Convergence of the calculation of $\mathcal{H}_\infty$-norms and related questions

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

Keywords: Stability radius, $\mathcal{H}_\infty$-norm, numerical methods, Hamiltonian pencils, symplectic pencils

Abstract

In this paper we analyze the convergence properties of an iterative algorithm based on cubic interpolation which was originally proposed for the calculation of the real stability radius [5]. When applied to the calculation of complex stability radii of generalized eigenvalue problems and of the $\mathcal{H}_\infty$-norm of an arbitrary rational transfer function, this new algorithm has global linear convergence and ultimate quartic convergence. It therefore compares favorably with earlier algorithms, analyzed e.g. in [1].

1 Introduction

The complex stability radius $r_\mathcal{C}$ of a stable matrix A with respect to a stability region Γ is the norm of the smallest perturbation Δ needed to "destabilize" $A+\Delta$, causing at least one eigenvalue of $A+\Delta$ to leave the region Γ :

$$r_\mathcal{C} = \inf\{||\Delta|| : \exists\lambda(A+\Delta)\in\Gamma_c\},$$

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 eigenvalues of perturbed matrices, the eigenvalue "leaving" Γ must actually lie on its boundary $\partial\Gamma$. This boundary (the $\jmath\omega$ axis or the unit circle $e^{\jmath\omega}$) can then be parameterized by a real variable ω. A simple argument now shows the equality

$$r_\mathcal{C}^{-1} = \sup_{\lambda\in\partial\Gamma} ||(\lambda I - A)^{-1}|| \tag{1}$$

for any subordinate norm [6]. For low rank perturbations and several other types of generalized eigenvalue problems (pencils and polynomial matrices), it was also shown in [6] that the corresponding complex stability radius can be reduced to computing $\sup_{\lambda\in\partial\Gamma}||G(\lambda)||$ for an appropriate *rational* transfer function $G(\lambda)$. If one chooses the 2-norm, then this is nothing but the $\mathcal{H}_\infty$-norm of the transfer function $G(\lambda)$:

$$\sigma_* = \sup_{\omega\in\mathcal{R}} \sigma_{\max}\{G(f(\omega))\}, \tag{2}$$

where $f(\omega)$ is the parametrization of $\partial\Gamma$ in terms of $\omega\in\mathcal{R}$, and $\sigma_{\max}\{M\}$ is the largest singular value of the matrix M. The computation of (2) can be performed iteratively using a test for the existence of real zeros ω_i of the matrix function

$$G(f(\omega))G(f(\omega))^* - \sigma_o^2 I. \tag{3}$$

It turns out that ω_i is a real zero of (3) iff σ_o is a singular value of $G(f(\omega_i))$, which then leads to a test for a bisection algorithm to find the maximum of the scalar function $\sigma(\omega) = \sigma_{\max}\{G(f(\omega))$ [3]. Later on, the points ω_i were also used to improve the convergence, which finally lead to algorithms with ultimate quadratic convergence [4], [2], [1]. Each of these methods uses an eigenvalue problem (with Hamiltonian or symplectic structure) to compute the zeros of (3). In the most general form these are derived for a generalized state-space model $G(\lambda) = C(\lambda E - A)^{-1}B + D$ and are given by :

$$\begin{bmatrix} A-\jmath\omega E & BB^* \\ 0 & A^*+\jmath\omega E^* \end{bmatrix} - \begin{bmatrix} BD^* \\ C^* \end{bmatrix} D_{\sigma_o}^{-1} \begin{bmatrix} C & DB^* \end{bmatrix}.$$

$$\begin{bmatrix} A-e^{\jmath\omega} E & BB^* \\ 0 & e^{\jmath\omega}A^* - E^* \end{bmatrix} - \begin{bmatrix} BD^* \\ e^{\jmath\omega}C^* \end{bmatrix} D_{\sigma_o}^{-1} \begin{bmatrix} C & DB^* \end{bmatrix}$$

where $D_{\sigma_o} \doteq DD^* - \sigma_o^2 I$. Notice that the $\jmath\omega$ pencil is Hamiltonian and that the $e^{\jmath\omega}$ pencil is symplectic.

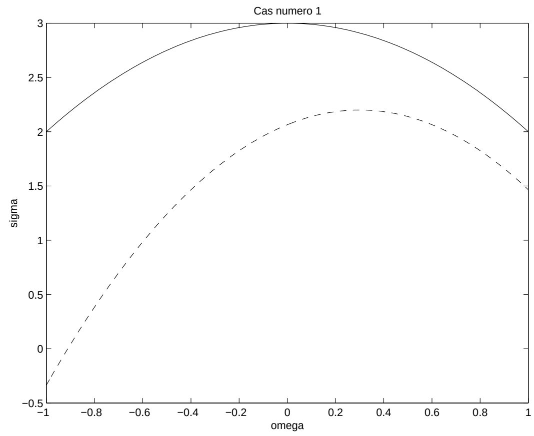
2 New results

More recently Sreedhar, Tits and Van Dooren [5] introduced a more efficient algorithm by taking benefit of the fact that the above eigenvalue problems not only give the intersection points of the function $\sigma(\omega)$ with a particular "level" σ_o but also the derivative of the function in these points, which is obtained at little extra cost from the generalized eigenvectors of the corresponding pencil. In a first phase, the set of subintervals of the real axis among which the optimum ω_* necessarily belongs, is computed by the algorithm for a given σ_o. These subintervals are determined by computing the real zeros of (3) corresponding to this value of σ_o. In the second phase, σ_o is increased to the largest value obtained from considering the successive midpoints of the above subintervals. This two phase process can then be iterated up to convergence an therefore delivers the supremum σ_* in a finite number of steps within any required degree of accuracy. The new algorithm presented in this paper can be viewed as a refinement, where the derivatives of the real eigenvalues at the endpoints of the subintervals are not simply used for their proper determination but also to speed up the convergence. For each subinterval, the point of interest is no longer its midpoint but the point corresponding to the maximum of the cubic polynomial interpolating the endpoints of the considered interval together with their derivatives. Let us explain all this in detail.

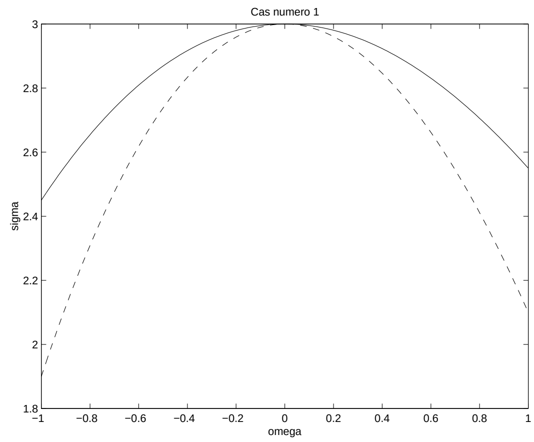
Two different situations need to be considered near the optimum. In the first situation, the endpoints of the subinterval of interest correspond to the same singular value $\sigma(\omega)$, which for values of ω sufficiently close to ω_* admits a Taylor expansion of the form

$$\sigma(\omega) = \sigma_* - D(\omega-\omega_*)^2 + E(\omega-\omega_*)^3 + F(\omega-\omega_*)^4 + G(\omega-\omega_*)^5 + \mathcal{O}((\omega-\omega_*)^6). \quad (4)$$

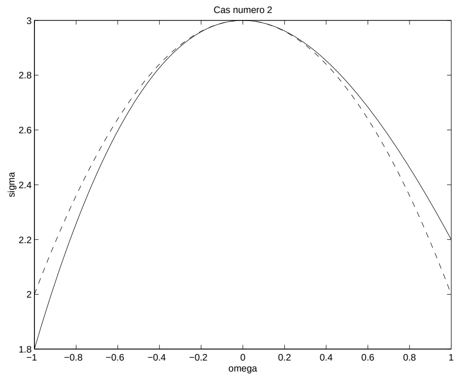
This pattern of singular values near σ_*, most frequently met in actual calculations, will be referred to as the *generic* case. It arises either when the optimal singular value is simple :



or when it is multiple with predominance of one of them :



In the second situation, the endpoints of the optimal subinterval turn out to be defined by two distinct singular values. In such cases however, there clearly exists a value of σ above which the two endpoints will correspond to the two same well defined singular values, say $\sigma_1(\omega)$ and $\sigma_2(\omega)$.



By continuity, these two singular values can be approximated as above in the neighborhood of ω_* by the expansions

$$\begin{aligned}\sigma_1(\omega) &= \sigma_* - D(\omega-\omega_*)^2 + E_1(\omega-\omega_*)^3 + F_1(\omega-\omega_*)^4 \\ &\quad + G_1(\omega-\omega_*)^5 + \mathcal{O}((\omega-\omega_*)^6), \\ \sigma_2(\omega) &= \sigma_* - D(\omega-\omega_*)^2 + E_2(\omega-\omega_*)^3 + F_2(\omega-\omega_*)^4 \\ &\quad + G_2(\omega-\omega_*)^5 + \mathcal{O}((\omega-\omega_*)^6).\end{aligned} \quad (5)$$

Note that one has necessarily $D = D_1 = D_2$ in view of the assumed singular value interlacing property.

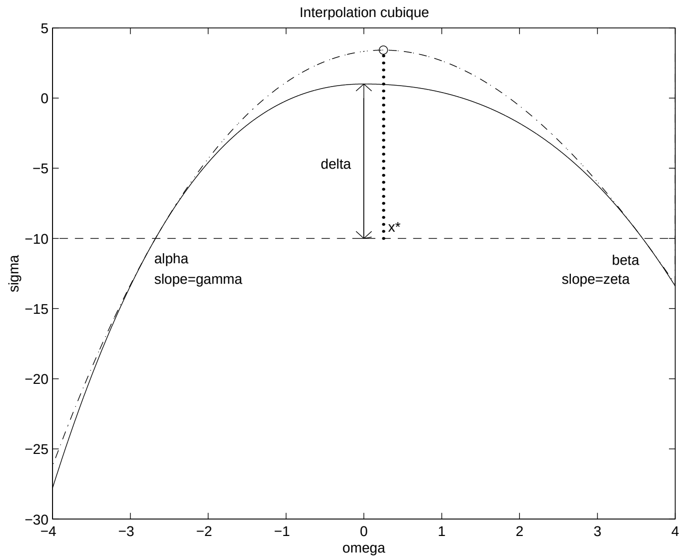
It appears that these two patterns of evolution of the singular values near the supremum σ_* can be discussed within the same theoretical approach. Indeed, the first situation can be viewed as a particular case of the second when $E_1 = E_2$, $F_1 = F_2$, $G_1 = G_2$, etc. For the sake of simplicity, we shall also assume, without loss of generality, that $\omega_* = 0$. Let us then call $\delta = \sigma_* - \sigma$ the positive difference between the supremum and the current value of σ and denote by $\alpha < 0$ and $\beta > 0$ the endpoints of the subinterval; therefore, one has by definition $\sigma_1(\alpha) = \sigma_2(\beta) = \sigma$. With $\gamma = d\,\sigma_1(\alpha)/d\,\omega > 0$ and $\zeta = d\,\sigma_2(\beta)/d\,\omega < 0$, let us consider the cubic polynomial interpolation problem defined as follows: find the polynomial $P(\omega)$ of degree 3 and satisfying the constraints $P(\alpha) = P(\beta) = -\delta$, $P'(\alpha) = \gamma$ and $P'(\beta) = \zeta$. Clearly, the resulting polynomial has a maximum at some point, say $\omega = \omega'$, belonging to the interval (α, β). The proposed algorithm consists in selecting $\omega = \omega'$ (instead of the interval midpoint) as the next point of computation of the singular values, which in turn yields the updated $\sigma-$level, say σ_{new}, and, hence, the next positive difference $\delta_{new} = \sigma_* - \sigma_{new}$. It turns out that the convergence order induced by this algorithmic refinement is 4 in the generic case and 3 otherwise.

To prove this result, let us first consider the solution of the cubic interpolating polynomial $P(\omega)$, which is verified by direct observation to have the expression

$$\begin{aligned}P(\omega) &= (\frac{\omega-\beta}{\alpha-\beta})^2[-\delta\,(1-2\,\frac{\omega-\alpha}{\alpha-\beta}) + \gamma\,(\omega-\alpha)] \\ &+ (\frac{\omega-\alpha}{\beta-\alpha})^2[-\delta\,(1-2\,\frac{\omega-\beta}{\beta-\alpha}) + \zeta\,(\omega-\beta)].\end{aligned} \quad (6)$$

Setting $\mu = (\alpha+\beta)/2$, $\epsilon = (\beta-\alpha)/2$, $z = (\omega-\mu)/\epsilon$, we introduce a normalized version $Q(z)$ of the interpolating polynomial $P(\omega)$ defined by $Q(z) = P(\epsilon\, z - \mu)$, i.e.

$$Q(z) = -\delta + \epsilon(z^2-1)\,[\gamma\,(z-1) + \zeta\,(z+1)]/4. \tag{7}$$



The point ω' (equivalently z') where $P(\omega)$ (equivalently $Q(z)$) reaches its maximum is then found to be given by

$$\begin{aligned} \omega' &= \frac{\alpha+\beta}{2} + \frac{\beta-\alpha}{2}\, z' \\ z' &= \frac{(\gamma-\zeta) - 2\sqrt{\gamma^2+\zeta^2+\gamma\,\zeta}}{3\,(\gamma+\zeta)} \end{aligned} \tag{8}$$

since one must have $Q'(z') = 0$ together with $Q''(z') < 0$. With $\sigma(\omega) = \sigma_1(\omega)$ for $\omega < 0$ and $\sigma(\omega) = \sigma_2(\omega)$ for $\omega > 0$ and therefore, $\sigma_{new} = \sigma'$ whence $\delta_{new} = \sigma_* - \sigma_{new}$, we are left with the problem of expanding δ_{new} as a function of δ to evaluate the convergence order of the proposed algorithm. To that aim, let us start from expressions (5) and their derivatives, which imply, in particular,

$$\begin{aligned} -\delta &= -C\,\alpha^2 + E_1\,\alpha^3 + F_1\,\alpha^4 + G_1\,\alpha^5 + \mathcal{O}(\alpha^6), \\ -\delta &= -C\,\beta^2 + E_2\,\beta^3 + F_2\,\beta^4 + G_2\,\beta^5 + \mathcal{O}(\beta^6), \end{aligned} \tag{9}$$

and

$$\begin{aligned} \gamma &= -2\,C\,\alpha + 3\,E_1\,\alpha^2 + 4\,F_1\,\alpha^3 + 5\,G_1\,\alpha^4 + \mathcal{O}(\alpha^5), \\ \zeta &= -2\,C\,\beta + 3\,E_2\,\beta^2 + 4\,F_2\,\beta^3 + 5\,G_2\,\beta^4 + \mathcal{O}(\beta^5). \end{aligned} \tag{10}$$

Therefore, it appears from inverting both expressions (9) that α and β take, up to the order δ^2, the values

$$\begin{aligned} \alpha &= -\frac{1}{D^{1/2}}\,\delta^{1/2} + \frac{E_1}{2\,D^2}\,\delta - \frac{5\,E_1^2 + 4\,D\,F_1}{8\,D^{7/2}}\,\delta^{3/2} \\ &\quad + \frac{2\,E_1^3 + 3\,D\,E_1\,F_1 + D^2\,G_1}{2\,D^5}\,\delta^2, \\ \beta &= \frac{1}{D^{1/2}}\,\delta^{1/2} + \frac{E_2}{2\,D^2}\,\delta + \frac{5\,E_2^2 + 4\,D\,F_2}{8\,D^{7/2}}\,\delta^{3/2} \\ &\quad + \frac{2\,E_2^3 + 3\,D\,E_2\,F_2 + D^2\,G_1}{2\,D^5}\,\delta^2, \end{aligned} \tag{11}$$

which, in turn, allows one via (10) to evaluate γ and ζ up to the order four in the variable $\delta^{1/2}$. By inserting the latter expansions in (8), one finds, after tedious but elementary algebraic operations, that the dominant terms in z' and ω' are given by

$$\begin{aligned} z' &= -\frac{E_1+E_2}{4\,D^{3/2}}\,\delta^{1/2} + \frac{11\,(E_2^2-E_1^2) + 12\,D\,(F_2-F_1)}{32\,D^3}\,\delta \\ &\quad + \mathcal{O}(\delta^{3/2}) \\ \omega' &= \frac{19\,(E_2^2-E_1^2) + 20\,D\,(F_2-F_1)}{32\,D^{7/2}}\,\delta^{3/2} + \mathcal{O}(\delta^2). \end{aligned}$$

Finally, as one has $\delta_{new} = -\,D\,{\omega'}^2 + \mathcal{O}({\omega'}^3)$, one establishes that δ_{new} can be written as

$$\delta_{new} = -\left(\frac{19\,(E_2^2-E_1^2) + 20\,D\,(F_2-F_1)}{32\,D^3}\right)^2 \delta^3 + \mathcal{O}(\delta^4) \tag{12}$$

and this proves the claimed property: the convergence order of the proposed algorithm is three in general ($\sigma_1 \neq \sigma_2$), four in the generic case ($\sigma_1 = \sigma_2$). Note incidentally that the same property holds true near the optimum for the successive interval lengths since one has $\beta - \alpha = 2(\delta/D)^{1/2} + \mathcal{O}(\delta)$ in view of (11).

As a result, it appears that the proposed algorithm significantly outperforms the two other existing procedures proposed in the literature for the same purpose. This is clear for the bisection method since its convergence is linear, as well as for the midpoint method that has quadratic convergence since it corresponds to the expansion

$$\frac{\alpha+\beta}{2} = \frac{E_1+E_2}{4\,D^2}\,\delta + \mathcal{O}(\delta^{3/2}). \tag{13}$$

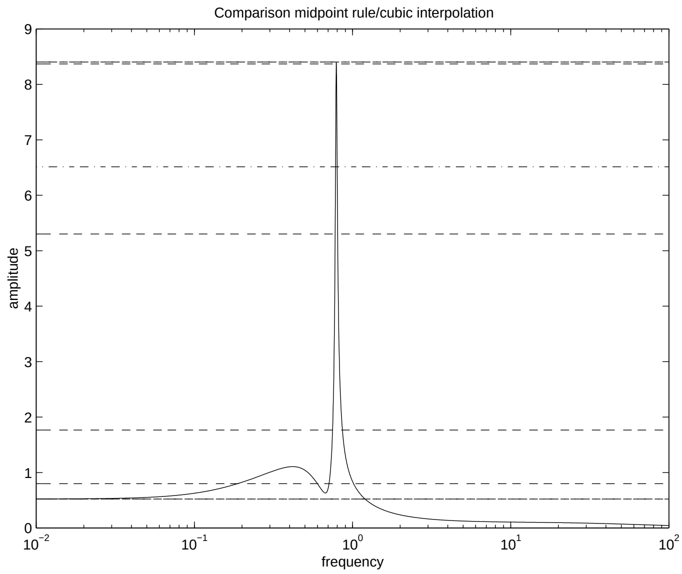
Let us also point out that our cubic interpolation algorithm does not entail any additional complexity with respect to the midpoint method since its only difference lies in the selection of the iteration point: ω' instead of the interval midpoint.

Finally, let us mention that the cubic interpolation method not only exhibits *local* accelerated convergence but also can be shown to converge *globally.* To see this, let us first consider the case of a single interval (α,β). As the reduced interval of interest in the next iteration will have necessarily ω' as one of its endpoints and since one derives from (8) the inequality

$$\left|\omega' - \frac{\alpha+\beta}{2}\right| = \frac{(\beta-\alpha)\,|\gamma-\zeta-2\sqrt{\gamma^2+\zeta^2+\gamma\,\zeta}\,|}{6\,|(\gamma+\zeta)|} \le \frac{\beta-\alpha}{6},$$

the length of this updated interval will be equal at most to $2/3$ of that of the previous interval. Next, let us recall that in the case of a collection of subintervals, the next σ-level results from selecting the largest singular value achieved at the successive ω' points relative to each of these subintervals. Moreover, as the subinterval length decreases with the σ-level, it turns out that each updated subinterval length is reduced by a factor 2/3 at least. Consequently, the global convergence of the proposed algorithm is guaranteed.

The above theoretical results are well illustrated in the next figure where the midpoint method on the one hand and the cubic interpolation method on the other hand are applied to the same numerical example. In this figure, the successive σ-levels are depicted in dash and dash-dotted lines for the midpoint and the cubic interpolation methods respectively.



Iteration	σ-level (midp.)	Intervals (midp.)
1	0.5224	[0,1.1991]
2	0.7980	[0.1867,0.5995]
		[0.7097,1.0153]
3	1.7669	[0.7472,0.8625]
4	5.3027	[0.7762,0.8048]
5	8.3691	[0.7884,0.7905]
6	8.4043	[0.78942,0.78943]

Iteration	σ-level (cubic)	Intervals (cubic)
1	0.5224	[0,1.1991]
2	6.5148	[0.7804,0.7994]
3	8.4043	[0.78942,0.78943]
4	8.4043	Convergence

The convergence acceleration due to the cubic interpolation method is clearly observed on this example. In fact, the convergence is so fast that its precise order (4 in the present case) is practically undetectable from the few successive numerical iterations.

3 Conclusion

We presented in this paper an improved algorithm for computing $\mathcal{H}_\infty$-norms and stability radii. It is based on cubic interpolation, has a convergence that is far superior to earlier algorithms and requires very little additional work per iteration step.

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] N. A. Bruinsma and M. Steinbuch, "A fast algorithm to compute the H_∞ norm of a transfer function matrix," *Systems & Control Letters*, vol. 14, pp. 287–293, 1990.

[3] R. Byers, "A bisection method for measuring the distance of a stable matrix to the unstable matrices," *SIAM J. Sc. Stat. Comp.*, vol. 9, pp. 875–881, Sept. 1988.

[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] J. Sreedhar, P. Van Dooren, and A. L. Tits, "A fast algorithm to compute the real structured stability radius," in *Stability Theory: Proceedings of Hurwitz Centenary Conference, Ticino, Switzerland, May 21–26, 1995* (R. Jeltsch, M. Mansour, ed.), Birkhauser Verlag AG, pp. 219-230, 1996.

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