On the Nonlinearly Structured Stability Radius Problem

Wei-Yong Yan*   James Lam[†§]

* School of Electrical and Electronic Engineering
Nanyang Technological University
Nanyang Avenue, Singapore 639798

[†] Department of Mechanical Engineering
University of Hong Kong
Pokfulam Road, Hong Kong

Abstract

This paper considers the problem of finding a perturbation matrix with the least spectral norm such that a matrix-valued function becomes singular, where the dependence of the function on the perturbation is allowed to be nonlinear. It is proved that such a problem can be approximated by a smooth unconstrained minimization problem with compact sublevel sets. A computational procedure proposed based on this result is demonstrated to be effective in both linear and nonlinear cases.

1 Introduction

Many control systems are subject to perturbations in terms of uncertain parameters. An important quantitative measure of stability robustness of a system against such perturbations is what is called the real or complex stability radius, depending on the nature of perturbations in concern. The computation of a stability radius is a subject which has attracted a lot of interest over recent decades, see e.g. [1, 2, 3, 4, 5, 6] and references therein.

Given a stable matrix A, the general stability radius problem is to determine a real or complex perturbation matrix Δ with a minimum spectral norm such that the perturbed matrix $A + S(\Delta)$ becomes unstable, where $S(\Delta)$ is a mapping describing the structure of the perturbation. It is well known that the basic subproblem of this problem is to determine a perturbation matrix Δ with a minimum spectral norm such that $A - j\omega + S(\Delta)$ becomes singular at each given frequency ω. This leads to the general linear algebra problem of determining a Δ with a minimum spectral norm such that $R(\Delta)$ becomes singular for a given mapping $R(\Delta)$.

Although the special case where $R(\Delta)$ is of the form $I - \Delta M$ has been successfully and completely treated recently by Qiu et al. in [5], the problem has yet to be

solved in the general case, including such a simple case where $R(\Delta)$ is of the form $I + G_1 \Delta H_1 + G_2 \Delta H_2$. It should also be mentioned that a nonlinear $R(\Delta)$ naturally arises when the dependence of a system matrix on uncertain parameters is nonlinear.

2 Problem Formulation

Throughout this section, it is assumed that $R(\Delta)$: $\mathbb{R}^{m \times p} \mapsto \mathbb{C}^{n \times n}$ is a continuous mapping with

$$h(\Delta) \triangleq \det(R(\Delta)R^*(\Delta)) \qquad (2.1)$$

The problem to be considered in this paper can be precisely stated as follows:

Problem P:

$$\text{minimize } J(\Delta) \triangleq \|\Delta\|$$

$$\text{over } \Sigma \triangleq \{\Delta \in \mathbb{R}^{m \times p} \mid h(\Delta) = 0\}$$

This problem is obviously very general as the mapping $R(\Delta)$ is allowed to take various forms. A typical linear form of $R(\Delta)$ could be the following

$$R(\Delta) = I + \sum_{i=1}^{N} G_i \Delta H_i$$

In the special case where $R(\Delta) = I - \Delta M$, the problem has been tackled in [5] by converting it into the problem of minimizing the second largest singular value of a certain matrix defined on the interval $(0, 1]$. Specifically, it has been shown there that

$$\min\{\|\Delta\| \mid \det(I - \Delta M) = 0\}$$

$$= \frac{1}{\inf_{\gamma \in (0,\,1]} \sigma_2 \left(\begin{bmatrix} \operatorname{Re} M & -\gamma \operatorname{Im} M \\ \gamma^{-1} \operatorname{Im} M & \operatorname{Re} M \end{bmatrix} \right)}$$

where $\sigma_2(\cdot)$ denotes the second largest singular value. A procedure has also been proposed for computing a worst real perturbation Δ after a minimum is found.

3622

One nonlinear example of $R(\Delta)$ is given by

$$R(\Delta) = I + \sum_{i=1}^{N} (G_i \Delta H_i)^2$$

although more complicated forms can easily be envisaged.

It should be emphasized that there is no loss of generality in restricting Δ to be real in the above problem formulation. This is because a complex Δ can be represented by a real matrix of larger size formed by the real and imaginary parts of Δ. As such, Problem P is closely related to the general complex stability radius problem as well.

Quite obviously, Problem P is a *nonsmooth* and constrained minimization problem, which is generally difficult to solve. On the other hand, a global minimum for the problem does exist unless the set Σ is empty. Our main result to be presented below shows that Problem P can be arbitrarily approximated by a *smooth* and *unconstrained* minimization problem.

Theorem 2.1 *Suppose that $R(\Delta)$ is continuous in $\mathbb{R}^{m \times p}$ with Σ nonempty. Let a function $J_k(\Delta)$: $\mathbb{R}^{m \times p} \mapsto \mathbb{R}$ be defined by*

$$J_k(\Delta) \triangleq f_k(\Delta) + kh(\Delta) \qquad (2.2)$$

with

$$f_k(\Delta) \triangleq \left\{ \text{trace} \left[(\Delta^T \Delta)^k \right] \right\}^{\frac{1}{2k}} \qquad (2.3)$$

If J^* is the global minimum for Problem P, then there holds

$$J^* = \lim_{k \to \infty} \min J_k(\Delta)$$

Proof.
It is known that for all $k \geq 1$ and $\Delta \in \mathbb{R}^{m \times p}$,

$$\lim_{k \to \infty} f_k(\Delta) = J(\Delta) \quad \text{and} \quad f_k(\Delta) \geq J(\Delta) \qquad (2.4)$$

Now with Δ^* denoting an optimal solution to Problem P and $J_k^* = \min J_k(\Delta)$, it is apparent that

$$J_k^* \leq f_k(\Delta^*) \qquad (2.5)$$

Next, let $\Delta_k^* \in \mathbb{R}^{m \times p}$ be such that $J_k^* = J_k(\Delta_k^*)$ and let Δ_k be a global minimum point of $J(\Delta)$ over the set

$$\Sigma_k \triangleq \left\{ \Delta \in \mathbb{R}^{m \times p} \mid h(\Delta) \leq \alpha/k \right\}$$

where α is such a constant that

$$f_k(\Delta^*) \leq \alpha, \quad \forall k \geq 1 \qquad (2.6)$$

Combining (2.5) with (2.6) yields

$$h(\Delta_k^*) \leq J_k^*/k \leq \alpha/k$$

which implies that $\Delta_k^* \in \Sigma_k$. Thus, one has

$$J_k^* \geq f_k(\Delta_k^*) \geq J(\Delta_k^*) \geq J(\Delta_k)$$

This together with (2.5) establishes that

$$J(\Delta_k) \leq J_k^* \leq f_k(\Delta^*) \qquad (2.7)$$

Quite obviously, $\{J(\Delta_k)\}$ is a nondecreasing sequence bounded by $J(\Delta^*)$ from above as $\Sigma_{k_1} \supseteq \Sigma_{k_2}$ for $k_1 < k_2$. As a consequence, there exists a convergent subsequence $\{\Delta_{n_k}\}$, whose limit is denoted by Δ_∞. From

$$h(\Delta_{n_k}) \leq \alpha/n_k$$

and the continuity of $h(\Delta)$, it follows that $\Delta_\infty \in \Sigma$, leading to

$$\lim_{k \to \infty} J(\Delta_k) = \lim_{k \to \infty} J(\Delta_{n_k}) = J(\Delta_\infty) \geq J(\Delta^*)$$

Hence, there holds

$$\lim_{k \to \infty} J(\Delta_k) = J(\Delta^*) \qquad (2.8)$$

The theorem is immediately concluded from (2.4), (2.7), and (2.8). □

From the above proof, we immediately conclude the following result, which says that an optimal solution to Problem P can be arbitrarily approximated by a global minimum point of $J_k(\Delta)$ provided k is sufficiently large.

Corollary 2.1 *Adopt the same assumption and notation as in Theorem 2.1. If Δ_k^* is a global minimum point of $J_k(\Delta)$, then there hold*

$$\lim_{k \to \infty} J(\Delta_k^*) = J^* \qquad (2.9)$$

$$\lim_{k \to \infty} h(\Delta_k^*) = 0 \qquad (2.10)$$

$$J(\Delta_k^*) \leq \min J_k(\Delta), \quad \forall k \geq 1 \qquad (2.11)$$

3 A Computational Procedure

In this section, $R(\Delta)$ will be assumed to be smooth, which is usually the case in practical applications. This smoothness will be exploited to suggest an efficient procedure for solving Problem P, whose optimal solution can be approached by minimizing the cost function $J_k(\Delta)$ with sufficiently large k by Theorem 2.1. In two specific cases, we will give a closed-form formula for the gradient of $J_k(\Delta)$.

By straightforward calculations, the gradient of $J_k(\Delta)$ can be found to be

$$\nabla J_k(\Delta) = [f_k(\Delta)]^{1-2k} \Delta \left(\Delta^T \Delta \right)^{k-1} + k \nabla h(\Delta)$$

Of course, the formula for the gradient $\nabla h(\Delta)$ cannot be obtained unless $R(\Delta)$ is known. In the following two cases which each cover the case treated in [5]:

Case 1: $R(\Delta) = F + \sum_{i=1}^{N} G_i \Delta H_i$

Case 2: $R(\Delta) = F + \Delta M_1 + \Delta M_2 \Delta^T$

it can be established that

$$\nabla h(\Delta) = \Psi(\Delta) + \bar{\Psi}(\Delta)$$

where

$$\Psi(\Delta) = \sum_{i=1}^{N} G_i^* \mathrm{adj}\big(R(\Delta)R^*(\Delta)\big)R(\Delta)H_i^* \qquad (3.1)$$

for Case 1 and

$$\Psi(\Delta) = \mathrm{adj}\big(R(\Delta)R^*(\Delta)\big)R(\Delta)(M_1^* + \Delta M_2^*)$$
$$+ R^*(\Delta)\mathrm{adj}\big(R(\Delta)R^*(\Delta)\big)\Delta M_2 \qquad (3.2)$$

for Case 2.

The gradient $\nabla J_k(\Delta)$ can be used to form the gradient flow

$$\boxed{\dot{\Delta}(t) = -\left[f_k(\Delta)\right]^{1-2k}\Delta\left(\Delta^T\Delta\right)^{k-1} - k\nabla h(\Delta)} \quad (3.3)$$

which enjoys several nice properties listed below.

- The ODE (3.3) has a unique solution $\Delta(t)$ on the interval $[0, \infty)$ for any initial condition.

- Any solution $\Delta(t)$ to (3.3) converges to a connected set of critical points as t goes to infinity.

- The cost function $J_k(\Delta)$ is nonincreasing along any solution $\Delta(t)$ to (3.3) and is strictly decreasing when $\Delta(0)$ is not a critical point of $J_k(\Delta)$.

These properties, which follow from the compactness of all the sublevel sets of $J_k(\Delta)$, suggest that a minimum of $J_k(\Delta)$ can be found by solving the ODE (3.3). Whether the obtain minimum is local or global may depend on the choice of an initial condition for (3.3). However, the global minimum must result regardless of the choice of a starting point if there is no other minimum.

Certainly, the ODE (3.3) can be easily integrated using an appropriate numerical routine, e.g., in Matlab on a digital computer as only standard matrix operations are involved. In the meantime, analog computing has recently gained renewed interest in view of advances in neural networks which allow massively parallel processing. As a result, it has recently become increasingly acceptable to make use of ODEs for solving various problems such as optimization and linear algebra problems, see e.g. [7, 8] and references therein.

We end this section by proposing the following general procedure for solving Problem P. The convergence of the algorithm is essentially guaranteed by Theorem 2.1.

Step 1 Choose an initial integer k and a point $\Delta_0 \in \mathbb{R}^{m \times p}$.

Step 2 Seek a minimum point $\bar{\Delta}$ of the cost function $J_k(\Delta)$ by finding a limiting solution to the ODE (3.3) with the initial condition $\Delta(0) = \Delta_0$.

Step 3 If both $|\det(R(\bar{\Delta}))|$ and $f_k(\bar{\Delta}) - J(\bar{\Delta})$ are less than a preset value, stop; otherwise, go back to step 2 with a larger k and $\Delta_0 = \bar{\Delta}$.

Remark 3.1 *In the second step, a general-purpose globally convergent search algorithm given in [9] could alternatively be employed to find a minimum point $\bar{\Delta}$ of the cost function $J_k(\Delta)$.*

Remark 3.2 *In order to achieve computational efficiency as well as to avoid numerical problems associated with large k, the gradient of $f_k(\Delta)$ should be computed as follows*

$$\nabla f_k(\Delta) = [f_k(\Delta)]^{1-2k}\Delta\left(\Delta^T\Delta\right)^{k-1}$$
$$= \left[\sum \left(\frac{s_i}{\|\Delta\|}\right)^{2k}\right]^{\frac{1-2k}{2k}} U S_k V^T$$

where Δ is assumed to have the singular value decomposition $\Delta = USV^T$ and S_k results from replacing every diagonal element s_i of S by $\left(\frac{s_i}{\|D\|}\right)^{2k-1}$.

4 Examples

Example 1: Assume that $R(\Delta) = I - \Delta M$ where

$$M = \begin{bmatrix} 2+j & 1 \\ 1 & 2+j \end{bmatrix}$$

As was indicated in [5], this case leads to the nonsmooth behavior of the cost function introduced there, which is defined to be the second largest singular vaule of an augmented matrix associated with M.

Let us implement the algorithm proposed in the last section to find a worst real perturbation which makes $R(\Delta)$ singular. The details of the implementation are described as follows. We use a randomly generated matrix Δ_0 as an initial condition for solving the ODE with $k = 50$. Figure 1 depicts the norm of the solution $\Delta_1(t)$ to the ODE as t goes from 0 to 0.14. By using the final value $\Delta_1(0.14)$ as an initial condition, we continue to solve the ODE with $k = 100$ to obtain the solution $\Delta_2(t)$ on $[0, 0.5]$. The norm $\|\Delta_2(t)\|$ as a function of t is depicted in Figure 2. With $\Delta_2(0.5)$ as a new starting point, the two solutions $\Delta_3(t)$ defined on $[0, 0.007]$ for $k = 1000$ and $\Delta_4(t)$ on $[0, 0.003]$ for $k = 3000$ are obtained by repeating the same process. Figure 3 shows a concatenation of the two functions $\|\Delta_3(t)\|$ and $\|\Delta_4(t)\|$ on the combined interval $[0, 0.01]$. The initial value Δ_0 and the values of the four solutions at the end point of the respective intervals are given in Table 1 along with their norms as well as their corresponding determinants of $R(\Delta)$. As is seen, the last matrix in Table 1 is identical with the worst perturbation obtained in [5]. The

total time taken to complete the whole simulation on an HP workstation is 12 seconds.

Table 1. Properties of initial, intermediate, and final perturbation matrices.

Δ	$\|\Delta\|$	$|\det(R(\Delta))|$
$\begin{bmatrix} 0.5129 & 0.3504 \\ 0.4605 & 0.0950 \end{bmatrix}$	0.7650	1.6394
$\begin{bmatrix} 0.2822 & -0.1649 \\ 0.3385 & 0.2471 \end{bmatrix}$	0.4538	0.0330
$\begin{bmatrix} 0.3321 & -0.2342 \\ 0.2359 & 0.3321 \end{bmatrix}$	0.4077	0.0023
$\begin{bmatrix} 0.3332 & -0.2355 \\ 0.2357 & 0.3332 \end{bmatrix}$	0.4082	0.0002
$\begin{bmatrix} 0.3333 & -0.2357 \\ 0.2357 & 0.3333 \end{bmatrix}$	0.4082	0.0001

Example 2: Consider $R(\Delta) = I - \Delta M$ where

$$M = \begin{bmatrix} 2+j & 0 \\ 0 & 1+j2 \end{bmatrix}$$

taken from [5]. There, an optimal solution to Problem P was found to be

$$\Delta^* = \begin{bmatrix} 0.3333 & -0.2981 \\ 0.2981 & 0.3333 \end{bmatrix}$$

and its norm i.e. the optimum is 0.4472. To test our algorithm, we use the randomly generated matrix

$$\Delta_0 = \begin{bmatrix} 0.4337 & 0.1160 \\ 0.7092 & 0.0781 \end{bmatrix}$$

as a starting point and numerically solve the ODE (3.3) successively for $k = 50, 280, 560, 840$ to obtain a sequence of 148 perturbation matrices. Figure 4 displays three curves associated with $|\det(R(\Delta))|$, $\|\Delta - \Delta^*\|$, and $\|\Delta\| - \|\Delta^*\|$, respectively. It is clear that the sequence converges to the optimal solution while the sequence of corresponding norms converges to the minimum. The actual time taken to complete the simulation is about 10 seconds.

Example 3: Consider a nonlinear $R(\Delta)$ given by

$$R(\Delta) = I - \Delta \begin{bmatrix} 1+j & 2 \\ 0 & 1 \end{bmatrix} - \Delta \begin{bmatrix} 2+j & 0 \\ 0 & 1+j2 \end{bmatrix} \Delta^T$$

Note that without the quadratic term the minimum for Problem P was found to be 0.4472 in [5]. By invoking our algorithm with the randomly generated starting point

$$\Delta_0 = \begin{bmatrix} 0.07745 & 0.10237 \\ 0.76290 & 0.57027 \end{bmatrix}$$

a worst real perturbation is found to be

$$\Delta^* = \begin{bmatrix} 0.02751 & -0.33775 \\ 0.37673 & 0.03563 \end{bmatrix}$$

with $\|\Delta^*\| = 0.37854$ and $|\det(R(\Delta^*))| = 5 \times 10^{-6}$.

5 Conclusions

This paper has proposed a unified approach to solving the real and complex stability radius problems in the face of nonlinearly structured perturbations. It has been proved and verified with numerical examples that an optimal solution can be approached by solving a smooth and unconstrained minimization problem.

References

[1] R. V. Patel and M. Toda, "Quantitative measures of robustness for multivariable systems," in *Proc. Automat. Contr. Conf.*, (San Francisco), 1980.

[2] R. K. Yedavalli, "Improved measures of stability robustness for linear state space models," *IEEE Trans. Autom. Control*, vol. 30, pp. 557–579, 1985.

[3] K. Zhou and P. P. Khargonekar, "Stability robustness bounds for linear state-space models with structured uncertainty," *IEEE Trans. Autom. Control*, vol. 32, pp. 621–623, 1987.

[4] R. Genesio and A. Tesi, "Results on the stability robustness of systems with state space perturbations," *Systems and Control Letters.*, vol. 11, pp. 39–46, 1988.

[5] L. Qiu, B. Bernhardsson, A. Rantzer, E. Davison, P. Young, and J. Doyle, "A foumula for computation of the real stability radius," *Automatica*, vol. 31, pp. 879–890, 1995.

[6] J. Sreedhar, P. V. Dooren, and A. L. Tits, "A fast algorithm to compute the real structured stability radius," in *Stability Theory*, pp. 219–230, Birkhäuser Verlag Basel, 1996.

[7] R. W. Brockett, "Dynamical systems that sort lists, diagonalise matrices and solve linear programming problems," *Linear Algebra and its Applications*, vol. 146, pp. 79–91, 1991.

[8] U. Helmke and J. B. Moore, *Optimization and Dynamical Systems.* London: Spring-Verlag, 1994.

[9] J. E. Dennis Jr. and R. B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations.* Englewood Cliffs: Prentice-Hall, 1983.
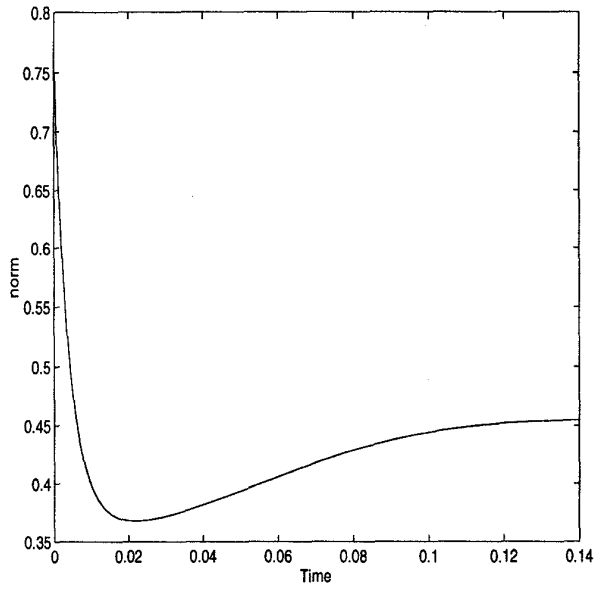
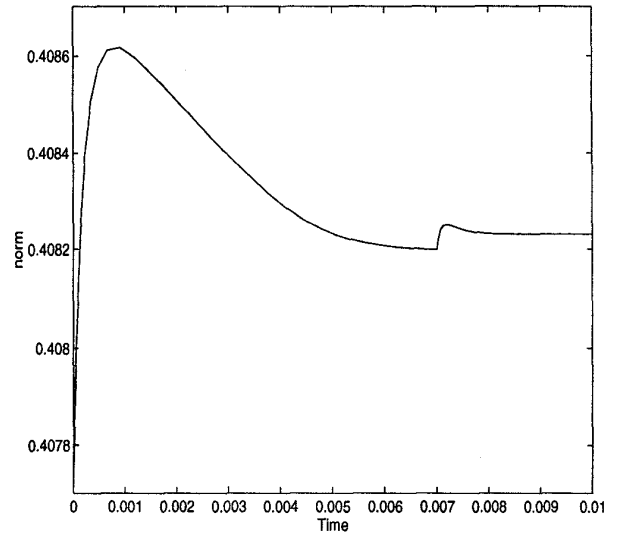Figure 1: The norm of the ODE solution versus $t \in [0, 0.14]$ for $k = 50$.

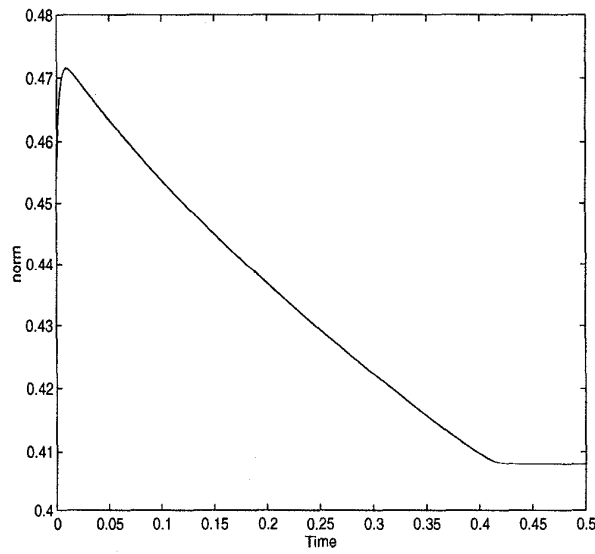Figure 3: The norm of the combined ODE solution versus $t \in [0, 0.01]$ for $k = 1000, 3000$.

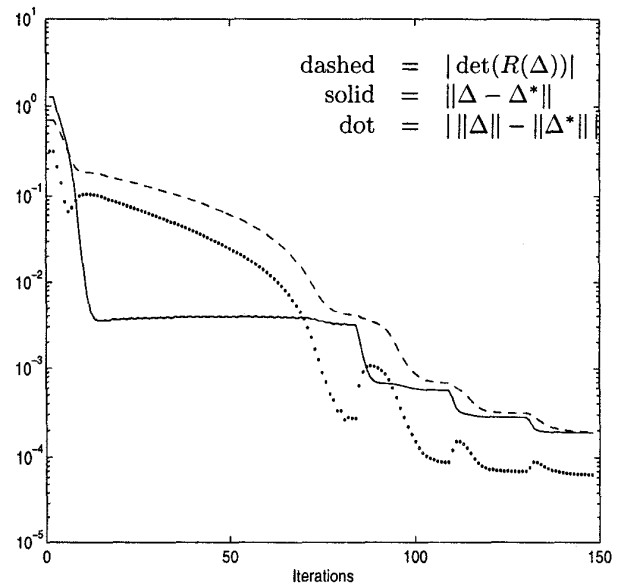Figure 2: The norm of the ODE solution versus $t \in [0, 0.5]$ for $k = 100$.

Figure 4: Convergence of a sequence of ODE solutions to a minimum point.