MP1-3 5:00

Proceedings of the 1999 IEEE
International Symposium on Computer Aided Control System Design
Kohala Coast-Island of Hawai'i, Hawai'i, USA • August 22-27, 1999

Stability radii of polynomial matrices

Yves Genin and Paul Van Dooren
CESAME, Université catholique de Louvain
Av. G. Lemaître, 4, B- 1348 Louvain-la-Neuve, BELGIUM
e-mail contact : vdooren@csam.ucl.ac.be

Abstract

We derive analytic expressions for the stability radius of polynomial matrices for all Hölder norms and discuss numerical issues for computing these stability radii for the 1, 2 and ∞ norm.

Keywords: Stability radius, numerical methods, Hamiltonian pencils, symplectic pencils

1 Introduction

In this paper we consider polynomial matrices

$$P(\lambda) \doteq P_0 + P_1\lambda + \cdots + P_k\lambda^k$$

which are square invertible and have zeros – i.e. roots of $\det P(\lambda)$ – inside a given region Γ. We say that $P(\lambda)$ is Γ-stable and call Γ the stability region. The complex stability radius r_C of such polynomial matrices is the norm of the smallest perturbation

$$\Delta P(\lambda) \doteq \Delta P_0 + \Delta P_1\lambda + \cdots \Delta P_k\lambda^k$$

needed to "destabilize" $P(\lambda)+\Delta P(\lambda)$ and hence causing at least one zero of $P(\lambda)+\Delta P(\lambda)$ to leave the region Γ. If we measure the perturbations via the norm of a constant matrix Δ depending on the coefficients of $\Delta P(\lambda)$:

$$\|\Delta\| \doteq g\left(\Delta P_0, \Delta P_1, \cdots, \Delta P_k\right) \tag{1}$$

then we have the expression

$$r_C = \inf\{\|\Delta\| : \exists\, \mathrm{root}(P(\lambda) + \Delta P(\lambda)) \in \Gamma_c\}, \tag{2}$$

where Γ_c is the complement of Γ. The two regions that are typically considered for Γ are the open left half plane and the open unit disc, which are both open and connected sets of the complex plane. By continuity of zeros of perturbed matrices, the root "leaving" Γ must actually lie on its boundary $\partial\Gamma$, which can be parameterized by a real variable ω. In this paper we prove for Hölder norms that

$$r_C^{-1} = \sup_{\lambda\in\partial\Gamma} \|G(\lambda)\|_p, \quad G(\lambda) = d(\lambda)\cdot P(\lambda)^{-1}, \tag{3}$$

where $d(\lambda)$ is a scalar function depending on the choice of function g in (1).

We also give in this paper some considerations about how to compute the expressions in a numerically sound way for the case of the 2-norm.

2 Polynomial matrices

In [6] we consider the calculation of stability radii of a first order polynomial matrix $P(\lambda) = P_0 + P_1\lambda$ with $P_0 = -A$ and $P_1 = E$. It turns out that the results obtained can be extended to polynomial matrices of arbitrary degree, i.e. $P(\lambda) = \sum_{i=0}^{k} P_i\lambda^i$. We consider here the perturbation matrices :

$$\begin{aligned}
\Delta_1 &= \begin{bmatrix} \Delta P_0 & \Delta P_1 & \cdots & \Delta P_k \end{bmatrix}, \qquad (4)\\
\Delta_2 &= \begin{bmatrix} \Delta P_0 \\ \Delta P_1 \\ \vdots \\ \Delta P_k \end{bmatrix}\\
\Delta_3 &= \begin{bmatrix} \Delta P_0 & & & \\ & \Delta P_1 & & \\ & & \ddots & \\ & & & \Delta P_k \end{bmatrix}.
\end{aligned}$$

The polynomial matrix perturbation $\Delta P(\lambda)$ can then be expressed as

$$\begin{aligned}
\Delta P(\lambda) &= \Delta_1 \begin{bmatrix} I \\ \lambda I \\ \vdots \\ \lambda^k I \end{bmatrix} = \begin{bmatrix} I & \lambda I & \ldots & \lambda^k I \end{bmatrix} \Delta_2\\
&= \begin{bmatrix} I & \alpha^{-1} I & \ldots & \alpha^{-k} I \end{bmatrix} \Delta_3 \begin{bmatrix} I \\ (\alpha\lambda) I \\ \vdots \\ (\alpha\lambda)^k I \end{bmatrix},
\end{aligned}$$

where α is arbitrary, as long as α^{-1} and $(\alpha\lambda)$ are bounded. It will prove useful later on for the calculation of Hölder norms to choose

$$\alpha^{-1} \doteq \lambda^{\frac{1}{q}} \Rightarrow (\alpha\lambda) = \lambda^{1-\frac{1}{q}} = \lambda^{\frac{1}{p}}. \qquad (5)$$

In order to prove something about Hölder norms for such perturbations, we first need the following lemma.

Lemma 1 *For every Hölder (or p) norm and vectors x and y one has the multiplicative property*

$$\|x \otimes y\|_p = \|x\|_p \|y\|_p. \qquad (6)$$

Proof: Since the elements of the vector $x \otimes y$ are $x_i y_j$ one has

$$\|x \otimes y\|_p = (\sum_i \sum_j |x_i y_j|^p)^{\frac{1}{p}}$$

$$= ((\sum_i |x_i|^p)(\sum_j |y_j|^p))^{\frac{1}{p}} = \|x\|_p \|y\|_p.$$

From this, it immediately follows that the induced p-norm of matrices with repetitions of scaled identities are easy to compute in terms of the scaling factors :

$$\|x \otimes I\|_p = \|x\|_p, \quad \|y^T \otimes I\|_p = \|y^T\|_p.$$

The first equality follows from the identity $\|(x \otimes I)y\|_p = \|x \otimes y\|_p = \|x\|_p \|y\|_p$ which holds for every vector y, and the second is obtained by duality (the p-norm of a matrix equals the q-norm of its transpose, where $1/p + 1/q = 1$, and both are Hölder norms [5]). For more complex matrices one derives a similar formula.

Corollary 1 *We have the following identity for the induced matrix p-norm :*

$$\begin{aligned} \|(x \otimes I)M(y^T \otimes I)\|_p &= \|x \otimes I\|_p \|M\|_p \|y^T \otimes I\|_p \\ &= \|x\|_p \|M\|_p \|y^T\|_p. \end{aligned}$$

Proof: The second equality follows from the above discussion, and for induced norms we have the multiplicative inequality

$$\|(x \otimes I)M(y^T \otimes I)\|_p \leq \|x \otimes I\|_p \|M\|_p \|y^T \otimes I\|_p.$$

We show that equality holds by constructing a vector for which the upper bound of the operator norm is achieved. Let $v = Mu$ be such that $\|v\|_p = \|Mu\|_p = \|M\|_p \|u\|_p$ and z be such that $|y^T z| = \|y^T\|_p \|z\|_p$. Then the vector $(z \otimes u)$ has norm $\|z\|_p \|u\|_p$ and achieves this upper bound since

$$\begin{aligned} &\|(x \otimes I)M(y^T \otimes I)(z \otimes u)\|_p = \|(x \otimes I)M(y^T z)u\|_p \\ &= \|(x \otimes v)(y^T z)\|_p = \|x\|_p \|M\|_p \|u\|_p \|y^T\|_p \|z\|_p. \end{aligned}$$

Theorem 1 *For all λ for which $P(\lambda)$ is invertible, we have*

$$\begin{aligned} &\inf_{\Delta_i} \{\|\Delta_i\|_p : \det(I + \Delta P(\lambda) P(\lambda)^{-1}) = 0\} \\ &\qquad = \|d_i(\lambda) P(\lambda)^{-1}\|_p^{-1}, \end{aligned}$$

where $d_i(\lambda)$ for Δ_i, $i = 1, 2, 3$ equals respectively

$$\begin{aligned} d_1(\lambda) &= \left(\sum_{i=0}^{k} |\lambda|^{ip}\right)^{\frac{1}{p}}, \\ d_2(\lambda) &= \left(\sum_{i=0}^{k} |\lambda|^{iq}\right)^{\frac{1}{q}}, \\ d_3(\lambda) &= \left(\sum_{i=0}^{k} |\lambda|^{i}\right), \end{aligned}$$

Proof: In [6] it is shown that for every p-norm one has

$$\det(I + \Delta M) = 0 \Rightarrow \|\Delta\|_p \geq \|M\|_p^{-1},$$

and equality is met if Δ is unconstrained. Since $P(\lambda)$ is assumed invertible we have

$$\det(P(\lambda) + \Delta P(\lambda)) = 0 \Leftrightarrow \det(I + \Delta P(\lambda) P(\lambda)^{-1}) = 0.$$

Simple algebraic manipulations then imply that this is also equivalent to

$$\begin{aligned} &\det(I + \Delta_1 \begin{bmatrix} I \\ \lambda I \\ \vdots \\ \lambda^k I \end{bmatrix} P(\lambda)^{-1}) = 0 \Leftrightarrow \\ &\det(I + \Delta_2 P(\lambda)^{-1} \begin{bmatrix} I & \lambda I & \dots & \lambda^k I \end{bmatrix}) = 0 \Leftrightarrow \\ &\det(I + \Delta_3 \begin{bmatrix} I \\ \lambda^{\frac{1}{p}} I \\ \vdots \\ \lambda^{\frac{k}{p}} I \end{bmatrix} P(\lambda)^{-1} \begin{bmatrix} I & \lambda^{\frac{1}{q}} I & \dots & \lambda^{\frac{k}{q}} I \end{bmatrix}) = 0. \end{aligned} \qquad (7)$$

Applying then the results of corollary 1, we immediately get the bounds

$$\|\Delta_1\|_p \geq (\|P(\lambda)^{-1} d_1(\lambda)\|_p)^{-1} = (\|P(\lambda)^{-1}\|_p \| \begin{bmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^k \end{bmatrix} \|_p)^{-1}$$

$$\|\Delta_2\|_p \geq (\|P(\lambda)^{-1} d_2(\lambda)\|_p)^{-1} = (\|P(\lambda)^{-1}\|_p \| \begin{bmatrix} 1 & \lambda & \dots & \lambda^k \end{bmatrix} \|_p)^{-1}$$

$$\|\Delta_3\|_p \geq (\|P(\lambda)^{-1}d_3(\lambda)\|_p)^{-1} =$$

$$(\|P(\lambda)^{-1}\|_p \| \begin{bmatrix} 1 & \lambda^{\frac{1}{q}} & \dots & \lambda^{\frac{k}{q}} \end{bmatrix} \|_p \| \begin{bmatrix} 1 \\ \lambda^{\frac{1}{p}} \\ \vdots \\ \lambda^{\frac{k}{p}} \end{bmatrix} \|_p)^{-1}.$$

This shows that the right hand sides are lower bounds for $\|\Delta_i\|_p$. From [6] it follows that these bounds are also equalities for Δ_1 and Δ_2 since these are unconstrained matrices, but we still need to establish equality for Δ_3 which is constrained to be block diagonal. We now construct a particular perturbation for each Δ_i which establishes equality for all of them (although this is not needed for Δ_1 and Δ_2). Let u and v be two vectors of unit p-norm such that $P(\lambda)^{-1}u = \|P(\lambda)^{-1}\|_p v$ and let v_D^T be the dual of v, hence $|v_D^T v| = 1$ with $\|v_D^T\|_p = 1$. The matrix entries ΔP_i defined by

$$\begin{aligned} \Delta_1 : \quad & \Delta P_i = -(\|d_1(\lambda)^p P(\lambda)^{-1}\|_p)^{-1} u v_D^T (\tfrac{|\lambda|^p}{\lambda})^i \\ \Delta_2 : \quad & \Delta P_i = -(\|d_2(\lambda)^q P(\lambda)^{-1}\|_p)^{-1} u v_D^T (\tfrac{|\lambda|^q}{\lambda})^i \\ \Delta_3 : \quad & \Delta P_i = -(\|d_3(\lambda) P(\lambda)^{-1}\|_p)^{-1} u v_D^T (\tfrac{|\lambda|}{\lambda})^i \end{aligned} \quad (8)$$

yield equalities in all bounds and also satisfy $\Delta P(\lambda) = -(\|P(\lambda)^{-1}\|_p)^{-1} u v_D^T$ from which it follows that $\Delta P(\lambda) P(\lambda)^{-1} u = -u$ and $(P(\lambda) + \Delta P(\lambda))v = 0$. This is now used in the following characterization of the stability radius of polynomial matrices.

Theorem 2 *The smallest perturbation of a polynomial matrix $P(\lambda)$ causing a zero of $P(\lambda) + \Delta P(\lambda)$ to reach the boundary $\partial\Gamma$ of the stability region Γ, is given by*

$$\begin{aligned} \inf_{\Delta_i} \{\|\Delta_i\|_p : \det(P(\lambda) + \Delta P(\lambda)) = 0,\ \lambda \in \partial\Gamma\} \\ = \inf_{\lambda \in \partial\Gamma} \|d_i(\lambda) P(\lambda)^{-1}\|_p^{-1} \\ = \left\{\sup_{\lambda \in \partial\Gamma} \|d_i(\lambda) P(\lambda)^{-1}\|_p\right\}^{-1}, \end{aligned}$$

where $d_i(\lambda)$ for Δ_i, $i = 1,2,3$ are defined as in Theorem 1.

3 Computational aspects

We point out here that the above complex stability radius can be computed efficiently provided $G(\lambda) = d(\lambda)P(\lambda)^{-1}$ is rational in λ for $\lambda \in \partial\Gamma$. This is obviously the case for the unit circle since there these functions are constant :

$$d_1(\lambda) = (k+1)^{\frac{1}{p}}, \quad d_2(\lambda) = (k+1)^{\frac{1}{q}},$$

$$d_3(\lambda) = (k+1).$$

For the $\jmath\omega$ axis, one can still replace $d_3(\lambda)$ by the following polynomials of the same amplitude

$$\begin{aligned} d(\lambda) &= \textstyle\sum_{i=0}^k (-\jmath\lambda)^i, \text{ for } \omega \geq 0 \\ d(\lambda) &= \textstyle\sum_{i=0}^k (\jmath\lambda)^i, \text{ for } \omega \leq 0, \end{aligned} \quad (9)$$

i.e. one has to consider two different rational functions for positive and negative values of ω. For $d_1(\lambda)$ and $d_2(\lambda)$ one can only do this for the special cases $p = 1, 2, \infty$. For $p = q = 2$ one has

$$|d_1(\jmath\omega)|^2 = |d_2(\jmath\omega)|^2 = (1 + \omega^2 + \cdots + \omega^{2k}) = |d(\jmath\omega)|^2$$

where $d(\lambda)$ is the (stable) spectral factor of $1 + \omega^2 + \cdots + \omega^{2k}$ and equals

$$d(\lambda) = \prod_{i=1}^{k} \left(\lambda + sin\frac{i\pi}{k+1} - \jmath cos\frac{i\pi}{k+1}\right). \quad (10)$$

For the cases $p = 1$ and $q = 1$, $d_1(\lambda)$ and $d_2(\lambda)$ reduce respectively to $d(\lambda)$ given in (9), while for $p = \infty$ and $q = \infty$, $d_1(\lambda)$ and $d_2(\lambda)$ reduce respectively to $\max\{1, |\omega|^k\}$ and hence has the same amplitude as the rational functions

$$d(\lambda) = 1, \text{ for } \omega \leq 1, \quad d(\lambda) = \lambda^k, \text{ for } \omega > 1.$$

Note that in each of these cases the constructed polynomial $d(\lambda)$ has degree k or less, i.e. $d(\lambda) = \sum_{i=1}^k d_i \lambda^i$. The transfer function matrix $d(\lambda)P(\lambda)^{-1}$ admits then a generalized state space realization of the form $C(\lambda E - A)^{-1}B$ where

$$A = \begin{bmatrix} 0_n & I_n & & & \\ & 0_n & I_n & & \\ & & \ddots & \ddots & \\ & & & 0_n & I_n \\ -P_0 & -P_1 & \cdots & \cdots & -P_k \end{bmatrix}, \quad E = I_{kn} \oplus 0_n;$$

$$B = [0_n\ 0_n\ \cdots\ 0_n\ I_n]^T, \quad C = [d_0 I_n\ d_1 I_n\ \cdots\ d_k I_n].$$

For the 2-norm the corresponding complex stability radius reduces to the $\mathcal{H}_\infty$-norm of the transfer function $G(\lambda)$:

$$\sigma_* = \sup_{\omega \in \mathcal{R}} \sigma_{\max}\{G(f(\omega))\}, \quad (11)$$

where $f(\omega)$ is the parameterization of $\partial\Gamma$ in terms of $\omega \in \mathcal{R}$, and $\sigma_{\max}\{M\}$ is the largest singular value of the matrix M. This can be obtained iteratively using the computation of the real zeros ω_i of the matrix function

$$G(f(\omega))G(f(\omega))^* - \sigma_o^2 I. \quad (12)$$

It turns out that ω_i is a real zero of (12) iff σ_o is a singular value of $G(f(\omega_i))$; this property leads to efficient algorithms to find the maximum of the scalar function $\sigma(\omega) = \sigma_{\max}\{G(f(\omega))\}$ [4], [1], [2]. Each of these methods uses an eigenvalue problem (with Hamiltonian or symplectic structure) to compute the zeros of (12). These apply to a generalized state-space model $G(\lambda) = C(\lambda E - A)^{-1}B$ and are given by :

$$\begin{bmatrix} A - \jmath\omega E & BB^* \\ C^*C\sigma_o^{-2} & A^* + \jmath\omega E^* \end{bmatrix} \quad (13)$$

$$\begin{bmatrix} A - e^{\jmath\omega}E & BB^* \\ e^{\jmath\omega}C^*C\sigma_o^{-2} & e^{\jmath\omega}A^* - E^* \end{bmatrix}. \quad (14)$$

Notice that the $\jmath\omega$ pencil is Hamiltonian and that the $e^{\jmath\omega}$ pencil is symplectic.

For $p = 1, \infty$ one still has a rational matrix to deal with, but the largest singular value is now replaced by the largest sum of absolute values in a column or row of $G(\lambda)$. This is a scalar piecewise rational function, which can be maximized using symbolic manipulation programs : each "piece" is rational and the "branching points" are zeros of some polynomial.

In the special case of scalar polynomials, obviously $\|P(\lambda)^{-1}\|_p = |p(\lambda)^{-1}|$. If moreover we choose $p = 1, 2, \infty$ then $d(\lambda)$ can also be chosen polynomial, hence we then have to find the maximum of the absolute value of a scalar rational function $G(\lambda) = d(\lambda)/p(\lambda)$ on $\partial\Gamma$. The zeros of $d'(\lambda)p(\lambda) - p'(\lambda)d(\lambda)$ are then the extrema of this function and it suffices to look for the largest of these $2(k-1)$ values. This can be obtained in $O(k^2)$ flops using polynomial root finding algorithms. Notice that other approaches have been proposed in the literature for complex stability radii of scalar polynomials [3].

4 Maximal stability radius

We complete this paper with a result on polynomials with optimal stability radius. We only consider the scalar case with leading coefficient 1. The question we try to address is which (stable) polynomial $p(\lambda) = p_0 + p_1\lambda + \ldots + p_{k-1}\lambda^{k-1} + \lambda^k$ has the largest possible stability radius.

For the discrete-time case the stability radius of a stable polynomial $p(\lambda)$ is proportional to the inverse of $\sup_{|\lambda|=1} |p(\lambda)|$. We thus have to minimize this over all possible stable polynomials. These can also be parametrized as $p(\lambda) = \prod_{i=1}^{k}(\lambda - \lambda_i)$ where $|\lambda_i| < 1$. But

$$\sup_{|\lambda|=1} |p(\lambda)| \leq \prod_{i=1}^{k} \sup_{|\lambda|=1} |(\lambda - \lambda_i)| = \prod_{i=1}^{k}(1 + |\lambda_i|)$$

which is minimized for $\lambda_i = 0$. The "most stable" discrete-time polynomial is thus λ^k. It is interesting to point out that it has a zero of highest possible multiplicity, and hence its zeros are very sensitive to infinitesimal perturbations. Nevertheless it is the most robust polynomial against large perturbations that preserve stability !

For the continuous-time case the stability radius of a stable polynomial $p(\lambda)$ is proportional to the inverse of $\sup_{\lambda=\jmath\omega} |p(\lambda)/d(\lambda)|$, where $d(\lambda)$ is also monic. We have to minimize this over all possible stable polynomials, but it is easy to see that for $\lambda = \pm\jmath\infty$ this function has value one and hence one will not be able to improve on this value. But there are many monic polynomials $p(\lambda)$ for which $|p(\jmath\omega)| \leq |d(\jmath\omega)|$. The obvious one is $d(\lambda)$ itself if it is stable and its stable inner factor otherwise. But moving all zeros of $p(\lambda)$ to the left in the complex plane will also do the job. The continuous-time version of this problem therefore has no unique solution.

5 Acknowledgments

This paper presents research results of the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. The scientific responsibility rests with its authors.

References

[1] S. Boyd and V. Balakrishnan, "A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its L_∞ norm," *Systems & Control Letters*, vol. 15, pp. 1–7, 1990.

[2] Y. Genin, P. Van Dooren and V. Vermaut, "Convergence of the calculation of $\mathcal{H}_\infty$-norms and related questions", Proceedings NTNS-98, Padua, July 1998.

[3] D. Hinrichsen and A.J. Pritchard, "Robustness measures for linear systems with applications to stability radii of Hurwitz nad Schur polynomials", *Int. J. Contr.*, vol. 55(4), pp. 809–844, 1992.

[4] D. Hinrichsen and N.K. Son, "The complex stability radius of discrete-time systems and symplectic pencils," in *Proc. IEEE 28th Conference on Decision and Control*, (Tampa, FL), pp. 2265–2270, 1989.

[5] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[6] P. Van Dooren and V. Vermaut, "On stability radii of generalized eigenvalue problems", *European Control Conf.*, paper FR-M-H6, 1997.