On the Convergence of the Sequence of Powers of a 2 × 2 Matrix

ROMAN W. WONG
Washington and Jefferson College
Washington, PA 15301

Introduction The sequence $\{r^n\}$, r real, is convergent if, and only if, $|r| < 1$ or $r = 1$. The same statement is true when r is a complex number, and $|r|$ is the modulus of r. We show an analogous result for the convergence of the sequence $\{A^n\}$, where A is a 2×2 matrix over $\mathbb{R}$. The sequence $\{A^n\}$ converges to the zero matrix $\mathbb{O}$ if, and only if, $|\det(A)| < 1$ and $|1 + \det(A)| > |\mathrm{tr}(A)|$. This result is then applied to prove a well-known theorem in Markov chains for 2×2 regular stochastic matrices and to obtain an explicit formula for the stationary matrix and eigenvector.

1. Convergence of $\{A^n\}$ Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be a 2×2 matrix over the complex field $\mathbb{C}$. Let α and β be the eigenvalues of A. Let $n \in \mathbb{Z}^+$. Denote the 2×2 identity matrix and the zero matrix by $\mathbb{I}$ and $\mathbb{O}$, respectively. In **[4]**, Kenneth Williams shows how to compute the powers of A using eigenvalues:

$$A^n = \begin{cases} \alpha^n\left(\dfrac{A - \beta\mathbb{I}}{\alpha - \beta}\right) + \beta^n\left(\dfrac{A - \alpha\mathbb{I}}{\beta - \alpha}\right), & \text{if } \alpha \neq \beta, \\ \alpha^{n-1}\big(nA - (n-1)\alpha\mathbb{I}\big), & \text{if } \alpha = \beta \end{cases}$$

$$= \begin{cases} \dfrac{1}{\alpha - \beta}\big((\alpha^n - \beta^n)A - \alpha\beta(\alpha^{n-1} - \beta^{n-1})\mathbb{I}\big), & \text{if } \alpha \neq \beta, \\ n\alpha^{n-1}A - (n-1)\alpha^n\mathbb{I}, & \text{if } \alpha = \beta. \end{cases}$$

As a consequence of this formula, we see that if A has distinct eigenvalues α and β, then $\{A^n\}$ converges if, and only if, $\{\alpha^n - \beta^n\}$ converges. If $\lim_{n \mapsto \infty}(\alpha^n - \beta^n)$ exists, say the limit is k, then

$$\lim_{n \mapsto \infty} A^n = \frac{k}{\alpha - \beta}\big(A - \alpha\beta\mathbb{I}\big).$$

Since both $\{\alpha^n\}$ and $\{\beta^n\}$ must converge to 0 or 1, we see that the values of k above can only be 0, 1, or -1 and that $k = 0$ if, and only if, $|\alpha| < 1$ and $|\beta| < 1$ since $\alpha \neq \beta$.

Now suppose that $\alpha = \beta$ and $|\alpha| < 1$, then $\{A^n\}$ also converges to $\mathbb{O}$ since both $\{n\alpha^{n-1}\}$ and $\{(n-1)\alpha^n\}$ approach zero by L'Hôpital's rule. Conversely, if we write, using Williams's formula,

$$A^n = \begin{pmatrix} n\alpha^{n-1}a - (n-1)\alpha^n & n\alpha^{n-1}b \\ n\alpha^{n-1}c & n\alpha^{n-1}d - (n-1)\alpha^n \end{pmatrix}$$

$$= \begin{pmatrix} \alpha^{n-1}(n(a - \alpha) + \alpha) & n\alpha^{n-1}b \\ n\alpha^{n-1}c & \alpha^{n-1}(n(d - \alpha) + \alpha) \end{pmatrix},$$

we see that if $\{A^n\}$ converges to $\mathbb{O}$, we must have $|\alpha| < 1$.

We have proved the following proposition.

PROPOSITION. *Let A be a 2×2 matrix over $\mathbb{C}$ and let α and β be the eigenvalues of A. Then $\{A^n\}$ converges to $\mathbb{O}$ if, and only if, $|\alpha| < 1$ and $|\beta| < 1$.*

Remark. This proposition is true in general for any $n \times n$ matrix A over the complex field. Using Jordan canonical form, one can show that the sequence $\{A^n\}$ converges to the $n \times n$ zero matrix $\mathbb{O}$ if, and only if, $\max\{|\lambda| : \lambda$ is an eigenvalue of $A\} < 1$. A proof of this can be found in **[1]**.

The case when the inequalities in the proposition become equalities will be discussed later. For the rest of this paper, we restrict A to be a 2×2 matrix over the reals $\mathbb{R}$. We now present the theorem that characterizes the convergence of $\{A^n\}$ directly in terms of A. As usual, denote the determinant of A by $\det(A)$ and the trace of A by $\operatorname{tr}(A)$.

THEOREM. *Let A be a 2×2 matrix over $\mathbb{R}$. Then $\{A^n\}$ converges to $\mathbb{O}$ if, and only if, $|\det(A)| < 1$ and $|1 + \det(A)| > |\operatorname{tr}(A)|$.*

The proof of the theorem uses the result of the previous proposition and a lemma that will be presented later. The theorem is not true over the complex number $\mathbb{C}$. A counterexample will be given later.

2. Examples We now illustrate the theorem by some numerical examples. Computer software can be used to show the convergence.

A	α and β	$\det(A)$	$1 + \det(A)$	$\operatorname{tr}(A)$	Convergence to $\mathbb{O}$
$\begin{pmatrix} 1/2 & 0 \\ 1 & -1/3 \end{pmatrix}$	$1/2, -1/3$	$-1/6$	$5/6$	$1/6$	yes
$\begin{pmatrix} 1 & -1 \\ 1 & -1.9 \end{pmatrix}$	$3/5, -3/2$	-0.9	0.1	-0.9	no
$\begin{pmatrix} 2 & -1 \\ 0 & -1/3 \end{pmatrix}$	$2, -1/3$	$-2/3$	$1/3$	$5/3$	no
$\begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}$	$\pm 2i$	4	5	0	no
$\begin{pmatrix} 0 & 1/2 \\ -1/2 & 0 \end{pmatrix}$	$\pm \frac{i}{2}$	$1/4$	$5/4$	0	yes
$\begin{pmatrix} -5.2 & 3 \\ -10.6 & 6.1 \end{pmatrix}$	$0.1, 0.8$	0.08	1.08	0.9	yes

In each of the above examples, when $\{A^n\}$ does not converge to $\mathbb{O}$, the entries in A^n actually tend to ∞. This may not be the case when the inequalities in the theorem become equalities. Here are some examples where one or both of the boundary conditions are attained.

A	α and β	det(A)	1 + det(A)	tr(A)	Convergence
$\begin{pmatrix} 1 & -1/2 \\ 1 & 1/2 \end{pmatrix}$	$\dfrac{3 \pm \sqrt{15}\,i}{4}$	1	2	3/2	no convergence: entries tend to ∞
$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	0, 1	0	1	1	$\{A^n\}$ converges to A since $A^2 = A$
$\begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$	± 1	-1	0	0	A^n equals A and A^2 $(=\mathbb{I})$ alternately
$\begin{pmatrix} 3 & 1 \\ -4 & -1 \end{pmatrix}$	1, 1	1	2	2	$A^n = \begin{pmatrix} 2n+1 & n \\ -4n & 1-2n \end{pmatrix}$ entries tend to $\pm\infty$

Another example on the boundary is $A = \begin{pmatrix} 5/2 & 1 \\ -3 & -1 \end{pmatrix}$ with det$(A) = 1/2$ and det$(A) + 1 = 3/2 = \text{tr}(A)$. The eigenvalues of A are 1 and 1/2. By the proposition, since $\{\alpha^n - \beta^n\}$ approaches 1 as a limit, $\{A^n\}$ is convergent. In fact, A^n converges to $\begin{pmatrix} 4 & 2 \\ -6 & -3 \end{pmatrix} \neq \mathbb{O}$.

Here is one more interesting example. Let $A = \begin{pmatrix} 1 & -1/2 \\ 1 & 1/2 \end{pmatrix}$ with det$(A) = 1$ and tr$(A) = 3/2 < 1 + \det(A)$. The eigenvalues of A are the complex conjugates $\bar{\beta} = \alpha = (3 + \sqrt{7}\,i)/4$ with modulus 1. To four decimal places,

$$A^{1000} = \begin{pmatrix} 1.0491 & -0.1264 \\ 0.2527 & 0.9228 \end{pmatrix}; \quad A^{1001} = \begin{pmatrix} 0.9228 & -0.5877 \\ 1.1755 & 0.3350 \end{pmatrix}; \quad A^{1002} = \begin{pmatrix} 0.3350 & -0.7532 \\ 1.5105 & -0.4202 \end{pmatrix}.$$

The entries of A never go to infinity, but do not stabilize at any values either. A closer look at these matrices reveals that this system is not quite chaotic. In fact, if A is considered as a linear transformation in the xy-plane, the orbit of a vector under A is shown to be elliptic. My colleague Michael Woltermann pointed out that for any vector $P = \begin{pmatrix} a \\ b \end{pmatrix}$, both P and $AP = \begin{pmatrix} a - b/2 \\ a + b/2 \end{pmatrix}$ lie on the ellipse $x^2 - \frac{1}{2}xy + \frac{1}{2}y^2 - a^2 + \frac{1}{2}ab - \frac{1}{2}b^2 = 0$ and that the equation remains unchanged when a and b are replaced by $a - b/2$ and $a + b/2$, respectively. FIGURE 1 shows the orbits of the vectors $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 0.3 \\ 0.4 \end{pmatrix}$, and $\begin{pmatrix} 0.6 \\ -1.4 \end{pmatrix}$ under the transformation A.

3. Stochastic matrices Recall that a square matrix $A = [a_{ij}]$ is called *stochastic* if $a_{ij} \geqq 0$ for all i and j and $\Sigma_i a_{ij} = 1$ for each j. A is called *regular* if for some positive integer n, all the entries in A^n are strictly positive. The fundamental theory of Markov chains says that the powers of every regular stochastic matrix approach a stationary matrix with identical columns, and that this column vector is indeed an eigenvector for the eigenvalue 1. (See, for instance, [2].) We give a different proof here for the special case of 2×2 regular stochastic matrices and provide an explicit formula for the stationary matrix.

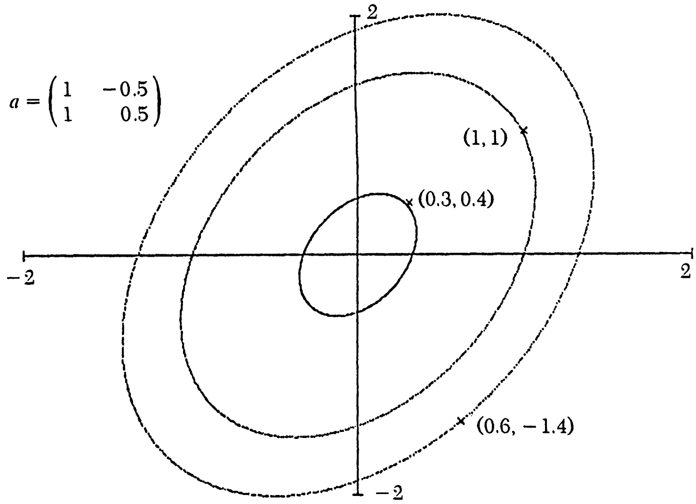

FIGURE 1

Let A be a 2×2 regular stochastic matrix with eigenvalues α and β. We can write A as

$$A = \begin{pmatrix} a & b \\ 1-a & 1-b \end{pmatrix}, \qquad 0 \le a, b \le 1.$$

We may assume that $\alpha = 1$ since one of the eigenvalues of a stochastic matrix must be 1. If β also equals to 1, we must have $\operatorname{tr}(A) = a + 1 - b = \alpha + \beta = 2$, which implies that $a = 1 + b$. Since $0 \leqq a, b \leqq 1$, we must have $b = 0$ and $a = 1$ and so $A = \mathbb{I}$. This is impossible since A is regular. Thus $\beta \neq 1$ and the two eigenvalues of A are distinct. Also $\beta = \alpha\beta = \det(A) = a - b$ and so $|\beta| < 1$. Thus $\lim_{n \mapsto \infty}(\alpha^n - \beta^n) = 1$ exists and we conclude that

$$\lim_{n \mapsto \infty} A^n = \frac{1}{1-a+b}\left(A - (a-b)\mathbb{I}\right) = \frac{1}{1-a+b}\begin{pmatrix} b & b \\ 1-a & 1-a \end{pmatrix}.$$

Thus

$$\frac{1}{1-a+b}\begin{pmatrix} b & b \\ 1-a & 1-a \end{pmatrix}$$

is the stationary matrix and the vector $\begin{pmatrix} b \\ 1-a \end{pmatrix}$ is an eigenvector for the eigenvalue 1.

As an example, consider the regular stochastic matrix $A = \begin{pmatrix} 1/4 & 1/3 \\ 3/4 & 2/3 \end{pmatrix}$. By the above formula, the stationary matrix for A would be

$$\frac{1}{1 - 1/4 + 1/3}\begin{pmatrix} 1/3 & 1/3 \\ 3/4 & 3/4 \end{pmatrix} = \frac{12}{13}\begin{pmatrix} 1/3 & 1/3 \\ 3/4 & 3/4 \end{pmatrix} = \frac{1}{13}\begin{pmatrix} 4 & 4 \\ 9 & 9 \end{pmatrix}$$

and the vector $\begin{pmatrix} 4 \\ 9 \end{pmatrix}$ is an eigenvector for the eigenvalue 1.

4. Proof of the theorem We first prove a lemma.

LEMMA. *Let both α and β be real numbers or be complex conjugates. Then $|\alpha| < 1$ and $|\beta| < 1$ if, and only if, $|\alpha\beta| < 1$ and $|1 + \alpha\beta| > |\alpha + \beta|$.*

Proof. First we assume that both α and β are reals. Suppose that $|\alpha| < 1$ and $|\beta| < 1$. Then $|\alpha\beta| < 1$ and

$$\begin{aligned} &1-\alpha^2>0 \quad \text{and} \quad 1-\beta^2>0 \\ &\Rightarrow (1-\alpha^2)(1-\beta^2)>0 \\ &\Leftrightarrow 1+\alpha^2\beta^2>\alpha^2+\beta^2 \\ &\Leftrightarrow 1+2\alpha\beta+(\alpha\beta)^2>\alpha^2+2\alpha\beta+\beta^2 \\ &\Leftrightarrow (1+\alpha\beta)^2>(\alpha+\beta)^2 \\ &\Leftrightarrow |1+\alpha\beta|>|\alpha+\beta|. \end{aligned}$$

Conversely, suppose that $|\alpha\beta| < 1$ and $|1+\alpha\beta| > |\alpha+\beta|$. Reversing the above argument, we have $(1-\alpha^2)(1-\beta^2) > 0$. This implies that either both $1 > \alpha^2$ and $1 > \beta^2$ or both $1 < \alpha^2$ and $1 < \beta^2$. Since $1 < \alpha^2$ and $1 < \beta^2$ would imply that $1 < |\alpha\beta|$, which is impossible, we thus conclude that $1 > \alpha^2$ and $1 > \beta^2$, that is, $|\alpha| < 1$ and $|\beta| < 1$.

Now suppose that $\alpha = r(\cos\theta + i\sin\theta)$ and $\beta = \bar{\alpha} = r(\cos\theta - i\sin\theta)$ are complex conjugates with $r < 1$. Then

$$|\alpha+\beta| = |2r\cos\theta| \leqq 2r < 1 + r^2 = |1+\alpha\beta|$$

since $1 + r^2 - 2r = (1-r)^2 > 0$. Also $|\alpha\beta| = r^2 < 1$ implies that $|\alpha| = |\beta| = r < 1$. Thus $|\alpha\beta| < 1$ if, and only if, $|\alpha| < 1$ and $|\beta| < 1$. This completes the proof of the lemma.

We now prove the theorem.

THEOREM. *Let A be a 2×2 matrix over $\mathbb{R}$. Then $\{A^n\}$ converges to $\mathbb{O}$ if, and only if,* $|\det(A)| < 1$ *and* $|1 + \det(A)| > |\operatorname{tr}(A)|$.

Proof. The characteristic polynomial of A can be written as

$$\begin{aligned} p(x) &= x^2 - (a+d)x + (ad-bc) \\ &= x^2 - \operatorname{tr}(A)x + \det(A) \\ &= x^2 - (\alpha+\beta)x + \alpha\beta \end{aligned}$$

where α and β are the eigenvalues of A. Since $p(x)$ is a polynomial with real coefficients, α and β are either both reals or are complex conjugates. The result then follows from the lemma and the proposition.

Notes. The lemma is not true in general for any complex numbers α and β. For example, consider $\alpha = \beta = \frac{1}{2}i$. Then $|\alpha| = |\beta| = \frac{1}{2} < 1$. But $|1+\alpha\beta| = |1 - \frac{1}{4}| = \frac{3}{4} < 1 = |i/2 + i/2| = |\alpha+\beta|$. This also shows that the theorem is not true if A is over the complex numbers. Let $A = \begin{pmatrix} i/2 & 0 \\ 0 & i/2 \end{pmatrix}$. Then the same computation as above shows that $|\det(A)| = |-\frac{1}{4}| = \frac{1}{4} < 1$ whereas $|1 + \det(A)| = \frac{3}{4} < 1 = |i| = |\operatorname{tr}(A)|$. Thus A does not satisfy the conditions in the theorem. However, since $A = (\frac{1}{2}i)\mathbb{I}$ is a scalar matrix, we see that $A^n = (\frac{1}{2}i)^n\mathbb{I}$. Since $|\frac{1}{2}i| < 1$, $\{A^n\}$ converges to $\mathbb{O}$. Observe that the convergence of $\{A^n\}$ to $\mathbb{O}$ is also a consequence of the proposition since the eigenvalues of A are $\alpha = \beta = i/2$. Of course, the theorem remains true as long as the characteristic polynomial of A is a polynomial with real coefficients.

Finally, we remark that the theorem is not true for 3×3 matrices. Consider the matrix

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Then $|\det(A)| = 0 < 1$ and $|1 + \det(A)| = 1 > |\mathrm{tr}(A)| = 0$ and A satisfies the conditions in the theorem. However, it is easy to see that

$$A^n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq \mathbb{O} \quad \text{for all } n.$$

Thus $\{A^n\}$ does not converge to $\mathbb{O}$.

Acknowledgement. The author would like to thank his colleague Michael Woltermann for finding the equation of the ellipse in one of the above examples. The author also wants to thank the referees and the former editor for their valuable suggestions.

REFERENCES

1. Charles Cullen, *Matrices and Linear Transformations*, Addison-Wesley Publishing Co., Reading, MA, 1966, p. 173.
2. John Kemeny and J. Laurie Snell, *Finite Markov Chains*, D. Van Nostrand Co., Inc., New York, NY, 1960.
3. Gareth Williams, *Linear Algebra with Applications*, 2nd edition, Wm. C. Brown Publishers, Dubuque, IA, 1991.
4. Kenneth S. Williams, The nth power of a 2×2 matrix, this MAGAZINE, **65** (1992), 336.