Real norms are not complex

L. Ridgway Scott*

October 17, 2012

Abstract

We provide alternate derivations of some results in numerical linear algebra based on a representation of the action of a real matrix on real vectors in the case that it has complex conjugate eigenvalues. We consider two applications of this representation. The first one is to the relationship between the spectral radius and the norm of a matrix. The second relates the spectral radius and the limits of norms of roots of powers of a matrix.

There has been significant recent interest in the distinction between real and complex operator norms [1, 2]. Given a vector space, there is a natural way to complexify it, but this does not carry over to norms. A natural, but not unique, complexification of norms is the Taylor norm [2]

$$\|x + iy\|_T = \sup_{\theta\in[0,2\pi]} \|(\cos\theta)x - (\sin\theta)y\|. \tag{1}$$

The relation to the current work will be evident, but we do not exploit this in any way.

It is possible to extend many results for complex matrix norms to the corresponding real norms, but often the extension is not straightforward. Our objective is to provide simpler derivations. Our results are based on a representation of the action of a real matrix on real vectors in the case that it has complex conjugate eigenvalues. We consider two applications of this representation. The first one is an inequality relating the spectral radius and the norm of a matrix. The second equates the spectral radius with the limits of norms of roots of powers of a matrix.

*The University of Chicago, Chicago, Illinois 60637, email: ridg@uchicago.edu

1 Spectral radius and real norms

It is commonly stated [3] that

$$\|A\|_{\mathcal{O}} \geq \rho(A) \tag{2}$$

for any matrix A and any norm, where $\rho(A)$ is the spectral radius of A and the operator norm is defined by

$$\|A\|_{\mathcal{O}} = \max_{v \neq 0} \frac{\|Av\|}{\|v\|}. \tag{3}$$

The "proof" is deceptively simple: pick any eigenvalue λ and corresponding eigenvector $X \neq 0$, so that $AX = \lambda X$. Choose $v = X$ in (3), and we find

$$\|A\|_{\mathcal{O}} \geq \frac{\|AX\|}{\|X\|} = \frac{\|\lambda X\|}{\|X\|} = |\lambda|. \tag{4}$$

The difficulty with this argument arises if the norm has been defined over real vector spaces, not complex ones [1]. The issue is that the maximum in (3) is taken over only real vectors v and yet we need to apply this with a complex vector. Here we present a proof of (2) that avoids this pitfall. We will prove the following.

Theorem 1 *Suppose A is a real matrix and $\|\cdot\|$ is any norm. Then there is a real vector X such that $\|AX\| = \rho(A)\|X\|$.*

If there is a real eigenvalue/eigenvector pair λ, X with $|\lambda| = \rho(A)$, then we are done. So suppose we have a real matrix A with a complex pair of eigenvalues, λ and $\overline{\lambda}$, where $|\lambda| = \rho(A)$ (the pair of eigenvalues are complex conjugates). Write $X = Y + iW$ where both Y and W are real vectors, and also $\lambda = \mu + i\nu$ where both μ and ν are real numbers. We can assume that $\nu \neq 0$ since $\nu = 0$ is the previously considered case.

We have $AX = AY + iAW$ since A is real. Writing out $AX = \lambda X$ we find

$$AY + iAW = AX = \lambda X = \mu Y - \nu W + i(\nu Y + \mu W). \tag{5}$$

Equating real and imaginary parts, we thus find

$$\begin{aligned} AY &= \mu Y - \nu W \\ AW &= \nu Y + \mu W. \end{aligned} \tag{6}$$

This says that A maps the space spanned by Y and W into itself. Call this space V. We will see that V has to be two-dimensional, although this is not essential for the proof. We will find $Z \in V$ such that $\|AZ\| = |\lambda| \|Z\|$.

If Y and W were collinear, that is, $W = \alpha Y$, with α real, then we would have

$$(1 + i\alpha)AY = A(Y + i\alpha Y) = \lambda(Y + i\alpha Y) = (1 + i\alpha)\lambda Y, \tag{7}$$

so that (dividing by $1 + i\alpha$) $AY = \lambda Y$. But this is not possible because both A and Y are real. Thus V must be two dimensional.

Any vector in V can be written as $c_1 Y + c_2 W$, and we easily compute from (6) that

$$A(c_1 Y + c_2 W) = d_1 Y + d_2 W \tag{8}$$

where

$$\begin{pmatrix} \mu & \nu \\ -\nu & \mu \end{pmatrix} c = d. \tag{9}$$

The matrix in (9) is a rotation and scaling:

$$\begin{pmatrix} \mu & \nu \\ -\nu & \mu \end{pmatrix} = \begin{pmatrix} |\lambda| \cos\theta & |\lambda| \sin\theta \\ -|\lambda| \sin\theta & |\lambda| \cos\theta \end{pmatrix} = |\lambda| \mathcal{R}(-\theta), \tag{10}$$

since $|\lambda|^2 = \mu^2 + \nu^2$. Here $\mathcal{R}(\theta)$ denotes the matrix that rotates a vector by an angle θ.

Let us consider a parameterization of vectors in V where

$$c(\phi) = (\cos\phi, \sin\phi) = R(\phi)E, \tag{11}$$

where

$$E = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{12}$$

The image vectors $d(\phi) = \mathcal{R}(\phi - \theta)E$.

Define $Z(\phi) = c_1(\phi)Y + c_2(\phi)W$ and

$$f(\phi) = \frac{\|AZ(\phi)\|}{\|Z(\phi)\|}. \tag{13}$$

It is helpful to define an induced norm

$$\|c\|_V = \|c_1 Y + c_2 W\| \tag{14}$$

for all $c \in \mathbb{R}^2$. Then

$$f(\phi) = \frac{\|AZ(\phi)\|}{\|Z(\phi)\|} = \frac{\|d(\phi)\|_V}{\|c(\phi)\|_V} = |\lambda| \frac{\|\mathcal{R}(\phi - \theta)E\|_V}{\|\mathcal{R}(\phi)E\|_V}. \tag{15}$$

Define $g(\phi) = \|\mathcal{R}(\phi)E\|$. Then g is periodic and continuous. Define

$$\phi_* = \text{argmin}_{\phi \in [0,2\pi]} g(\phi) \quad \text{and} \quad \phi^* = \text{argmax}_{\phi \in [0,2\pi]} g(\phi). \tag{16}$$

Then

$$g(\phi_*) \leq g(\phi_* - \theta) \quad \text{and} \quad g(\phi^* - \theta) \leq g(\phi^*), \tag{17}$$

and thus $f(\phi_*) \geq |\lambda| \geq f(\phi^*)$. By the intermediate value theorem, there is a $\hat{\phi}$ such that $f(\hat{\phi}) = |\lambda|$. Therefore $Z = Z(\hat{\phi})$ is the desired vector. **QED**

2 Limiting relations

The following is frequently called Gelfand's equality:

$$\rho(A) = \lim_{k \to \infty} \|A^k\|^{1/k}. \tag{18}$$

Here we let $\|A\|$ denote any norm, not necessarily one derived as an operator norm. To prove this, we begin with the following result.

Theorem 2 *Suppose A is a real $n \times n$ matrix and $\|A\|$ denotes any norm on $\mathbb{R}^{n^2}$. If $\rho(A) > 1$, then*

$$\lim_{k \to \infty} \|A^k\| = \infty. \tag{19}$$

First we assume that there is a complex pair of eigenvalues as in the proof of Theorem 1. Iterating the representation (8), we find

$$A^k(c_1 Y + c_2 W) = d_1^{(k)} Y + d_2^{(k)} W, \tag{20}$$

where

$$d^{(k)} = \begin{pmatrix} \mu & \nu \\ -\nu & \mu \end{pmatrix}^k c = |\lambda|^k \mathcal{R}(-k\theta)c. \tag{21}$$

Suppose that for some (finite or infinite) set of positive integers K, we have

$$\|A^k\| \leq C_0 \text{ for all } k \in K. \tag{22}$$

Then the same bound is true for any norm, with a different constant, by the equivalence of norms on a finite dimensional vector space. In particular, $\|A^k\|_\infty \leq C_1$, where this is the component-wise max norm, for all $k \in K$. That is, $|(A^k)_{ij}| \leq C_1$ for all i, j and for all $k \in K$. Thus we conclude that

$$|d_1 Y + d_2 W|_\infty \leq C_2 |c|_2, \tag{23}$$

where the latter norm is the Euclidean norm in $\mathbb{R}^2$. The first term in (23) defines a norm on two-space:

$$|d|_* = |d_1 Y + d_2 W|_\infty. \tag{24}$$

Thus

$$|\lambda|^k |\mathcal{R}(-k\theta)c|_* \leq C_2 |c|_2 = C_2 |\mathcal{R}(-k\theta)c|_2, \tag{25}$$

since the 2-norm is invariant under rotation. Since the 2-norm and the *-norm are equivalent, we have

$$|\lambda|^k \leq C_3 \tag{26}$$

for all $k \in K$. If $|\lambda| > 1$, K must be a finite set. Since C_0 was arbitrary, (19) holds.

If there is a real eigenvalue $\lambda = \pm\rho(A)$, with eigenvector $X \neq 0$, the result follows from the fact that $A^k X = \lambda^k X$ by similar arguments. **QED**

The remainder of the proof of (18) is standard. Let $\epsilon > 0$ be arbitrary, but with $\epsilon < \rho(A)$. (If $\rho(A) = 0$, the result is trivial.) We define

$$A_\pm = \frac{1}{\rho(A) \pm \epsilon} A. \tag{27}$$

We have $\rho(A_+) < 1$ and $\rho(A_-) > 1$.

Applying Theorem 2, there is an N such that

$$\|A_-^k\| \geq 1, \tag{28}$$

for all $k \geq N$. Thus

$$\|A^k\|^{1/k} = (\rho(A) - \epsilon)\|A_-^k\|^{1/k} \geq (\rho(A) - \epsilon), \tag{29}$$

for all $k \geq N$. Thus

$$\liminf_{k\to\infty} \|A^k\|^{1/k} \geq (\rho(A) - \epsilon). \tag{30}$$

Since this holds for any $\epsilon > 0$, we must have

$$\liminf_{k\to\infty} \|A^k\|^{1/k} \geq \rho(A). \tag{31}$$

Since $\rho(A_+) < 1$, we have $A_+^k \to 0$ as $k \to \infty$. This is a deep result, but it is independent of any norm. Therefore there is an N such that

$$\|A_+^k\| \leq 1, \tag{32}$$

for all $k \geq N$, by the continuity of the norm. Thus

$$\|A^k\|^{1/k} = (\rho(A) + \epsilon)\|A_+^k\|^{1/k} \leq (\rho(A) + \epsilon), \tag{33}$$

for all $k \geq N$. Thus

$$\limsup_{k\to\infty} \|A^k\|^{1/k} \leq (\rho(A) + \epsilon). \tag{34}$$

Since this holds for any $\epsilon > 0$, we must have

$$\limsup_{k\to\infty} \|A^k\|^{1/k} \leq \rho(A). \tag{35}$$

Since $\liminf_{k\to\infty} a_k \leq \limsup_{k\to\infty} a_k$ for any nonnegative real numbers a_k, (31) and (35) imply that

$$\lim_{k\to\infty} \|A^k\|^{1/k} = \rho(A). \tag{36}$$

QED

3 Other proof

We can prove (2) using (18) as follows for submultiplicative norms, that is, ones that satisfy $\|AB\| \leq \|A\|\|B\|$:

$$\rho(A) = \lim_{k\to\infty} \|A^k\|^{1/k} \leq \|A\|. \tag{37}$$

4 Remarks

Theorem 2 is often proved using the Jordan decomposition, together with careful analysis of powers of Jordan blocks. The approach taken here shows that this result is fairly elementary, susceptible to a two-dimensional analysis.

5 Acknowledgements

We thank Nick Trefethen and Lek-Heng Lim for valuable contributions to this note.

References

[1] Olga Holtz and Michael Karow. Real and complex operator norms. *Arxiv preprint math/0512608*, 2005.

[2] Natalia Sabourova. *Real and complex operator norms.* PhD thesis, Luleå University of Technology, 2007.

[3] L. R. Scott. *Numerical Analysis.* Princeton University Press, 2011.