

ON THE CONDITION OF A COMPLEX EIGENVALUE UNDER REAL PERTURBATIONS

R. BYERS^{1*} and D. KRESSNER^{2 †}

¹ *Department of Mathematics, University of Kansas, 405 Snow Hall, Lawrence, KS 66045.
email: byers@math.ukans.edu*

² *Institut für Mathematik MA 4-5, TU Berlin, D-10623 Berlin, Germany.
email: kressner@math.tu-berlin.de*

Abstract.

We investigate the condition number for a complex eigenvalue of a real matrix under real perturbations. Based on an explicit formula, it is shown that this number is never smaller than $1/\sqrt{2}$ times the corresponding condition number with respect to complex perturbations. This result can be generalized to the condition number of an arbitrary complex-valued function under real perturbations. This extends to related condition numbers.

AMS subject classification (2000): 65F15, 65F35.

Key words: Condition number, eigenvalue.

1 Introduction

Data are typically corrupted by errors. There may be measurement errors, linearizations or other simplifications in a mathematical model, equipment may wear or be out of adjustment, operating conditions may vary, etc. . . . Numerical methods for computing the eigenvalues of a general matrix A are affected by rounding errors, which are a consequence of working in finite precision arithmetic. An algorithm is *backward stable* if it computes the *exact* eigenvalues and invariant subspaces of a perturbed matrix $A + E$, where $\|E\|_F \leq \varepsilon \|A\|_F$ with ε is not much larger than the unit roundoff and, for mathematical convenience, $\|M\|_F = \sqrt{\text{trace}(M^*M)}$ is the Frobenius matrix norm. Such backward rounding errors are typically much smaller than other *a priori* errors that may corrupt the data A .

It is of interest to bound the effect of the *backward error* E on the eigenvalues of A . The condition number of an eigenvalue λ provides a tight-to-first-order bound on the worst-case effect of a backward error on the particular eigenvalue λ . For complex eigenvalues, textbook derivations of this condition number implicitly assume that the backward error E may be a general complex matrix [1, 5]. However, in many contexts it is known with certainty that the matrix A is real

*This author was partially supported by National Science Foundation grants 0098150 and 0112375.

†This work was supported by the DFG Research Center “Mathematics for Key Technologies” (FZT 86) in Berlin.

and that errors or uncertainties in the data are also real. Common numerical methods like the QR algorithm and Krylov subspace methods, avoid complex arithmetic when A is real and so rounding errors also yield a real backward error E . The question arises whether the eigenvalue condition number with respect to real backward errors may be significantly smaller than the textbook condition number with respect to general complex backward errors. In this short paper, we show that restricting the backward error E to be real decreases the condition number at most by a factor of $1/\sqrt{2}$. This extends to related condition numbers, e.g., for eigenvectors and invariant subspaces.

2 Real and Complex Eigenvalue Condition Numbers

Our analysis is based on the perturbation expansion of a simple eigenvalue λ , which can be found, e.g., in [4]. Let $E \in \mathcal{B}(0)$ be a perturbation of $A \in \mathbb{R}^{n \times n}$, where $\mathcal{B}(0) \subset \mathbb{C}^{n \times n}$ is a sufficiently small open neighborhood of the origin. Then there exists a uniquely defined analytic function $f_\lambda : \mathcal{B}(0) \rightarrow \mathbb{C}$ so that $\lambda = f_\lambda(0)$ and $\hat{\lambda} = f_\lambda(E)$ is an eigenvalue of $A + E$. Moreover, we have the expansion

$$(2.1) \quad \hat{\lambda} = \lambda + \frac{1}{y^H x} y^H E x + \mathcal{O}(\|E\|_F^2),$$

where x and y are right and left eigenvectors belonging to λ normalized so that $\|x\|_2 = \|y\|_2 = 1$ and $y^H x = |y^H x|$.

Following [3], a norm-wise, absolute condition number for λ can be defined as follows.

$$(2.2) \quad c^{\mathbb{F}}(\lambda) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \sup \left\{ |\hat{\lambda} - \lambda| : E \in \mathbb{F}^{n \times n}, \|E\|_F \leq \varepsilon \right\},$$

where $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. The perturbation expansion (2.1) reveals

$$(2.3) \quad \begin{aligned} c^{\mathbb{F}}(\lambda) &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \sup \left\{ |y^H E x| / |y^H x| : E \in \mathbb{F}^{n \times n}, \|E\|_F \leq \varepsilon \right\} \\ &= 1/|y^H x| \cdot \sup \left\{ |y^H E x| : E \in \mathbb{F}^{n \times n}, \|E\|_F = 1 \right\}. \end{aligned}$$

This readily implies $c^{\mathbb{F}}(\lambda) \leq 1/|y^H x|$. If $\mathbb{F} = \mathbb{C}$, then $E = yx^H$ attains this upper bound showing that $c^{\mathbb{C}}(\lambda) = 1/|y^H x|$, which coincides with the textbook absolute condition number of an eigenvalue. For $\mathbb{F} = \mathbb{R}$ and real λ , this equality is still true as the eigenvectors x and y can be chosen to be real. For $\mathbb{F} = \mathbb{C}$ and real λ , however, this upper bound is generally not attained and $c^{\mathbb{R}}(\lambda)$ can be less than $c^{\mathbb{C}}(\lambda)$. The following theorem provides an explicit formula for $c^{\mathbb{R}}(\lambda)$ and a bound on how much $c^{\mathbb{C}}(\lambda)$ can exceed $c^{\mathbb{R}}(\lambda)$. Here, ' \otimes ' denotes the Kronecker product of two matrices and 'vec' is the operator which stacks the columns of a matrix into one long vector [1].

THEOREM 2.1. *Let $\lambda \in \mathbb{C}$ be a simple eigenvalue of $A \in \mathbb{R}^{n \times n}$ with normalized right and left eigenvectors $x = x_R + ix_I$ and $y = y_R + iy_I$, respectively, where $x_R, x_I, y_R, y_I \in \mathbb{R}^n$. Then the condition number $c^{\mathbb{R}}(\lambda)$ as defined in (2.2) satisfies*

$$c^{\mathbb{R}}(\lambda) = \frac{1}{|y^H x|} \sqrt{\frac{1}{2} + \sqrt{\frac{1}{4}(b^T b - c^T c)^2 + (b^T c)^2}},$$

where $b = x_R \otimes y_R + x_I \otimes y_I$ and $c = x_I \otimes y_R - x_R \otimes y_I$. In particular, we have the inequality

$$(2.4) \quad c^{\mathbb{R}}(\lambda) \geq c^{\mathbb{C}}(\lambda)/\sqrt{2}.$$

PROOF. Starting from (2.3), we have

$$(2.5) \quad \begin{aligned} c^{\mathbb{R}}(\lambda) &= 1/|y^H x| \cdot \sup_{\substack{E \in \mathbb{R}^{n \times n} \\ \|E\|_F = 1}} \left\| \begin{bmatrix} y_R^T E x_R + y_I^T E x_I \\ y_R^T E x_I - y_I^T E x_R \end{bmatrix} \right\|_2 \\ &= 1/|y^H x| \cdot \sup_{\substack{E \in \mathbb{R}^{n \times n} \\ \|\text{vec}(E)\|_2 = 1}} \left\| \begin{bmatrix} (x_R \otimes y_R + x_I \otimes y_I)^T \\ (x_I \otimes y_R - x_R \otimes y_I)^T \end{bmatrix} \text{vec}(E) \right\|_2. \end{aligned}$$

The maximum of the second factor is given by the largest singular value of the $n^2 \times 2$ matrix

$$(2.6) \quad X = \begin{bmatrix} x_R \otimes y_R + x_I \otimes y_I & x_I \otimes y_R - x_R \otimes y_I \end{bmatrix}.$$

A vector $\text{vec}(E)$ attaining the supremum in (2.5) is a left singular vector belonging to this singular value. The square of the largest singular value of X is given by the larger root θ_* of the polynomial

$$\det(X^T X - \theta I_2) = \theta^2 - (b^T b + c^T c)\theta + (b^T b)(c^T c) - (b^T c)^2.$$

Because the eigenvectors x and y are normalized, it can be shown by direct calculation that $b^T b + c^T c = 1$ and $1/4 - (b^T b)(c^T c) = 1/4 \cdot (b^T b - c^T c)^2$, implying

$$\theta_* = \frac{1}{2} + \sqrt{\frac{1}{4}(b^T b - c^T c)^2 + (b^T c)^2},$$

which concludes the proof. \square

For the matrix $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, we have $c^{\mathbb{R}}(\iota) = c^{\mathbb{R}}(-\iota) = 1/\sqrt{2}$ and $c^{\mathbb{C}}(\iota) = c^{\mathbb{C}}(-\iota) = 1$, revealing that the bound $c^{\mathbb{R}}(\lambda) \geq c^{\mathbb{C}}(\lambda)/\sqrt{2}$ can actually be attained. It is the use of the Frobenius norm in the definition (2.2) of $c^{\mathbb{R}}(\lambda)$ that leads to the esthetically displeasing effect that this condition number may become less than the norm of A .

Textbooks often define conditions numbers using the spectral norm $\|\cdot\|_2$. Define $c_2^{\mathbb{F}}(\lambda)$ by (2.2) with $\|E\|_F$ replaced by $\|E\|_2$. It follows from the inequality $\|E\|_2 \leq \|E\|_F$ (with equality for rank one matrices E) that

$$\frac{1}{\sqrt{2}}c^{\mathbb{C}}(\lambda) \leq c^{\mathbb{R}}(\lambda) \leq c_2^{\mathbb{R}}(\lambda) \leq c_2^{\mathbb{C}}(\lambda) = c^{\mathbb{C}}(\lambda).$$

3 Real-Part and Imaginary-Part Condition Numbers

In the case that the perturbation E is real, (2.1) may be expressed in terms of real and imaginary parts as

$$(3.1) \quad \hat{\lambda}_R = \lambda_R + \frac{y_R^T E x_R + y_I^T E x_I}{y_R^T x_R + y_I^T x_I} + \mathcal{O}(\|E\|_F^2)$$

$$(3.2) \quad \hat{\lambda}_I = \lambda_I + \frac{y_R^T E x_I - y_I^T E x_R}{y_R^T x_R + y_I^T x_I} + \mathcal{O}(\|E\|_F^2)$$

where $\lambda = \lambda_R + i\lambda_I$, $\hat{\lambda} = \hat{\lambda}_R + i\hat{\lambda}_I$, $x = x_R + ix_I$ and $y = y_R + iy_I$ with $\lambda_R, \lambda_I, \hat{\lambda}_R, \hat{\lambda}_I \in \mathbb{R}$ and $x_R, x_I, y_R, y_I \in \mathbb{R}^n$. Here, as in the previous section, we use right and left eigenvectors x and y normalized so that $\|x\|_2 = \|y\|_2 = 1$ and $y^H x = |y^H x|$. Following [3], norm-wise, absolute condition number for the real and imaginary parts of an eigenvalue λ with respect to a real perturbation E , can be defined as follows.

$$(3.3) \quad c_{\Re}^{\mathbb{R}}(\lambda) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \sup \left\{ |\hat{\lambda}_R - \lambda_R| : E \in \mathbb{R}^{n \times n}, \|E\|_F \leq \varepsilon \right\},$$

$$(3.4) \quad c_{\Im}^{\mathbb{R}}(\lambda) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \sup \left\{ |\hat{\lambda}_I - \lambda_I| : E \in \mathbb{R}^{n \times n}, \|E\|_F \leq \varepsilon \right\}.$$

The perturbation expansions (3.1) and (3.2) imply

$$(3.5) \quad c_{\Re}^{\mathbb{R}}(\lambda) = \sup_{\|E\|_F=1} \left(\frac{1}{y_R^T x_R + y_I^T x_I} \right) (x_R \otimes y_R + x_I \otimes y_I)^T \text{vec}(E)$$

$$(3.6) \quad = \frac{\sqrt{\|x_R\|^2 \|y_R\|^2 + \|x_I\|^2 \|y_I\|^2 + 2(x_R^T x_I)(y_R^T y_I)}}{y_R^T x_R + y_I^T x_I},$$

$$(3.7) \quad c_{\Im}^{\mathbb{R}}(\lambda) = \sup_{\|E\|_F=1} \left(\frac{1}{y_R^T x_R + y_I^T x_I} \right) (x_I \otimes y_R - x_R \otimes y_I)^T \text{vec}(E)$$

$$(3.8) \quad = \frac{\sqrt{\|x_I\|^2 \|y_R\|^2 + \|x_R\|^2 \|y_I\|^2 - 2(x_R^T x_I)(y_R^T y_I)}}{y_R^T x_R + y_I^T x_I}.$$

Note that the condition numbers $c_{\Re}^{\mathbb{R}}(\lambda)$ and $c_{\Im}^{\mathbb{R}}(\lambda)$ may have different magnitudes. The imaginary part of an eigenvalue may be well-conditioned while the real part is not. For example, if λ , x and y are real, then $c_{\Re}^{\mathbb{R}}(\lambda) = 1/(y^T x)$ and $c_{\Im}^{\mathbb{R}}(\lambda) = 0$ which reflects the fact that $\hat{\lambda}$ is real for all sufficiently small $E \in \mathbb{R}^{n \times n}$.

An alternate proof of Theorem 2.1 may be obtained from the observation

$$(3.9) \quad c^{\mathbb{C}}(\lambda) = \sqrt{(c_{\Re}^{\mathbb{R}}(\lambda))^2 + (c_{\Im}^{\mathbb{R}}(\lambda))^2} = \|X\|_F$$

where X is the matrix in (2.6). Equation (3.9) also shows that at least one of (3.3) and (3.4) must be within a factor of $1/\sqrt{2}$ of the complex perturbation condition number $c^{\mathbb{C}}(\lambda) = 1/|y^H x|$.

A general framework of eigenvalue perturbations of real matrices allowing the use of a broad class of norms has been developed by Karow [2] based on the theory of spectral value sets and real μ -functions. However, it is not known how a simple bound relating the sensitivities of λ to real and complex perturbations can be obtained from this theory.

4 Generalizations

The arguments above generalize without modification to real perturbations of complex matrices, because the expansion (2.1) depends only on the eigenvalue λ and its corresponding left and right eigenvectors. Consequently, the conclusion of Theorem 2.1 as well as (3.5) and (3.7) carry over to this case.

The results in the previous two sections also generalize immediately to perturbations E whose entries lie on a line in the complex plane through the origin, for example, perturbation matrices E whose entries have zero real part. Such perturbations take the form $E = e^{i\theta} \hat{E}$ for some $\theta \in \mathbb{R}$ and $\hat{E} \in \mathbb{R}^{n \times n}$, so the conditioning and perturbation properties of $A + E$ are identical to those of $e^{-i\theta} A + \hat{E}$ which is a real perturbation of a complex matrix.

More importantly, inequality (2.4) also generalizes to an arbitrary analytic function $f : \mathbb{C}^n \rightarrow \mathbb{C}^m$. At a given point $x \in \mathbb{C}^n$, f satisfies the expansion

$$(4.1) \quad f(x + \Delta x) = f(x) + L \cdot \Delta x + \mathcal{O}(\|\Delta x\|_2^2),$$

for all perturbations $\Delta x \in \mathbb{C}^n$ in a sufficiently small neighborhood around zero. Here, L is the complex $m \times n$ Jacobian matrix depending on x . Following [3], define the absolute condition number of f at x as

$$(4.2) \quad c^{\mathbb{F}}(x) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \sup \{ \|f(x + \Delta x) - f(x)\|_2 : \Delta x \in \mathbb{F}^n, \|\Delta x\|_2 \leq \varepsilon \}.$$

The perturbation expansion (4.1) implies

$$c^{\mathbb{F}}(x) = \sup_{\substack{\Delta x \in \mathbb{F}^n \\ \|\Delta x\|_2 = 1}} \|L \cdot \Delta x\|_2.$$

For $\mathbb{F} = \mathbb{R}$, decompose $L = L_R + iL_C$ with $L_R, L_C \in \mathbb{R}^{m \times n}$. Since $\|L \cdot \Delta x\|_2^2 = \|L_R \cdot \Delta x\|_2^2 + \|L_C \cdot \Delta x\|_2^2$, we have

$$c^{\mathbb{R}}(x) = \left\| \begin{bmatrix} L_R \\ L_C \end{bmatrix} \right\|_2.$$

On the other hand, $c^{\mathbb{C}}(x)$ coincides with

$$\|L\|_2 = \left\| \begin{bmatrix} L_R & -L_C \\ L_C & L_R \end{bmatrix} \right\|_2.$$

This shows $1/\sqrt{2} \cdot c^{\mathbb{C}}(x) \leq c^{\mathbb{R}}(x) \leq c^{\mathbb{C}}(x)$.

5 Conclusions

The real-part and imaginary-part eigenvalue condition numbers (3.3) and (3.4) show that under real perturbations of a real matrix, the real part or the imaginary part of an eigenvalue may be ill-conditioned while the other is not. However, at least one of (3.3) and (3.4) must be within a factor of $1/\sqrt{2}$ of the complex perturbation condition number $c^C(\lambda) = 1/|y^H x|$. Hence, taken as a whole, real and imaginary parts together, eigenvalues of real matrices subject to real perturbations are not much better conditioned than eigenvalues of real matrices subject to complex perturbations.

Real arithmetic algorithms preserve the pairing of complex conjugate eigenvalues and eigenvectors, are typically computationally more efficient and use less computer memory than complex arithmetic algorithms applied to real matrices. Backward stable real arithmetic algorithms also calculate the real and imaginary parts of eigenvalues at least as accurately as the condition numbers (3.3) and (3.4) suggest, so they may calculate the real part or imaginary part of some eigenvalues more accurately than complex arithmetic algorithms. However, regarding eigenvalues as complex numbers, real and imaginary parts together, backward stable real arithmetic algorithms are not in general much more accurate than backward stable complex arithmetic algorithms.

6 Acknowledgments

The authors thank Michael Karow for stimulating discussions and the referees for helpful remarks.

REFERENCES

1. G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, Johns Hopkins University Press, Baltimore, MD, third ed., 1996.
2. M. KAROW, *Geometry of Spectral Value Sets*, PhD thesis, Universität Bremen, Fachbereich 3 (Mathematik & Informatik), Bremen, Germany, 2003.
3. J. R. RICE, *A theory of condition*, SIAM J. Numer. Anal., 3 (1966), pp. 287–310.
4. G. W. STEWART AND J.-G. SUN, *Matrix Perturbation Theory*, Academic Press, New York, 1990.
5. J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.