The distance between regularity and strong regularity *

Siegfried M. Rump

Abstract

In this paper we give very sharp bounds for the distance between regularity and strong regularity. The solution of this problem uses a general matrix analogue [18] of the Perron-Frobenius Theory. This theory has been graded as a challenge problem for future research by the International Linear Algebra Society [7].

An interval matrix $[A] \in \mathbb{IM}_n(\mathbb{R})$ is called *regular*, if every $A \in [A]$ is nonsingular, whereas $[A]$ is called *strongly regular*, if $M := \mathrm{mid}\big([A]\big) \in M_n(\mathbb{R})$ is nonsingular and $\rho(M^{-1} \cdot \mathrm{rad}\big([A]\big) < 1$. Strong regularity implies regularity.

Consider a system of linear equations, the data of which are afflicted with tolerances. The solution complex $\Sigma\big([A],[b]\big) := \{\, x \in \mathbb{R}^n \;\big|\; \exists\, A \in [A] \;\exists\, b \in [b] : Ax = b \,\}$ is bounded if $[A]$ is regular. Self-validating algorithms provide methods to compute an inclusion of the solution complex. However, many of those, e.g. the methods based on preconditioning and the Krawczyk operator, require $[A]$ to be strongly regular [9], [11], [16]. This raises the question: "How far is strong regularity from regularity?" More precisely, let matrices A and nonnegative Δ be given and define

$$\rho_{\mathrm{sreg}}(A,\Delta) := \sup\{\, r \in \mathbb{R} \;\big|\; [A - r\cdot\Delta, A + r\cdot\Delta] \quad \text{strongly regular} \,\}$$

and

$$\rho_{\mathrm{reg}}(A,\Delta) := \sup\{\, r \in \mathbb{R} \;\big|\; A - r\cdot\Delta, A + r\cdot\Delta] \quad \text{regular} \,\},$$

where the values may range within $[0,\infty]$. Then the question is: are there finite bounds for the ratio $\rho_{\mathrm{reg}}(A,\Delta)/\rho_{\mathrm{sreg}}(A,\Delta)$ independent on A and Δ and only depending on the dimension? And if so, how sharp are the bounds? In this note we present an analysis of this question and, up to a small constant factor, a complete answer.

0 Introduction

Let $\mathbb{IR}^n$ denote the set of real interval vectors with n components, let $M_n(\mathbb{R})$ and $\mathbb{IM}_n(\mathbb{R})$ denote the set of $n \times n$ matrices with real and real interval entries, respectively.

*In G. Alefeld, A. Frommer, and B. Lang, editors, Scientific Computing and Validated Numerics, volume 90 of Mathematical research, pp. 105–117, 1996; Part of the invited talk given at the SCAN-95 conference in Wuppertal; the other parts will be published in the papers "Improved Iteration Schemes for Validation Algorithms for Dense and Sparse Nonlinear Systems", to appear in COMPUTING, "Expansion and Estimation of the Range of Nonlinear Functions", to appear in MATH. of COMPUTATION, and "Inclusion of zeros of nowhere differentiable n-dimensional functions", submitted to RELIABLE COMPUTING.

Let some $[z] \in \mathbb{IR}^n$ and $[C] \in \mathbb{IM}_n(\mathbb{R})$ be given. For some $[X^0] \in \mathbb{IR}^n$ define the iteration

$$[Y^k] := [X^0] + [-\varepsilon, +\varepsilon]; \quad [X^{k+1}] := [z] + [C] \cdot [Y^k] \quad \text{for } 0 \le k \in \mathbb{N}, \tag{1}$$

where $0 < \varepsilon \in \mathbb{R}$ is some constant. All operations in (1) are interval operations (cf. [1], [12]). It is well known (cf. [16]) that the following statements are equivalent.

$$\begin{array}{rl} i) & \exists\, k \in \mathbb{N}: \quad [X^{k+1}] \subseteq \operatorname{int}\big([Y^k]\big). \\ ii) & \rho\Big(\big|[C]\big|\Big) < 1. \end{array} \tag{2}$$

The absolute value of an interval matrix is defined by $\big|[C]\big|_{ij} := \max\big\{|c| \;\big|\; c \in [C]_{ij}\big\}$. Note that the equivalence is independent of the choice of ε and of $[X^0]$. Also, the ε-inflation is crucial for the equivalence statement.

Our observation has a well known application to the numerical computation of enclosures of $\Sigma\big([A],[b]\big)$. For some preconditioner $R \in M_n(\mathbb{R})$ and some approximate solution $\tilde{x}$ (e.g. of the midpoint system) set $[z] := R \cdot \big([b] - [A] \cdot \tilde{x}\big)$ and $[C] := I - R \cdot [A]$. Then

$$[z] + [C] \cdot [X] \subseteq \operatorname{int}\big([X]\big)$$

for some $[X] \in \mathbb{IR}^n$ implies

$$[A] \quad \text{is regular and} \quad \Sigma\big([A],[b]\big) \subseteq \tilde{x} + [X]$$

(cf. [16] and papers cited over there). In view of (2) this has two interesting implications:

I) If $\big|I - R \cdot [A]\big|$ is convergent, then we will obtain a validated inclusion of the solution complex $\Sigma\big([A],[b]\big)$ after a finite number of steps.

On the other hand we know that the precise midpoint inverse is the optimal preconditioner ([12], Chapter 4 and [14]) and $\rho\Big(\big|[C]\big|\Big) < 1$ implies strong regularity of $[A]$. This leads to the second implication:

II) The ansatz for computing a validated inclusion of the solution complex $\Sigma\big([A],[b]\big)$ will fail if $[A]$ is not strongly regular.

This means, for $R := \operatorname{mid}\big([A]\big)^{-1}$ we have the following dichotomy:

Either, $[A]$ is strongly regular and *any* starting interval $[X^0]$ and *any* value of ε will lead to validated bounds for $\Sigma\big([A],[b]\big)$,

or, the method will fail for *any* $[X^0]$ and ε.

This raises two questions. First, are there other methods which compute validated bounds for $\Sigma\big([A],[b]\big)$ for $[A]$ *not* being strongly regular? Second, "how far" is strong regularity of $[A]$ from regularity?

A method answering the first question in the affirmative is used for large systems of linear and nonlinear equations. It uses

$$\sigma_n\Big(\mathrm{mid}\big([A]\big)\Big) > \sigma_1\Big(\mathrm{rad}\big([A]\big)\Big) \quad \Rightarrow \quad [A] \quad \text{is regular,} \tag{3}$$

where σ_i denotes the i-th singular value of a matrix in decreasing order. It can be shown ([17]) that there are examples of interval matrices not being strongly regular where the radius matrix can be amplified by a factor up to $n^{1/2}$, where regularity can still be verified using (3). On the other hand, the criterion (3) can be arbitrarily weak compared to strong regularity [17].

The method proposed by Shary [20] *assumes* regularity of the interval matrix. The method proposed by Jansson [8] seems promising to calculate the true interval hull of the solution complex of an interval linear system. It proves regularity of an interval matrix, and it has polynomially bounded computing time in the number of orthants with nonempty intersection with the solution set of the linear interval system. The worst case computing time, however, is exponential in the number of unknowns. It is very likely, unless $NP = P$, that no polynomially bounded algorithm exists, because Poljak and Rohn showed that checking regularity of a general interval matrix is NP-hard [13].

Now to answer the second question. It turns out that this question has much attention in modern numerical analysis. The buzz word in this context is componentwise distances and componentwise error estimates. Consider the componentwise distance of some matrix $A \in M_n(\mathbb{R})$ to the nearest singular matrix weighted by some nonnegative weight matrix $E \in M_n(\mathbb{R})$:

$$\sigma(A,E) := \min\{\, \alpha \in \mathbb{R} \;\big|\; \exists\ \tilde{E} \in M_n(\mathbb{R}) : |\tilde{E}| \le \alpha \cdot E \quad \text{and } A + \tilde{E} \text{ singular}\}\ . \tag{4}$$

If no such α exists, we set $\sigma(A,E) := \infty$. Absolute value and order relation are to be understood componentwise. It is $\sigma(A,E) = \rho_{\mathrm{reg}}(A,E)^{-1}$. Such a *componentwise* distance is of high practical interest. For a normwise distance we have explicit formulas, for example

$$\min\{\, \|\Delta\|_2 \;\big|\; A + \Delta \text{ singular}\,\} = \sigma_n(A).$$

This can be extended to other norms (cf. [21]), and also a corresponding Δ, which is of rank 1, can be explicitly calculated. However, a normwise distance favours large matrix entries. Consider Hamming's example [5]

$$A = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2\varepsilon & 2\varepsilon \\ 1 & 2\varepsilon & -\varepsilon \end{pmatrix} \quad \Rightarrow \quad \sigma_n(A) \approx 2 \cdot \varepsilon \ , \quad \text{but} \quad \sigma\big(A, |A|\big) > 0.3.$$

That means, a very small normwise perturbation suffices to run into a singular matrix, but any relative perturbations of the entries of A less than 30 % produces only regular matrices. Moreover, specific components of A like system zeros or constants may be left unaltered by the componentwise approach by setting the corresponding entry in the weight matrix E to zero.

The connection between our question and $\sigma(A, E)$ is the following. Let $A \in M_n(\mathrm{I\!R})$ and nonnegative $E \in M_n(\mathrm{I\!R})$ be given, Then $\alpha := \sigma(A, E)$ is the smallest value such that $[A - \alpha \cdot E, A + \alpha \cdot E]$ is singular. On the other hand, for $\beta := \rho\big(|A^{-1}| \cdot E\big)^{-1}$, any matrix $[A - \beta' \cdot E, A + \beta' \cdot E]$ with $\beta' < \beta$ is strongly regular. The ratio $\sigma(A, E)/\rho\big(|A^{-1}| \cdot E\big)^{-1}$ can therefore be interpreted as the distance between strong regularity and regularity. There is another interesting connection to the condition number of A. In a normwise sense, an ill-conditioned matrix is “not too far” from a singular matrix. For a componentwise distance we need a condition number reflecting the weight matrix E. This is the Bauer-Skeel condition number

$$\mathrm{cond}_{BS}(A, E) := \| \, |A^{-1}| \cdot E \|$$

for some matrix norm $\| \cdot \|$. The Bauer-Skeel condition number may be large due to improper scaling. Therefore we consider the *optimal* Bauer-Skeel condition number

$$\mathrm{cond}_{OBS}(A, E) := \inf_D \ \mathrm{cond}_{BS}(AD, ED),$$

where D is a nonsingular diagonal matrix. Interestingly, there is an explicit formula for this number for all p-norms [4]:

$$\mathrm{cond}_{OBS}(A, E) = \rho\big(|A^{-1}| \cdot E\big). \tag{5}$$

This is exactly the inverse of the supremum of all β' such that $[A - \beta' \cdot E, A + \beta' \cdot E]$ is strongly regular. Any strongly regular matrix is regular, and therefore

$$\frac{1}{\rho\big(|A^{-1}| \cdot E\big)} \le \sigma(A, E).$$

The question of distance between strong regularity and regularity, or the ratio $\sigma(A, E)/\rho\big(|A^{-1}| \cdot E\big)^{-1}$, has therefore the following interpretation: If the optimal Bauer-Skeel condition number is large, is it then true that not too far away in a componentwise sense there exists a singular matrix? This is the pendant to the corresponding statement for norms. That means, the question is whether there exist real constants $\gamma(n) \in \mathrm{I\!R}$, only depending on the dimension, with

$$\frac{1}{\rho\big(|A^{-1}| \cdot E\big)} \le \sigma(A, E) \le \frac{\gamma(n)}{\rho\big(|A^{-1}| \cdot E\big)} \tag{6}$$

for all A, E. For relative perturbations this has been conjectured by J. Demmel and N. J. Higham [4]. We define $\gamma(n)$ by

$$\gamma(n) := \sup\{\, \sigma(A,E) \cdot \rho\big(|A^{-1}| \cdot E\big) \;\big|\; A \in M_n(\mathbb{R}),\ 0 \le E \le M_n(\mathbb{R}) \,\}\ . \tag{7}$$

We show that $\gamma(n)$ is well defined and is finite, and that they therefore satisfy (6). Furthermore, we give very sharp lower and upper bounds for $\gamma(n)$. The bounds differ only by a small constant factor.

The major difficulty in finding bounds for $\gamma(n)$ is that we need *upper bounds* α for $\sigma(A,E)$, thus proving existence of a singular matrix within $[A - \alpha \cdot E, A + \alpha \cdot E]$. Note that computation of $\sigma(A,E)$ is NP-hard [13].

1 A Perron-Frobenius Theory for general matrices

For singular $A \in M_n(\mathbb{R})$, we have $\sigma(A,E) = 0$ for any weight matrix E, and there is nothing to prove. Therefore, we assume A to be regular for this chapter. For $e \in M_n(\mathbb{R})$, $|e| \le E$, there holds

$$A - e = A \cdot (I - A^{-1} \cdot e). \tag{8}$$

That means, singularity of $A - e$ is equivalent to the fact that $A^{-1} \cdot e$ has the eigenvalue 1. Let $\rho_0(A)$ denote the *real* spectral radius of A (cf. [15]), i.e.

$$\rho_0(A) := \max\{\, |\lambda| \;\big|\; \lambda \textit{ real} \text{ eigenvalue of } A \,\}.$$

If A has no real eigenvalue, we define $\rho_0(A) := 0$. Then (8) implies

$$\sigma(A,E) = \{\, \max_{|e| \le E} \ \rho_0(A^{-1}e) \,\}^{-1}. \tag{9}$$

Formula (9) is not suitable for a computation of $\sigma(A,E)$, because infinitely many matrices $|e| \le E$ have to be checked. There is an explicit formula by Rohn, where only finitely many matrices have to be checked [15]:

$$\sigma(A,E) = \{\, \max_{S_1,S_2} \ \rho_0(S_1 A^{-1} S_2 E) \,\}^{-1}. \tag{10}$$

Here $S_1, S_2 \in \mathcal{S}$, the set of all signature matrices, that is diagonal matrices with diagonal entries $+1$ or -1. In short notation $S \in \mathcal{S} \quad \Leftrightarrow \quad |S| = I$. The exponential number of matrices to be checked in (10) corresponds to the NP-hardness of computation of $\sigma(A,E)$.

For our purpose, for the estimation of the distance between strong regularity and regularity of an interval matrix, it turns out that (9) is useful. However, not in this form because the real spectral radius is an unpleasant number. For example, it is not continuous in the entries of A because real eigenvalues may become complex under arbitrary small perturbations. Consider $A(\varepsilon) := \begin{pmatrix} 1 & -\varepsilon \\ 1 & 1 \end{pmatrix}$. Then $\rho_0\big(A(\varepsilon)\big) = 0$ for any $\varepsilon > 0$, whereas $\rho\big(A(0)\big) = 1$.

We define the *signature-real* spectral radius by

$$\rho_0^S(A) := \max_{S \in \mathcal{S}} \ \rho_0(S \cdot A) \ . \tag{11}$$

This number turns out to have very interesting (and pleasant) properties, and it helps to solve our problem. In fact, the signature-real spectral radius allows to extend a number of properties of the Perron root (the spectral radius) of a nonnegative matrix to an arbitrary matrix. We first mention some basic properties of $\rho_0^S(A)$. The corresponding theorems are proved in [19].

For $A \in M_n(\mathbb{R})$, signature matrices $S_1, S_2 \in \mathcal{S}$, permutation matrix P and regular diagonal matrix D there holds

$$\begin{array}{l} \rho_0^S(A) = \rho_0^S(S_1 A S_2) = \rho_0^S(A^T) = \rho_0^S(P^T A P) = \rho_0^S(D^{-1} A D), \\ \rho_0^S(AD) = \rho_0^S(DA) \quad \text{and} \quad \rho_0^S(\alpha A) = |\alpha| \cdot \rho_0^S(A) \quad \text{for} \quad \alpha \in \mathbb{R}. \end{array} \tag{12}$$

For lower or upper triangular A it is $\rho_0^S(A) = \max\limits_i \ |A_{ii}|$. On the other hand, even for orthogonal Q, in general $\rho_0^S(Q^T A Q) \neq \rho_0^S(A)$. Also, in general $\rho_0^S(AB) \neq \rho_0^S(BA)$.

But it can be proved [18] that the signature-real spectral radius depends *continuously* on the entries of A. Furthermore, we can characterize $\rho_0^S(A) = 0$ by

$$\rho_0^S(A) = 0 \quad \Leftrightarrow \quad \begin{array}{l} A \text{ is permutationally similar to a strictly} \\ \text{upper triangular matrix.} \end{array}$$

Moreover, there exist always signature matrices S_1, S_2 such that $x \in \mathbb{R}^n$, $x \geq 0$ is eigenvector of $S_1 A S_2$ to the eigenvalue $\rho_0^S(A)$:

$$S_1 A S_2 \cdot x = \rho_0^S(A) \cdot x.$$

Another characterisation is

$$\rho_0^S(A) = \inf\{\, 0 \leq b \in \mathbb{R} \;\big|\; b \cdot I - S \cdot A \quad \text{is } P\text{-matrix for all} \quad S \in \mathcal{S}\,\},$$

where $B \in M_n(\mathbb{R})$ is called P-matrix, if all principle minors are positive. The Perron root of a nonnegative matrix has the inheritance property, that is it cannot increase when going to a principle submatrix. The same is true for the signature-real spectral radius:

$$\rho_0^S(A) \geq \rho_0^S\big(A[\omega]\big) \quad \text{for all} \quad \omega \in Q_{kn}, 1 \leq k \leq n \tag{13}$$

(we use standard notation from matrix theory, cf. [6], [10]: Q_{kn} denotes the set of strictly increasing sequences of k integers out of $\{1, \ldots, n\}$, and for $\omega \in Q_{kn}$, $A[\omega]$ is the $k \times k$ principle submatrix of A with entries A_{ij}, $i, j \in \omega$). Especially,

$$\rho_0^S(A) \geq \max_i \ |A_{ii}|. \tag{14}$$

The following explicit relation between the componentwise distance to the nearest singular matrix and the signature-real spectral radius

$$\sigma(A,E) = \rho_0^S \begin{pmatrix} 0 & E \\ A^{-1} & 0 \end{pmatrix}^{-2}$$

shows that the computation of ρ_0^S is also NP-hard. $\rho_0^S(A) \leq \rho(A)$ is not true in general, but

$$\rho_0^S(A) \leq \|A\|_p \text{ for } 1 \leq p \leq \infty, \quad \max_{Q \text{ orthogonal}} \rho_0^S(QA) = \|A\|_2 \quad \text{and}$$
$$A = A^T \quad \Rightarrow \quad \rho_0^S(A) = \|A\|_2 = \rho(A).$$

We can prove [18] the following max min characterisation, which is almost exactly the same as for nonnegative matrices [3]:

$$\rho_0^S(A) = \max_{x \in \mathbb{R}^n} \min_{x_i \neq 0} \left| \frac{(Ax)_i}{x_i} \right|. \tag{15}$$

As for general nonnegative matrices the corresponding min max equality is, in general, not true. The characterisation (15) will be the key to the solution of our problem. For nonnegative A, decreasing a single entry cannot increase the Perron root. The corresponding property for $\rho_0^S(A)$ is

$$\rho_0^S(A + \alpha \cdot e_i e_j^T) \geq \rho_0^S(A)$$

either for all $\alpha \geq 0$, or for all $\alpha \leq 0$ and any $1 \leq i, j \leq n$.

There are many more interesting properties of the signature-real spectral radius. For details the reader is referred to [18].

2 The distance between strong regularity and regularity

Our main result can be derived by combining the results of the previous chapter. From (9), (12) and (15) we know for regular A and any nonnegative E,

$$\begin{aligned} \sigma(A,E) = & \frac{1}{\max\limits_{|\tilde{E}| \leq E} \rho_0(A^{-1}\tilde{E})} = \frac{1}{\max\limits_{|\tilde{E}| \leq E} \rho_0^S(A^{-1}\tilde{E})} = \\ = & \frac{1}{\max\limits_{|\tilde{E}| \leq E} \max\limits_{x \in \mathbb{R}^n} \min\limits_{x_i \neq 0} \left| \dfrac{(A^{-1}\tilde{E}x)_i}{x_i} \right|} . \end{aligned} \tag{16}$$

The key is to construct a proper matrix $\tilde{E}$ and proper x in order to obtain an upper bound for $\sigma(A, E)$.

For this purpose we derive lower bounds for the signature real spectral radius of A using cycles of A, which in turn will imply an upper bound for $\sigma(A, E)$. A *cycle* $\omega = (i_1, \ldots, i_k)$ is a collection of mutually different elements out of $\{1, \ldots, n\}$. The cycle ω defines a *cycle product* $\Pi_\omega A := A_{i_1 i_2} \cdot \ldots \cdot A_{i_{k-1} i_k} \cdot A_{i_k i_1}$ on A. We denote $|\omega| := k$, such that $|\Pi_\omega A|^{1/|\omega|}$ is the geometric mean of the elements of the cycle ω of A.

A key observation is that for any cycle ω, with proper choice of a matrix $\tilde{E}$ on the boundary of $[-E, +E]$, a matrix $\tilde{C} := A^{-1}\tilde{E}$ can be obtained with $|\tilde{C}| \leq |A^{-1}| \cdot E =: C$ and $|\tilde{C}_\omega| = C_\omega$. That means, for $\omega := \{i_1, \ldots, i_k\}$ there holds $|\tilde{C}_{i_1 i_2}| = C_{i_1 i_2}, \ldots, |\tilde{C}_{i_{k-1} i_k}| = C_{i_{k-1} i_k}$ and $|\tilde{C}_{i_k i_1}| = C_{i_k i_1}$. Moreover, we are free in the choice of the signs of the elements $\tilde{C}_{i_\nu i_{\nu+1}}$ of a cycle. Next we need the following lower bound on the signature real spectral radius based on the geometric mean of cyclic products.

Lemma 1. Let $A \in M_n(\mathrm{I\!R})$ and $\omega := \{i_1, \ldots, i_k\} \in \{1, \ldots, n\}^k$ be given. Denote the geometric mean of the elements of $|A|$ defined by the cycle ω by $|\Pi_\omega A|^{1/|\omega|}$. Then

$$\rho_0^S(A) \geq (3 + 2 \cdot \sqrt{2})^{-1} \cdot |\Pi_\omega A|^{1/|\omega|}.$$

It is remarkable that the constant $(3 + 2\sqrt{2})^{-1}$ in Lemma 1 does not depend on k or n. The constant can be improved for specific values of k, for example

$$\rho_0^S(A) \geq |\Pi_\omega A|^{1/|\omega|} \quad \text{for } |\omega| \in \{1, 2\}. \tag{17}$$

For $|\omega| = 1$ this has been observed in (14). Note that (17) is not valid for $|\omega| \geq 3$. Consider (cf. [17])

$$A := \begin{pmatrix} -0.3 & 1 & -0.8 \\ -0.8 & -0.3 & 1 \\ 1 & -0.8 & -0.3 \end{pmatrix} \qquad \text{and} \qquad \omega := \{1, 2, 3\}.$$

Then $\Pi_\omega A = 1$, whereas $\rho_0^S(A) < 0.95$.

Now we have the tools to prove bounds for the distance between strong regularity and regularity. First, we observe that $\sigma(A, E)$ depends continuously on the entries of A and E (cf. [17]). This allows to transform A and E in such a way that $|A^{-1}| \cdot E$ is row stochastic. Next, we can prove that there exists a cycle of $|A^{-1}| \cdot E$ with geometric mean not less than n^{-1}. Combining this with Lemma 1 and the previous results we obtain the following theorem.

Theorem 2. Let $A, E \in M_n(\mathrm{I\!R})$, A regular and $E \geq 0$. Then

$$\frac{1}{\rho\big(|A^{-1}| \cdot E\big)} \leq \sigma(A, E) \leq (3 + 2\sqrt{2}) \cdot n \cdot \frac{1}{\rho\big(|A^{-1}| \cdot E\big)} .$$

Furthermore, the constants $\gamma(n)$ in (7) satisfy

$$n \le \gamma(n) \le (3 + 2\sqrt{2}) \cdot n. \tag{18}$$

The lower bound in (18) cannot be improved for every $n \in \mathbb{N}$.

In terms of interval analysis and self-validating algorithms this result means the following. Given an interval matrix $[A - \Delta, A + \Delta]$, which is not strongly regular. That means, self-validating algorithms based on the methods described in Chapter 0 cannot work. Then increasing the radius by at most the factor $(3 + 2\sqrt{2}) \cdot n$ produces an interval matrix which is not regular.

The interpretation in terms of traditional numerical analysis is that the inverse of the optimal Bauer-Skeel condition number is a lower bound for the componentwise distance to the nearest singular matrix weighted by E, whereas $(3 + 2\sqrt{2}) \cdot n$ times this number is an upper bound:

$$\frac{1}{\rho\big(|A^{-1}| \cdot E\big)} \le \sigma(A, E) \le \frac{(3 + 2\sqrt{2}) \cdot n}{\rho\big(|A^{-1}| \cdot E\big)}.$$

In a practical application this bound can be improved by using Lemma 1 and the following bounds for smaller values of n.

Theorem 3. Let $A, E \in M_n(\mathbb{R})$, A regular and $E \ge 0$. Then

$$\frac{1}{\rho\big(|A^{-1}| \cdot E\big)} \le \sigma(A, E) \le \frac{\psi_{|\omega|}}{(\Pi_\omega \ |A^{-1}| \cdot E)^{1/|\omega|}} \tag{19}$$

for any cycle ω, where (cf. [19])

$$\psi_1 = 1, \quad \psi_2 = 2, \quad \psi_3 = 1.59, \quad \psi_4 = 1.97, \quad \psi_5 = 2.30 \ldots$$

In its simplest form, (19) reads for $|\omega| = 1$ and $|\omega| = 2$ as follows.

Corollary 4. Let $A, E \in M_n(\mathbb{R})$, A regular and $E \ge 0$. Then

$$\frac{1}{\rho(C)} \le \sigma(A, E) \le \frac{1}{\max\limits_i \ |C_{ii}|} \quad \text{and} \tag{20}$$

$$\frac{1}{\rho(C)} \le \sigma(A, E) \le \frac{1}{\max\limits_{i,j} \ \sqrt{|C_{ij} \cdot C_{ji}|}}\ , \tag{21}$$

where $C := |A^{-1}| \cdot E$.

The factors in (20) and (21) are easy to calculate, and frequently they give reasonable bounds for the distance between regularity and strong regularity. It can be combined with the following observation [17]. Let $S \in M_n(\mathbb{R})$ be a matrix with

$$S_{ij} := \begin{cases} +1 & \text{if } (A^{-1})_{ij} > 0 \\ -1 & \text{if } (A^{-1})_{ij} < 0 \\ -1 \text{ or } +1 & \text{otherwise} \end{cases} .$$

In other words, S is the matrix of signs of the elements of A^{-1}, where the sign of 0 can be interpreted to be -1 or $+1$. The latter gives possible freedom in choosing S. For *any* such matrix S there holds

$$\operatorname{rank}(S) = 1 \quad \Rightarrow \quad \sigma(A, E) = \Big\{ \rho\big(|A^{-1}| \cdot E\big) \Big\}^{-1} . \tag{22}$$

This is true, for example, for M-matrices. In other words, for M-matrices regularity and strong regularity are identical.

As an example, we estimated the distance between regularity and strong regularity for all interval matrices in the test examples in the paper by Shary [20]. We used (20) and (21), and possibly (22) for the estimation.

Shary's first example is taken from [2]

$$\begin{pmatrix} [2,4] & [-2,1] \\ [-1,2] & [2,4] \end{pmatrix} . \tag{23}$$

His second example is the same with a different right hand side; the next example is taken from [20]

$$\begin{pmatrix} [2,4] & [-5,-1] & [-2,3] \\ [-3,1] & [5,7] & [4,6] \\ [-1,1] & [-2,1] & [-7,-2] \end{pmatrix} . \tag{24}$$

The other examples are

$$\begin{pmatrix} [3,4] & [-5,-2] & [-2,2] \\ [-3,-1] & [6,7] & [5,6] \\ [-1,0] & [-1,1] & [-4,1] \end{pmatrix} , \tag{25}$$

and finally

$$\begin{pmatrix} [4,6] & [-9,0] & [0,12] & [2,3] & [5,9] & [-23,-9] & [15,23] \\ [0,1] & [6,10] & [-1,1] & [-1,3] & [-5,1] & [1,15] & [-3,-1] \\ [0,3] & [-20,-9] & [12,77] & [-6,30] & [0,3] & [-18,1] & [0,1] \\ [-4,1] & [-1,1] & [-3,1] & [3,5] & [5,9] & [1,2] & [1,4] \\ [0,3] & [0,6] & [0,20] & [-1,5] & [8,14] & [-6,1] & [10,17] \\ [-7,-2] & [1,2] & [7,14] & [-3,1] & [0,2] & [3,5] & [-2,1] \\ [-1,5] & [-3,2] & [0,8] & [1,11] & [-5,10] & [2,7] & [6,82] \end{pmatrix} . \tag{26}$$

First, we set $E := \mathrm{rad}(A)$, i.e. we consider the original examples given in Shary's paper. The following table shows from left to right the

number of the example,
$\Big\{\rho\big(|A^{-1}| \cdot E\big)\Big\}^{-1}$, the lower bound for $\sigma\big(A, E\big)$,
$\sigma\big(A, E\big)$, if known,
the upper bound β_1 for $\sigma\big(A, E\big)$ by criterion in (20),
the upper bound β_2 for $\sigma\big(A, E\big)$ by criterion in (21).

example	$\Big\{\rho\big(\vert A^{-1}\vert \cdot E\big)\Big\}^{-1}$	$\sigma\big(A, E\big)$	β_1	β_2
(23)	1.06	1.23	2.47	1.85
(24)	0.38	0.38	0.85	0.85
(25)	0.15	0.15	0.30	0.30
(26)	0.11		0.41	0.41

Table 1. Estimated distance between regularity and strong regularity

Except in the first example, the bounds β_1 and β_2 are equal. In examples (24) and (25) we could use (22) to show equality between regularity and strong regularity. We did not calculate $\sigma\big(A, E\big)$ for the last example, but the bounds show that regularity and strong regularity cannot be too far apart. Note that the ratio between β_1, β_2 and $\Big\{\rho\big(|A^{-1}| \cdot E\big)\Big\}^{-1}$ is better than $n = 7$, which is, due to Theorem 2, an achievable ratio.

Next, we treated the same examples for relative perturbations, i.e. $E = |A|$.

example	$\Big\{\rho\big(\vert A^{-1}\vert \cdot \vert A\vert\big)\Big\}^{-1}$	$\sigma\big(A, \vert A\vert\big)$	β_1	β_2
(23)	0.76	1	1	1
(24)	0.32	0.32	0.58	0.58
(25)	0.11	0.11	0.26	0.26
(26)	0.07		0.24	0.24

Table 2. Estimated distance between regularity and strong regularity for relative perturbations

The same examples for absolute perturbations, i.e. $E_{ij} = 1$ for all i, j, look as follows.

example	$\left\{\rho\left(\lvert A^{-1}\rvert \cdot (1)\right)\right\}^{-1}$	$\sigma\left(A,(1)\right)$	β_1	β_2
(23)	1.32	1.54	2.64	2.64
(24)	0.56	0.56	1.06	1.06
(25)	0.16	0.16	0.41	0.41
(26)	0.30		1.09	1.09

Table 3. Estimated distance between regularity and strong regularity for absolute perturbations

The amount of overestimation is comparable. In the last example the difference between $\left\{\rho\left(\lvert A^{-1}\rvert \cdot E\right)\right\}^{-1}$ and β_1, β_2 is not too big, as it was the case for relative perturbations. As before, (22) is applicable to examples (24) and (25). Note that validity of the assumptions of (22) is *independent* of the weight matrix E.

For test matrices like Hilbert matrices

$$H_{ij} := 1/(i+j-1),$$

Pascal matrices

$$P_{ij} := \binom{i+j}{i} \quad \text{or} \quad P_{ij} = \binom{i+j-1}{i},$$

Boothroyd matrices

$$B_{ij} := \frac{\binom{n+i-1}{i-1} \cdot n \cdot \binom{n-1}{n-j}}{i+j-1},$$

the criterion (22) applies, that means regularity and strong regularity are identical for all of those matrices. This is also true for the inverses of all of those matrices.

Finally, we tested randomly generated matrices. The entries A_{ij} are chosen randomly out of $[-1,1]$ with uniform distribution. The results are as follows, where in the left half of Table 3 we chose $E := \lvert A\rvert$, whereas in the right half it is $E := (1)$. The ratio is $(\max C_{ii})^{-1}/\left(\rho\left(\lvert A^{-1}\rvert \cdot E\right)^{-1}\right)$. For each dimension we calculated 10 test cases and display the arithmetic mean of the results.

	relative perturbations			absolute perturbations		
n	$\rho\left(\lvert A^{-1}\rvert \cdot E\right)^{-1}$	$(\max C_{ii})^{-1}$	ratio	$\rho\left(\lvert A^{-1}\rvert \cdot E\right)^{-1}$	$(\max C_{ii})^{-1}$	ratio
10	2.24e-02	1.19e-01	5.1	1.14e-02	6.59e-02	5.4
20	8.40e-03	8.83e-02	9.5	4.10e-03	3.99e-02	9.1
50	2.97e-03	6.42e-02	21.5	1.48e-03	3.25e-02	22.0
100	8.74e-04	3.82e-02	38.3	4.34e-04	1.86e-02	37.3
150	4.97e-04	2.85e-02	52.6	2.48e-04	1.42e-02	52.8
200	1.97e-04	1.31e-02	64.6	9.87e-05	6.50e-03	64.0

Table 3. Estimated distance between regularity and strong regularity for random matrices

The bounds achieved by estimation (21) are almost identical with the displayed numbers achieved by the bound (20).

Finally, we note that in many other examples (20) gave already reasonable bounds for the distance between regularity and strong regularity, although, as has been mentioned, the bound may be arbitrarily weak.

References

[1] G. Alefeld and J. Herzberger. *Introduction to Interval Computations.* Academic Press, New York 1983.

[2] W. Barth and E. Nuding. Optimale Lösung von Intervallgleichungssystemen. *Computing 12*, pages 117–125, 1974.

[3] L. Collatz. Einschließungssatz für die charakteristischen Zahlen von Matrizen. *Math. Z.*, 48:221–226, 1942.

[4] J. Demmel. The Componentwise Distance to the Nearest Singular Matrix. *SIAM J. Matrix Anal. Appl.*, 13(1):10–19, 1992.

[5] R. Hamming. *Introduction to Applied Numerical Analysis.* McGraw Hill, New York, 1971.

[6] R. A. Horn and Ch. Johnson. *Matrix Analysis.* Cambridge University Press, 1985.

[7] Call for "Challenges in Matrix Theory". LAA Vol. 233, p. 1 (1996)

[8] C. Jansson. Calculation of Exact Bounds for the Solution Set of Linear Interval Equations, to appear in LAA.

[9] R. Krawczyk. Newton-Algorithmen zur Bestimmung von Nullstellen mit Fehlerschranken. *Computing 4*, pages 187–201, 1969.

[10] M. Marcus and H. Minc. *A survey of matrix theory and matrix inequalities.* Dover publications, New York, 1992.

[11] R. E. Moore. A Test for Existence of Solutions for Non-Linear-Systems. *SIAM J. Numer. Anal. 4*, 1977.

[12] A. Neumaier. *Interval Methods for Systems of Equations, Encyclopedia of Mathematics and its Applications.* Cambridte University Press, 1990.

[13] S. Poljak and J. Rohn. Checking Robust Nonsingularity Is NP-Hard. *Math. of Control, Signals, and Systems 6*, pages 1–9, 1993.

[14] G. Rex and J. Rohn. A Note on Checking Regularity of Interval Matrices. *Linear and Multilinear Algebra 39*, pages 259–262, 1995.

[15] J. Rohn. Systems of Linear Interval Equations. *Linear Algebra Appl. 126*, pages 39–78, 1989.

[16] S. M. Rump. New Results on Verified Inclusions. In W. L. Miranker and R. Toupin, editors, *Accurate Scientific Computation*, pages 31–69. Springer Lecture Notes in Computer Science 235, 1986.

[17] S. M. Rump. Bounds for the Componentwise Distance to the Nearest Singular Matrix. to appear in SIMAX.

[18] S. M. Rump. Perron-Frobenius-like Theorems for not sign-restricted Matrices. submitted for publication.

[19] S. M. Rump. Almost sharp bounds for the componentwise distance to the nearest singular matrix. submitted for publication.

[20] S. P. Shary. Algebraic approach to the interval linear static identification, tolerance and control problems, to appear in Reliable Computing (1996).

[21] G. W. Stewart and J. Sun. *Matrix Perturbation Theory.* Academic Press, 1990.

Addresses:

S. M. Rump, Department of Computer Science III, Technical University Hamburg-Harburg, Eißendorfer Straße 38, 21071 Hamburg, Germany.