ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 44, Number 4, 2014

A FIXED POINT APPROACH TO THE STEADY STATE FOR STOCHASTIC MATRICES

ROBERT KANTROWITZ AND MICHAEL M. NEUMANN

ABSTRACT. We provide two conditions, both in the spirit of classical regularity, that are equivalent to the existence of the steady state for a stochastic matrix. Our development of these characterizations sidesteps Perron-Frobenius theory for non-negative matrices, hinging instead on an elementary fixed point result that complements Banach's contraction mapping theorem.

1. Introduction. This note addresses a certain convergence property associated with a class of stochastic matrices. To fix notation, let $n \in \mathbf{N}$ be given, and let $\mathbf{P}^n$ denote the set of all probability vectors in $\mathbf{R}^n$. Thus, $\mathbf{P}^n$ consists of all column vectors $x \in \mathbf{R}^n$ that satisfy $x \geq 0$ in the componentwise order of $\mathbf{R}^n$ as well as $x_1 + \cdots + x_n = 1$. An $n \times n$ matrix A is said to be *stochastic*, provided that all column vectors of A belong to $\mathbf{P}^n$. If A is stochastic, then it is immediate that $Ax \in \mathbf{P}^n$ for all $x \in \mathbf{P}^n$. In particular, all powers of a stochastic matrix are stochastic. Also, a stochastic matrix is said to be *regular* (or primitive) if all entries of one of its powers are strictly positive.

Our main interest lies in those stochastic $n \times n$ matrices A for which there exists some $x \in \mathbf{P}^n$ such that, for each choice of $u \in \mathbf{P}^n$, the componentwise convergence $A^k u \to x$ as $k \to \infty$ obtains. In this case, the vector x is certainly unique and is called the *steady state* of A. Note that x satisfies $Ax = x$, since $A^k x \to x$, and hence $A^{k+1}x \to Ax$, as $k \to \infty$. Steady states play a fundamental role in the theory of Markov chains, as witnessed, for instance, by [**2, 7, 8**].

It is well known that the steady state exists for every regular stochastic matrix; see [**7,** Theorem 15.3.2] or [**8,** Theorem 4.2]. This

---

2010 AMS *Mathematics subject classification.* Primary 15B51, Secondary 37C25.

Keywords and phrases. Stochastic matrix, regular matrix, scrambling matrix, steady state, Perron-Frobenius theory, contractive mappings and fixed points.

Received by the editors on March 10, 2010, and in revised form on April 24, 2011.

DOI:10.1216/RMJ-2014-44-4-1243

classical result has become known as the fundamental theorem for Markov chains. It provides, among many other things, the theoretical background for some of the commonly used search algorithms on the internet, see [**6**]. The traditional proof of this result is based on the Perron-Frobenius theory for non-negative matrices and may be found in [**4, 7, 8**]. For this approach, the decisive point is that every regular stochastic matrix A has 1 as both a simple and dominant eigenvalue with a strictly positive eigenvector. The existence of the steady state may then be obtained from an inspection of the Jordan canonical form of A.

However, simple examples illustrate that not all stochastic matrices admit a steady state and that the classical condition of regularity is far from necessary in this context. In the present note, we establish that two weakened versions of the regularity condition on a stochastic matrix A are each both necessary and sufficient for the existence of the steady state. One of these two conditions just means that some power of A contains a strictly positive row. The other one amounts to $(A^p)^T A^p$ being strictly positive for some $p \in \mathbf{N}$. In geometric terms, the latter condition holds precisely when A^p is contractive on $\mathbf{P}^n$ with respect to a canonical metric.

Our elementary and natural approach completely avoids the spectral theory of non-negative matrices and any specifics of the theory of Markov chains. Instead, we employ a simple fixed point theorem that may be viewed as a counterpart of Banach's classical contraction principle for the case of compact metric spaces. Moreover, while the computation of eigenvalues for matrices of high order tends to be a difficult task, it turns out that our conditions of weak regularity type are remarkably easy to check.

2. A fixed point theorem. Given a non-empty metric space (S, ρ), a function $f : S \to S$ is said to be *contractive*, provided that

$$\rho\,(f(u), f(v)) < \rho(u, v) \quad \text{for all } u, v \in S \text{ with } u \neq v.$$

For arbitrary $m \in \mathbf{N}$, let $f^m = f \circ \cdots \circ f$ denote the m-fold composition of f on S, and let

$$f^\infty(S) = \bigcap_{m=1}^{\infty} f^m(S)$$

stand for the *generalized range* of f.

Theorem 2.1 below is an extension of a classical fixed point result due to Nemytzki and Edelstein, see Theorem 2.2 of either [**1**] or [**3**]. Our elementary proof employs the generalized range as a useful new tool in this context. Notice that the characteristic function of the rationals on $[0, 1]$ provides an example of a nowhere continuous function for which some iterate is contractive. Moreover, it will become clear that, even in the setting of stochastic matrices, there are natural examples of non-contractive mappings f for which f^p is contractive for some $p \in \mathbf{N}$.

Theorem 2.1. *Let (S, ρ) be a compact metric space, and suppose that $f : S \to S$ is a function for which f^p is continuous for some $p \in \mathbf{N}$. Then the following assertions are equivalent*:

(a) *f^p is contractive on $f^\infty(S)$*;
(b) *$f^\infty(S)$ is a singleton*;
(c) *f has a unique fixed point $x \in S$, and the convergence $f^k(u) \to x$ as $k \to \infty$ holds uniformly over all $u \in S$.*

Proof. We first show that (a) implies (b). Let $g = f^p$, and observe that $g^\infty(S)$ is invariant under g. Also, as the intersection of a decreasing sequence of non-empty compact sets, $g^\infty(S)$ is non-empty and compact.

To see that g induces a surjection on $g^\infty(S)$, let $u \in g^\infty(S)$ be given, and choose elements $u_k \in g^{k-1}(S)$ such that $u = g(u_k)$ for all $k \in \mathbf{N}$ with $k \geq 2$. By the compactness of S, there exists a subsequence $(u_{k(j)})_{j\in\mathbf{N}}$ that converges to some $v \in S$. For arbitrary $m \in \mathbf{N}$, we obtain $u_{k(j)} \in g^m(S)$ whenever $j > m$, and therefore $v \in g^m(S)$. Thus, $v \in g^\infty(S)$, and, by continuity, $g(u_{k(j)}) \to g(v)$ as $j \to \infty$. Because $g(u_{k(j)}) = u$ for all $j \in \mathbf{N}$, we conclude that $u = g(v)$ for some $v \in g^\infty(S)$, as desired.

Now, by the continuity of ρ on the compact set $g^\infty(S) \times g^\infty(S)$, there exist points $x, y \in g^\infty(S)$ with the property that $\rho(x, y) \geq \rho(u, v)$ for all $u, v \in g^\infty(S)$. If $x \neq y$, then $x = g(u)$ and $y = g(v)$ for distinct points $u, v \in g^\infty(S)$ and, hence, by condition (a),

$$\rho(u, v) \leq \rho(x, y) = \rho(g(u), g(v)) < \rho(u, v),$$

which is impossible. Thus, $g^\infty(S)$ is a singleton when (a) holds. Since $f^\infty(S)$ coincides with $g^\infty(S)$, we see that (a) implies (b).

Next suppose that (b) holds. Thus, $g^\infty(S) = f^\infty(S) = \{x\}$ for some $x \in S$. Since $f^\infty(S)$ is invariant under f and contains all the fixed points of f, it follows that f has the unique fixed point x. To establish the last part of (c), let $\varepsilon > 0$ be given, and let U consist of all $u \in S$ for which $\rho(u, x) < \varepsilon$. Since the sets U and $S \setminus g^m(S)$ for all $m \in \mathbf{N}$ form an open cover of S and $g^{m+1}(S) \subseteq g^m(S)$ for each $m \in \mathbf{N}$, we obtain, by compactness, some $m \in \mathbf{N}$ for which $U \cup (S \setminus g^m(S)) = S$, and therefore $g^m(S) \subseteq U$. For each $k \in \mathbf{N}$ with $k \geq mp$, we conclude that $f^k(S) \subseteq g^m(S) \subseteq U$, and therefore $\rho(f^k(u), x) < \varepsilon$ for all $u \in S$, as desired. Thus (b) implies (c).

Finally, suppose that (c) holds. Clearly, $x \in f^\infty(S)$. Moreover, given an arbitrary point $u \in f^\infty(S)$, we choose $u_k \in S$ for which $u = f^k(u_k)$ for all $k \in \mathbf{N}$. Since

$$\rho(u, x) = \rho(f^k(u_k), x) \leq \sup\{\rho(f^k(w), x) : w \in S\} \longrightarrow 0$$

as $k \to \infty$, we conclude that $f^\infty(S) = \{x\}$. Thus, (c) implies (b), and (a) is immediate from (b). $\square$

In general, the equivalent conditions of the preceding result are strictly stronger than the requirement that f admits a unique fixed point. Indeed, if $f(u) = 1 - u$ for all $u \in [0, 1]$, then $x = 1/2$ is the sole fixed point of f, whereas $f^k(u) \to x$ as $k \to \infty$ only when $u = x$ and $f^\infty([0, 1]) = [0, 1]$, since f is surjective.

Moreover, if f satisfies the condition that $\rho(f(u), f(v)) \leq \rho(u, v)$ for all $u, v \in S$, then it is easily seen that $f^\infty(S)$ consists precisely of all partial limits of $(f^k(u))_{k\in\mathbf{N}}$ for arbitrary $u \in S$. Given this description of $f^\infty(S)$, the condition of uniform convergence in part (c) of Theorem 2.1 turns out to be equivalent to that of pointwise convergence.

3. Steady states. We now explore the extent to which stochastic $n \times n$ matrices are contractive on $\mathbf{P}^n$ with respect to the norm given by $\|x\|_1 = |x_1| + \cdots + |x_n|$ for all $x \in \mathbf{R}^n$.

There are two elementary facts about inequalities for real numbers that will be used several times. First, if the vectors $u, v \in \mathbf{R}^n$ satisfy $u \leq v$, then the identity $u_1 + \cdots + u_n = v_1 + \cdots + v_n$ ensures that $u = v$. Second, if $u \in \mathbf{R}^n$ satisfies $|u_1 + \cdots + u_n| = |u_1| + \cdots + |u_n|$, then $u \geq 0$ or $u \leq 0$. As a simple application, we obtain the following characterization.

Lemma 3.1. *For arbitrary $u, v \in \mathbf{R}^n$ with $u, v \geq 0$, the identity $\|u - v\|_1 = \|u\|_1 + \|v\|_1$ holds precisely when $u_j v_j = 0$ for $j = 1, \ldots, n$.*

Proof. If the latter condition is satisfied, then the sets $J(u)$ and $J(v)$ consisting of all $j \in \{1, \ldots, n\}$ for which, respectively, $u_j > 0$ and $v_j > 0$ are disjoint, so that

$$\|u - v\|_1 = \sum_{j \in J(u)} u_j + \sum_{j \in J(v)} v_j = \|u\|_1 + \|v\|_1 .$$

Conversely, suppose that $\|u - v\|_1 = \|u\|_1 + \|v\|_1$. Then the inequalities $|u_j - v_j| \leq |u_j| + |v_j|$ for $j = 1, \ldots, n$ have to be identities, which implies that $u_j, -v_j \geq 0$ or $u_j, -v_j \leq 0$ for $j = 1, \ldots, n$. Because $u, v \geq 0$, this entails that $u_j v_j = 0$ for $j = 1, \ldots, n$, as claimed. □

Every stochastic $n \times n$ matrix $A = (a_{jk})$ satisfies condition $(*)$:

$$\|Ax\|_1 = \sum_{j=1}^{n} \left| \sum_{k=1}^{n} a_{jk} x_k \right| \leq \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk} \, |x_k| = \sum_{k=1}^{n} |x_k| \sum_{j=1}^{n} a_{jk} = \|x\|_1$$

for all $x \in \mathbf{R}^n$, with equality holding, for instance, when x is any of the standard basis vectors $e_1, \ldots, e_n$ of $\mathbf{R}^n$. While the strict estimate $\|Au - Av\|_1 < \|u - v\|_1$ fails for arbitrary distinct $u, v \in \mathbf{R}^n$, it is still possible that A is contractive on $\mathbf{P}^n$. This issue is addressed in the following result.

Proposition 3.2. *For each stochastic $n \times n$ matrix $A = (a_{jk})$, the following assertions are equivalent:*

(a) $\|Au - Av\|_1 < \|u - v\|_1$ *for all $u, v \in \mathbf{P}^n$ with $u \neq v$;*
(b) $\|Ae_k - Ae_\ell\|_1 < \|e_k - e_\ell\|_1$ *for all distinct $k, \ell \in \{1, \ldots, n\}$;*
(c) *for all $k, \ell \in \{1, \ldots, n\}$, there exists some $j \in \{1, \ldots, n\}$ for which $a_{jk}, a_{j\ell} > 0$;*
(d) *all entries of $A^T A$ are strictly positive.*

In particular, these equivalent conditions hold when A contains a row with strictly positive entries.

Proof. Clearly, (a) implies (b). Moreover, since $\|e_k - e_\ell\|_1 = 2 = \|Ae_k\|_1 + \|Ae_\ell\|_1$ for all distinct $k, \ell \in \{1, \dots, n\}$, the equivalence of (b) and (c) is immediate from Lemma 3.1, while (d) is just a convenient reformulation of (c). Finally, assume that condition (c) is satisfied, but not (a). Then there exist distinct vectors $u, v \in \mathbf{P}^n$ for which $\|Au - Av\|_1 = \|u - v\|_1$. This means that equality obtains in $(*)$ for the non-zero vector $x = u - v$. We conclude that, for each $j \in \{1, \dots, n\}$, the identity

$$\left| \sum_{k=1}^{n} a_{jk} x_k \right| = \sum_{k=1}^{n} |a_{jk} x_k|$$

holds, which implies that $a_{j1}x_1, \dots, a_{jn}x_n \geq 0$ or $a_{j1}x_1, \dots, a_{jn}x_n \leq 0$. Now, since x is non-zero, we have $x_k \neq 0$ for some k. Given an arbitrary $\ell \in \{1, \dots, n\}$, condition (c) yields some index j such that $a_{jk}, a_{j\ell} > 0$. Thus $x_k, x_\ell \geq 0$ or $x_k, x_\ell \leq 0$, and hence $x \geq 0$ if $x_k > 0$, while $x \leq 0$ if $x_k < 0$. It follows that $u \geq v$ or $u \leq v$. Because $u_1 + \cdots + u_n = v_1 + \cdots + v_n$, in either case, we arrive at $u = v$, the desired contradiction. $\square$

Matrices with property (c) play an important role in ergodic theory and are called *scrambling* in **[8]**. As an immediate consequence of the preceding results, we obtain the following characterizations for the existence of the steady state.

Theorem 3.3. *For every stochastic $n \times n$ matrix A, the following assertions are equivalent*:

(a) *there exists a unique $x \in \mathbf{P}^n$ such that $A^k u \to x$ as $k \to \infty$ for all $u \in \mathbf{P}^n$*;
(b) *some power of A contains a row with strictly positive entries*;
(c) *some power of A is scrambling*;
(d) *some power of A is contractive on $\mathbf{P}^n$.*

Moreover, if these equivalent conditions are satisfied, then the convergence $A^k u \to x$ as $k \to \infty$ holds uniformly over all $u \in \mathbf{P}^n$.

Proof. If (a) holds, then $A^k e_j \to x$ as $k \to \infty$ for $j = 1, \dots, n$. Hence, the powers of A converge entrywise to the $n \times n$ matrix that has x as each of its column vectors. When $j \in \{1, \dots, n\}$ satisfies $x_j > 0$, it follows that the jth row of A^k is strictly positive for sufficiently large

$k \in \mathbf{N}$. Thus, (a) entails (b), and the rest of the proof is clear from Proposition 3.2 and Theorem 2.1. $\square$

Theorem 3.3 complements Theorem 4.7 and Exercise 4.9 of [**8**], which characterize the existence of the steady state in terms of a certain Markov-theoretic condition. The proof of this result relies heavily on the Perron-Frobenius theory.

In the setting of Theorem 3.3, it turns out that the jth component of the steady state x is strictly positive precisely when the jth row of some and hence almost all powers of A are strictly positive. Indeed, this is immediate from the preceding proof and the identity $A^k x = x$ for all $k \in \mathbf{N}$. In particular, x is strictly positive if and only if A is regular.

As mentioned in [**5**], condition (b) of Theorem 3.3 holds precisely when 1 is both a simple and dominant eigenvalue of A. It is also shown in [**5**] that condition (b) forces the (n^2-3n+3)th power of A to contain a strictly positive row and that this exponent is optimal in the case of the Wielandt-type matrix

$$W_n = \begin{bmatrix} 0 & \alpha & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & \beta & 0 & \cdots & 0 \end{bmatrix},$$

where $\alpha, \beta > 0$ satisfy $\alpha + \beta = 1$. By Theorem 3.3 and Proposition 3.2, we conclude that condition (c) of Theorem 3.3 implies that the $(n^2 - 3n + 3)$th power of A is contractive on $\mathbf{P}^n$, but we are not aware of an optimal result in this direction. For partial results, see [**5**].

The problem that we are facing here is illuminated by the Wielandt-type matrix. By Problem 8.5.4 of [**4**], the pth power of W_n is strictly positive precisely when $p \geq n^2 - 2n + 2$. In the case $n = 10$, this means that $p \geq 82$. On the other hand, a computer algebra system may be used to show that, at least for $n \leq 10$, the optimal contractivity exponent for W_n is given by the smallest integer greater than or equal to $(n^2 - 2n + 2)/2$. In particular, it follows that the pth power of W_{10} is contractive on $\mathbf{P}^{10}$ if and only if $p \geq 41$, while the pth power of W_{10} admits a strictly positive row precisely when $p \geq 73$.

REFERENCES

1. V. Berinde, *Iterative approximation of fixed points*, Springer, Berlin, 2007.

2. A. Berman and R.J. Plemmons, *Nonnegative matrices in the mathematical sciences*, Academic Press, New York, 1979.

3. K. Goebel and W.A. Kirk, *Topics in metric fixed point theory*, Cambridge University Press, Cambridge, 1990.

4. R.A. Horn and C.R. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge, 1985.

5. R. Kantrowitz, M.M. Neumann, and T.J. Ransford, *Regularity, scrambling, and the steady state for stochastic matrices*, Contemp. Math. **547** (2011), 153–164.

6. A.N. Langville and C.D. Meyer, *Google's PageRank and beyond*: *The science of search engine rankings*, Princeton University Press, Princeton, NJ, 2006.

7. C.R. Rao and M.B. Rao, *Matrix algebra and its applications to statistics and econometrics*, World Scientific, Singapore, 1998.

8. E. Seneta, *Non-negative matrices and Markov chains*, 2nd edition, Springer, New York, 1981.

Department of Mathematics, Hamilton College, 198 College Hill Road, Clinton, NY 13323
Email address: **rkantrow@hamilton.edu**

Department of Mathematics and Statistics, Mississippi State University, Mississippi State, MS 39762
Email address: **neumann@math.msstate.edu**