International Journal of Statistics and Applied Mathematics 2018; 3(2): 285-288

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2018; 3(2): 285-288

www.mathsjournal.com
Received: 19-01-2018
Accepted: 22-02-2018

Nitin Chouhan
Department of Mathematics, Jaipur National University, Rajasthan, India

Mohan Singh Dhaked
Department of Mathematics, Jaipur National University, Rajasthan, India

Hoshiyar Singh
Department of Mathematics, Jaipur National University, Rajasthan, India

Yogesh Khandelwal
Department of Mathematics, Jaipur National University, Rajasthan, India

Correspondence
Mohan Singh Dhaked
Department of Mathematics, Jaipur National University, Rajasthan, India

A study of measuring the distance of a stable matrix to the unstable matrices

Nitin Chouhan, Mohan Singh Dhaked, Hoshiyar Singh and Yogesh Khandelwal

Abstract
This paper determine the 2-norm and frobenius norm distance form a given matrix A to the nearest matrix, we explain a bisection method due to Byers, with an Eigen value on the imaginary axis. If A has an Eigen value plane, then, this distance measures how "nearly unstable" A is. Each bisection step provides a rigorous upper bound or lower bound on the distance. Here, we use Bayer's Bisection method to estimate the distance to the nearest matrix with an Eigen value on the unit circle.

Keywords: Hamiltonian matrix, rounding error analysis, matrix pencil

Introduction

If A is a square matrix, a non-zero vector v is an eigenvector of A if there is a scalar λ such that
$Av = \lambda v$

The scalar λ is said to be the eigenvalue of A corresponding to v. The singular value decomposition of an m × n real or complex matrix M is a factorization of the form
$M = U\Sigma V^*$,

Where U is an m×m real or complex unitary matrix, Σ is an m×n diagonal matrix with nonnegative real numbers on the diagonal, and V is an n × n real or complex unitary matrix. The diagonal entries $\Sigma_{i,i}$ of Σ are known as the singular values of M. The m columns of U and the n columns of V are called the left singular vectors and right singular vectors of M, respectively. Suppose that $A \in \mathbb{C}^{n\times n}$ has no Eigen value on the imaginary axis. Let $U \subset \mathbb{C}^{n\times n}$ be the set of matrices with at least one eigenvalue on the imaginary axis. The stability radius of A is defined by

$$\beta(A) = \min\{\|E\| \mid A + E \in U\} \qquad (1)$$

If A is stable, that is, all eigenvalues of A have negative real part, then β(A) is the distance to the set of unstable matrices. It is a measure of how "nearly unstable" is the stable matrix A. In this report we describe a bisection algorithm due to Byers for calculating β(A). It can be shown that

$$\beta(A) = \min_{\omega\in\mathbb{R}} \sigma_{\min}(A - \omega i I) \qquad (2)$$

Where $\sigma_{\min}(A - \omega i I)$ is the smallest singular value of $(A - \omega i I)$. So, for any real ω, an upper bound on β(A) is

$$\beta(A) \leqq \sigma_{\min}(A - \omega i I) \qquad (3)$$

The bisection method described in this report gives upper and lower bounds on the global minimum of $f(\omega) = \sigma_{\min}(A - \omega i I)$. It is not affected by the number of local minima f (ω) may have, nor does it require any starting value in order to begin. The technique can be used to bracket β(A) to any accuracy, but it converges too slowly to be used to calculate β(A) many significant digits. Fortunately, β(A) rarely needs to be calculated more accurately than within an order of magnitude. Throughout, λ(A) represent the set of eigenvalues of A. The norm ||.|| may represent either the operator 2-norm $\|A\| = max_{x\neq 0}(\|Ax\|_2/\|x\|_2)$ or the Frobenius norm $\|A\| = \sqrt{trace(A^H A)}$.

Hamiltonian Matrix

In mathematics, a Hamiltonian matrix is a 2n-by-2n matrix A such that JA is symmetric, where J is the skew-symmetric matrix

$$J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$$

and I_n is the n-by-n identity matrix. In other words, A is Hamiltonian if and only if $(JA)^T = JA\ where\ (\ )^T$ denotes the transpose.

Symplectic Matrix

In mathematics, a Symplectic matrix is a 2n×2n matrix M with real entries that satisfies the condition

$$M^T \Omega\, M = \Omega$$

Where M^T denotes the transpose of M and Ω is a fixed 2n×2n nonsingular, skew-symmetric matrix.

Bisection Method

Given $\sigma \geqq 0$ and an n-by-n matrix A, define the 2n-by-2n matrix H = H(σ)
By

$$H(\sigma) = \begin{bmatrix} A & -\sigma I_n \\ \sigma I_n & -A^H \end{bmatrix} \quad (4)$$

Here I_n denotes the n-by-n identity matrix. Theorem 1 given below shows how the eigenvalue of H(σ) distinguish the case $\sigma \geqq \beta(A)$ from the case $\sigma < \beta(A)$.

Theorem 1: H(σ) has an Eigen value whose real part is zero if and only if $\sigma \geqq \beta(A)$.

Proof. If for some ω ∈ℝ, ωi∈λ(H), then there are nonzero vector u,v∈$\mathbb{C}_n$ such that

$$\begin{bmatrix} A & -\sigma I_n \\ \sigma I_n & -A^H \end{bmatrix} \begin{bmatrix} v \\ u \end{bmatrix} = \omega i \begin{bmatrix} v \\ u \end{bmatrix} \quad (5)$$

Which gives

$$Av - \sigma I u = \omega i v \quad (6)$$
$$\sigma I n v - A^H u = \omega i u \quad (7)$$

After solving (1.6) and (1.7), we have

$$(Av - \omega iI)v = \sigma u \quad (8)$$
$$(A - \omega iI)^H u = \sigma v \quad (9)$$

Thus σ is a singular value of $(A - \omega iI)$ and A implies σ ≥ β(A). Conversely, suppose that $\sigma \geqq \beta(A)$. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be the function such that $f(\alpha) = \sigma_{min}(A - \alpha iI)$.

The function f is continuous and $lim_{\alpha\rightarrow\infty} f(\alpha) = \infty$. so, at some $\alpha \in \mathbb{R}$, f attains its minimum value $f(\alpha) = \beta(A) \leqq \sigma$. The intermediate value theorem implies that for some $\omega \in \mathbb{R}$, f(ω) = σ. So σ is the singular value of A−ωiI and there are unit vectors $u \in \mathbb{C}^n$ and $v \in \mathbb{C}^n$ satisfying (8) and (9) But this implies (5) and ωi∈λ(H).

Suppose that α is a lower bound and γ is an upper bound on β(A). The bounds can be improved by choosing a number σ between α and γ and checking to see if H(σ) has an eigenvalue with zero real part. Theoretically, any accuracy can be obtained, but the bisection method converges only linearly, So highly accurate calculations of β(A) may be expensive. All that is usually needed is an estimate of β(A) that is correct to within a factor of 10. The Algorithm given below estimates β(A) to within a factor of 10 or indicates that β(A) is less than a small tolerance. The algorithm makes use of the naive upper bound $\beta(A) \leqq \frac{1}{2}\|A + A^H\|$.

Algorithm 1

Input: An n-by-n matrix A and tolerance τ > 0

Output: α ∈ℝ and γ ∈ℝ such that either $\gamma/10 \leqq \alpha \leqq \beta(A) \leqq \gamma$ or

$0 = \alpha \leqq \beta(A) \leqq \gamma \leqq 10\tau$.
$\alpha := 0; \gamma := \frac{1}{2}\|A + A^H\|$.

While $\gamma > 10$ MAX (τ,α)

$\sigma := \sqrt{\gamma MAX(\tau, \alpha)}$

IF H(σ) has an eigenvalue with zero real part THEN γ := σ ELSE α := σ. If $\tau = \frac{1}{2}10^{-\Gamma}\|A + A^H\|$, then at most $\lceil log_2 p \rceil$ bisection steps are required. For the choice $\tau = \frac{1}{2}10^{-8}\|A + A^H\|$, at most three bisection steps are required. The bulk of the work in Algorithm 1 is deciding whether H(σ) has an eigenvalue with zero real part.

Rounding Error Analysis

The success of the Algorithm 1 depends on its being able decide, despite rounding errors, whether H(σ) has an eigenvalue with zero real part. As long as the decision is made correctly, the upper and lower bounds obtained by the algorithm are correct. In this section we show that if the decision is made in a numerically stable way, then it is correct except possibly when σ is within the rounding error of β(A). So, in the worst case, β(A) may lie outside the bounds given by the algorithm by an amount proportional to the precision of the arithmetic. Define $J \in \mathbb{R}^{2n\times 2n}$ by

$$J = \begin{bmatrix} o_n & I_n \\ -I_n & 0_n \end{bmatrix} \quad (10)$$

where I_n is the n-by-n identity matrix and 0_n is the n-by-n zero matrix.

Theorem 2: Let J be as in (1) $H \in \mathbb{C}^{2n\times 2n}$ be Hamiltonian and $\sigma \in \mathbb{R}$. If $H - \sigma J$ has an eigenvalue on the imaginary axis then | σ |= β(H). If | σ |= β(H) then at least one of $H + \sigma J$ and $H - \sigma J$ has an eigenvalue on the imaginary axis.

Proof.
If ω ∈ℝ and $\omega i \in \lambda(H - \sigma J)$, then

$$0 = det(H - \sigma J - \omega iI) \quad (11)$$

$$0 = det(-JJH + \sigma JJJ + \omega iIJJ) \quad (12)$$

$$0 = det(-J(JH + \sigma I - \omega IJ)) = det(-J)det(JH - \omega iJ + \sigma I) \quad (13)$$

Now $det(-J) = -1$, so $-\sigma \in \lambda(JH - \omega iJ)$. Because H is Hamiltonian, $JH - \omega iJ$ is Hermitian. The singular values of a Hermitian matrix are the absolute values of its eigenvalues. So, | σ | is a singular value of $H - \omega iI$ and $\beta(A) \leqq \sigma_{min}(A - \omega iI)$ implies $|\sigma| \geqq \beta(H)$.
Conversely, suppose that $|\sigma| \geqq \beta(H)$. An application of the intermediate value theorem to $f(\alpha) = \sigma_{min}(H - \alpha iI)$ guarantees the existence of a number ω ∈ℝ such that | σ | is a singular value $H - \omega iI$. Since J is unitary, | σ | is also a singular value of the Hermitian matrix $JH - \omega iJ$ and $\pm\sigma \in \lambda(JH - \omega iJ)$. So $0 = det(JH - \omega iJ \pm \sigma I)$
$= det(J)det(H - \omega iI \pm \sigma I)$. ...(14)
Hence $\omega i \in \lambda(H \pm \sigma J)$.
The matrix H(σ) in (4) and in Algorithm 1 is Hamiltonian. The next theorem shows that, despite rounding errors, the decision that H(σ) does or does not have an Eigen value on the imaginary axis is essentially correct, providing that Hamiltonian structure is preserved.

Theorem3: Let $\sigma \in \mathbb{R}$ be non-negative and let H (σ) be as in (4). Suppose $E \in \mathbb{C}^{2n\times 2n}$ is Hamiltonian and $K\sigma) = H(\sigma) + E$. If $K(\sigma)$ has an eigenvalue with zero real part, then $(A) \leq \sigma + 2||E||$. If $K(\sigma)$ does not have an eigenvalue with zero real part, then $\beta(A) \geq \sigma - 2||E||$.

Proof
Let $\omega \in \mathbb{R}$ be such that $\omega i \in \lambda[H(\beta(A))]$. so

$$0 = det(J)\det[H(\beta(A)) - \omega iI]$$
$$0 = \det[JH(0) - \omega iJ + \beta(A)I] \quad (15)$$

i.e. $-\beta(A) \in \lambda(JH(0) - \omega iJ)$. The Wielandt-Hoffman theorem [10] implies that there is an eigenvalue $-\gamma \in \mathbb{R}$ of $JH(0) - \omega iJ + JE = JK(0) - \omega iJ$ such that $\gamma - \beta(A) \leq ||E||$. Theorem 2 implies that $|\gamma| \geq \beta(K(0))$. Hence $\beta(K(0)) \leq \beta(A) + ||E||$. A similar argument with the roles of $H(0)$ and $K(0)$reversed shows that $\beta(A) \leq \beta(K(0)) + ||E||$, so

$$\beta(A) - ||E|| \leq \beta(K(0)) \leq \beta(A) + ||E||$$

If $K(\sigma)$ has an eigenvalue on the imaginary axis, then

$$\sigma \geq \beta(K(0)) \geq \beta(A) - ||E||.$$

If neither $K(\sigma)$ nor $H(\sigma)$ have an eigenvalue on the imaginary axis, then

$$\sigma < \beta(A) < \beta(A) + 2||E||$$

If $K(\sigma)$ has no Eigen value on the imaginary axis and $H(\sigma)$ does, then there is a number $\omega \in \mathbb{R}$ such that $H(\sigma) - \omega iI$ is singular. So,

$$\sigma - \beta(K(0)) \leq \sigma_{min}[K(0) - \omega iI - \sigma J] = \sigma_{min}[H(\sigma) - \omega iI + E]$$

$$\leq \sigma_{min}[H(\sigma) - \omega iI] + ||E|| = ||E||.$$

Since β(K(0)) ≤ β(A) + ||E ||, we have σ ≤ β(A) + 2||E ||. Using t digit base b arithmetic, the square reduced algorithm [11] applied to a Hamiltonian matrix H delivers the Eigen values of a Hamiltonian matrix K = H + E, where ||E || = O($b^{t/2}$||H ||). Theorem 3 shows that if the square reduced algorithm is used to test H (σ) for a pure imaginary Eigen value in Algorithm 1, then the results may be "taken at face value".

Generalizations
The bisection method extends to estimating

$$\gamma(A) = \min\{||E|| \text{ for some } \theta \in \mathbb{R};\ e^{i\theta} \in \lambda(A+E)\} \quad (16)$$

i.e., the distance from A to the nearest matrix with an Eigen value on the unit circle. If A is stable in the sense that all its Eigen values lie inside the unit circle, then (A) is a measure of how “nearly unstable” A is. By applying the method to simple modifications of A, the unit circle can be replaced by an arbitrary circle and the imaginary axis can be replaced by an arbitrary line.

Matrix Pencil
Definition: A matrix pencil is the set of all matrices of the form G−λH, where G and H are given n-by-n matrices and $\lambda \in \mathbb{C}$. A number $\mu \in \mathbb{C}$ is a generalized Eigen value of the pencil G−λH if μ is a root of the polynomial $\psi(\lambda)$ = det(G−λH). The following matrix pencil theorem is an analogue of Theorem 1.

Theorem 4: For $A \in \mathbb{C}^{nxn}$, there is a number r(A) ϵ R such that r(A) $\geqq \gamma$(A) and for r(A) $\geqq \sigma \geqq \gamma$(A), the 2n-by-2n matrix pencil.

$$F(\sigma) - \lambda G(\sigma) = \begin{bmatrix} -\sigma I_n & A \\ I_n & 0_n \end{bmatrix} - \lambda \begin{bmatrix} 0_n & I_n \\ A^T & -\sigma I_n \end{bmatrix} \quad (17)$$

has a generalized eigen value of magnitude 1. Furthermore, if $\sigma < \gamma$(A), then matrix pencil has no generalized eigen value of magnitude 1.

Proof. Define the function g : [0, 2π) → R by g(θ) = σ_{min}(A − $e^{i\theta}I_n$).
This is a continuous function on a compact set, so it attains a maximum and a minimum value. Let r(A) be the maximum value. An easy singular-value argument shows the minimum value is γ(A).
The intermediate value theorem implies that for every σ such that r(A) $\geqq \sigma \geqq \gamma$(A), there is a $\theta \in [0, 2\pi]$ such that σ = g(θ). It follows that r(A) $\geqq \sigma \geqq \gamma$(A) if and only if there is a number $\theta \in [0, 2\pi]$ such that σ is an eigen value of

$$K(\theta) = \begin{bmatrix} 0_n & (A - e^{i\theta}I_n) \\ (A^H - e^{i\theta}I_n) & 0_n \end{bmatrix} \quad (18)$$

Now $\sigma \in$ R is an eigen value of K(θ) if and only if there is an vector $u \in \mathbb{C}^{2n}$
Such that

K(θ) u = σu.

but this is easily rearranged to give

F(σ)u = $e^{i\theta}$G(σ)u.

The following algorithm estimates γ(A) within a factor of 10 or indicates that γ(A) is less than a small tolerance. It makes use of the naive bound

r(A) ≧ σ_{min}(A − I_n).

Algorithm 2
Input: An *n-by-n* matrix A and tolerance $\tau > 0$
Output: $\alpha \epsilon \mathbb{R}$ and $\delta \epsilon \mathbb{R}$ such that $\delta/10 \leqq \alpha \leqq \gamma(A) \leqq \delta$ or
$0 = \alpha \leqq \gamma(A) \leqq \delta \leqq 10\tau$.

$\alpha := 0; \delta := \sigma_{min}$(A − I_n).

While $\gamma >$ 10 MAX (τ,α)

$\sigma := \sqrt{\delta MAX(\tau, \alpha)}$

IF matrix pencil has a generalized eigen value of magnitude1, THEN δ := σ ELSE α := σ.
Let Re (z) represent the real part of z ϵ ℂ and Arg(z) represant the argument of z. It is easy to show that for r > 0 and $\rho \epsilon \mathbb{C}$, $r^{-1}\gamma$(r(A + ρI)) is the distance from A to the nearest matrix with an eigen value on the circle {z ϵ ℂ ‖ z $-\rho$|= r^{-1}}. Algorithm 2 applied to aA+ ρI estimates this quantity. Similarly, if $\rho \epsilon \mathbb{C}$ and $\theta \epsilon \mathbb{R}$, then $\beta(e^{i\theta}($ A $- \rho I))$ is the distance from A to the nearest matrix with an eigen value on the line Arg(z $-\rho$) = $\pi - \theta$. Note that all lines in the complex plane take this form. Furthermore, if all eigen values of A lie in a convex region R bounded by the p lines, Arg(z $-\rho_j$) = $\pi - \theta_j$.,j = 1,2,3,......,p, then the magnitude of the smallest perturbation that drives an eigen value outside R is minj $\beta(e^{i\theta_j}$A $- \rho_j I)$. This quantity is estimated by p applications of Algorithm 1. As with Algorithm 1, the success of Algorithm 2 depends on its being able to decide, despite rounding errors, if a generalized eigenvalue of (17) has magnitude 1. If F(σ) is nonsingular, then G(σ)F(σ)$^{-1}$ is symplectic.

Conclusion
We have presented a bisection method due to Byers for computing β(A). A heuristic estimate of β(A) is

$$\beta(A) \approx \{\sigma_{min}(A - Re(\lambda)iI) | \lambda \ \forall\epsilon \ \lambda(A)\} \quad (19)$$

The other method consists of applying a general nonlinear minimization algorithm to

$$f(w) = \sigma_{min}(A - wiI)) \quad (20)$$

The heuristic gives an upper bound that is often within an order of magnitude of β(A). However, there are examples for which the bound is greater than β(A) by an arbitrary amount. The bisection method does not fail to approximate β(A) to within an order of magnitude. Furthermore, it delivers both an upper and a lower bound on β(A). In the absence of a good initial guess, a nonlinear minimization algorithm applied to (20) may start slowly. Accurate function values using a singular value decomposition cost O(n^3) flops each and so are too expensive except for low-dimensional problems. The function values may be economically estimated with a condition estimator, but this can destroy the smoothness of f(w) on which many minimization algorithms depend. It is not always the case that f(w) is an "easy" function to minimize. When β(A) = 0, f(w) may not be smooth; the minimum lies on a cusp. Furthermore, for an n-by-n matrix A, f(w) may have as many as n local minima and as few as one. A nonlinear minimization algorithm must be used O(n) times to search for all possible local minima. If fewer than n local minima are discovered, there remains some doubt that all were found. In contrast, the bisection algorithm does not require good starting values. Cusps in f(w) do not affect the bisection algorithm nor does the number of local minima.
Bisection is competitive with the heuristic (19) and nonlinear minimization (20). Bisection has no possibility of failure. It produces upper and lower bounds that differ by no more than a factor of 10. It enjoys a favorable.

References

1. Laub AJ, Meyer KR. Canonical Forms for Symplectic and Hamiltonian Matrices, Celestial Mechanics, 1974; 9:213-238.
2. Spence A, Poulton C. Photonic band structure calculations using non-linear eigenvalue techniques. J. Comput. Phys. 2005; 204:65-61.
3. A symplectic method for approximating all eigenvalues of a Hamiltonian matrix, Linear Algebra Appl., 1984; 61:233-251.
4. Van CF. Loan how near is a stable matrix to an unstable matrix? in: Liner algebra and its role in system theory(Brunswick, maine), Contemp. Math, Amer. Math. Soc. Providence, RI, Prividence, 1985; 47:465-478.
5. Golub GH, Loan Van CF. Matrix Computatinos, The Johns Hopkins Press, Baltimorem Maryland, 1983.
6. Ralph Byers The bisection method for measuring the distance of a stable matrix to the unstable matrices, Siam J. Sci. Stat. Comput. 1988, 9.
7. Ralph Byers. A Hamiltonian QR algorithm, Siam J. Sci. Statist. Comput. 1986; 7:212-229.
8. Byers R. Hamilition and symplectic algorithms for the algebraic riccati eqation, Phd dissertation, cornell university, Ithaca, new York, 1983, 14856.