theta_k = 1$ is selected. However if the initial point is taken as one of the eigenvalues of F in (4.4), say $0.8625 + 0.9749i$, then the Newton's algorithm with $\theta_k = 1$ converges to s^* in 5 steps.

Example 3. Consider

$$A = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

In this example, A is a 10×10 matrix, B a 10×1 matrix. Let $f(s)$ be the real function of the real parameter s, then $f(s)$ has three zero points in real axis. They are 0 and -0.959492, $+0.959492$. It is interesting to know that $f(s)$ is no longer a continuous function. It has a big jump at zero point (see Figure 2).

The reason is that $[A - sI, B]$ has a multiple least singular value 1 at $s = 0$, so its singular vector $u_n(s)$ may not be continuous at zero. In such case, one can change the initial point to run the Newton algorithm again. Fortunately this extreme case never happens at s^*.

6. Multiple singular values and conclusions

Our theorems about the differentiability of the least singular value $\sigma_n(s)$ of $G(s)$, a matrix with $\text{Re}(G(s))$ and $\text{Im}(G(s))$ being real analytic matrix-valued function of a real parameter s, are based on the assumption of $\sigma_n(s)$ being simple. In this section, we discuss the case of $\sigma_n(s)$ being a multiple singular value of $G(s)$. The problem of minimizing the least singular value of $G(s)$ is very different from that of minimizing the largest one of $G(s)$. The solution of the latter is usually at a point where singular values coalesce, i.e. at a nondifferentiable point, since the minimization will drive several singular values to the same minimum value [8]. But for the former problem, its local minimum does not happen at the cross singular values in general. So at its local minimum point, $\sigma_n(s)$ is generally simple and differentiable.

When the minimum of $\sigma_n(s)$ happens at the point s^* (this is extremely unusual), where $\sigma_{n-r+1}(s^*) = \cdots = \sigma_n(s^*)$, $\sigma_n(s)$ is also differentiable at s^*, and its derivative is zero. This comes from the fact that the left and right limits of $\dot{\sigma}_n(s)$ at s^* always exist [11], and they are equal to zero when s^* is a local minimum point (see the graph below). What about the second derivative of $\sigma_n(s)$ at s^*? We claim that the second derivative of $\sigma_n(s)$ at s^* always exists too. Considering $\sigma_j(s)$ and $\sigma_n(s)$, where j is one number of $\{n - r + 1, \ldots, n - 1\}$, we define two new functions $p_1(s)$ and $p_2(s)$ near s^* as the original singular value functions without ordering them, so $p_1(s)$ and $p_2(s)$ are analytic near s^*. The relations of $p_1(s)$, $p_2(s)$ and $\sigma_j(s)$ and $\sigma_n(s)$ are $\sigma_j(s) = \max\{p_1(s), p_2(s)\}$ and $\sigma_n(s) = \min\{p_1(s), p_2(s)\}$. Moreover $p_1(s^*) = p_2(s^*)$ and $\dot{p}_1(s^*) = \dot{p}_2(s^*) = 0$, since $\sigma_j(s^*) = \sigma_n(s^*)$ and $\dot{\sigma}_j(s^*) = \dot{\sigma}_n(s^*) = 0$. Let us assume

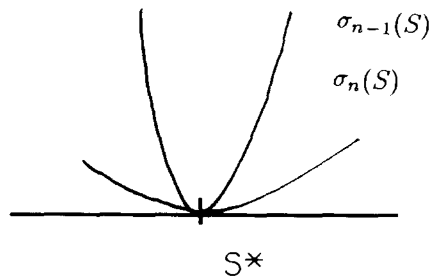

Fig. 3.

that $(d^2p_1/ds^2)(s^*) \geq (d^2p_2/ds^2)(s^*)$. When $(d^2p_1/ds^2)(s^*) > (d^2p_2/ds^2)(s^*)$, we have $p_1(s) - p_2(s) \geq 0$ near to s^*. By the definitions of $p_1(s)$ and $p_2(s)$, we know that $p_1(s) = \sigma_j(s)$ and $p_2(s) = \sigma_n(s)$, so $\sigma_n(s)$ is analytic. When $(d^2p_1/ds^2)(s^*) = (d^2p_2/ds^2)(s^*)$, we know that the second derivative of $\sigma_n(s)$ exists near s^*. Hence the Newton method can be used at the minimum points of $\sigma_n(s^*)$. See Figure 3.

We have presented a new method to compute the distance to uncontrollability $d(A, B)$, which is based on the explicit expressions of the first and second derivatives of $\sigma(x, y)$. Numerical examples show that this method works well.

References

[1] D.L. Boley and W. Lu, Measuring how far a controllable system is from an uncontrollable one, *IEEE Trans. Automat. Control* **31** (1986) 249–251.
[2] D.L. Boley, Computing rank-deficiency of rectangular matrix pencials, *Systems Control Lett.* **9** (1987) 207–214.
[3] D.L. Boley, Estimating the sensitivity of the algebraic structure of pencils with simple eigenvalue estimates, *SIAM J. Matrix Anal. Appl.* (Oct. 1990).
[4] R. Byers, Detecting nearly uncontrollable pairs, *Proceedings of the International Symposium MTNS-89* (Amsterdam, 1989).
[5] R. Eising, The distance between a system and the set of uncontrollable systems, *Proc. MTNS*, Beer-sheva, June 1983 (Springer-Verlag, Berlin–New York, 1984) 303–314.
[6] R. Eising, Between controllable and uncontrollable, *Systems Control Lett.* **4** (1984) 263–264.
[7] C. Kenney and A.J. Laub, Controllability and stability radii for companion form systems, *Math. Control Signals Systems* (1988).
[8] M.L. Overton, On minimizing the maximum eigenvalue of a symmetric matrix, *SIAM J. Matrix Anal. Appl.* **9** (1988) 256–268.
[9] C.C. Paige, Properties of numerical algorithms relating to controllability, *IEEE Trans. Automat. Control* **26** (1981) 130–138.
[10] Ji-guang Sun, A note on simple non-zero singular values, *J. Comput. Math.* **6** (1988) 259–266.
[11] Ji-guang Sun, Sensitivity analysis of zero singular values and multiple singular values *J. Comput. Math.* **6** (1988) 325–335.
[12] M. Wicks and R. DeCarlo, Computing the distance to uncontrollable system, *IEEE Trans. Automat. Control* (to appear).
[13] J.H. Wilkinson, *The Algebraic Eigenvalue Problem* (Clarendon Press, Oxford, 1965).