

# Somewhat Stochastic Matrices

Branko Ćurgus and Robert I. Jewett

## 1 Introduction.

The notion of a Markov chain is ubiquitous in linear algebra and probability books. For example see [3, Theorem 5.25] and [2, p. 173]. Also, see [5, p. 131] for the history of the subject. A *Markov* (or *stochastic*) *matrix* is a square matrix whose entries are non-negative and whose column sums are equal to 1. The term stochastic matrix seems to prevail in current literature and therefore we use it in the title. But, since a Markov matrix is a transition matrix of a Markov chain, we prefer the term Markov matrix and we use it from now on. The theorem below gives some of the standard results for such matrices.

**Theorem 1.1.** *Let  $M$  be an  $n \times n$  Markov matrix. Suppose that there exists  $p \in \mathbb{N}$  such that all the entries of  $M^p$  are positive. Then the following statements are true.*

(a) *There exists a unique  $E \in \mathbb{R}^n$  such that*

$$ME = E \quad \text{and} \quad \text{sum } E = 1.$$

(b) *Let  $P$  be the square matrix each of whose columns is equal to  $E$ . Then  $P$  is a projection and  $PX = (\text{sum } X)E$  for each  $X \in \mathbb{R}^n$ .*

(c) *The powers  $M^k$  tend to  $P$  as  $k$  tends to  $+\infty$ .*

The statement that all the entries of some power of  $M$  are positive is usually abbreviated by saying that  $M$  is *regular*. The fact that all the entries of  $E$  are positive is easily shown, since  $M^p E = E$ ,  $\text{sum } E = 1$ , and  $M^p$  is positive.

Theorem 1.1 follows readily from Theorem 4.2, the main result in this article. In Theorem 4.2 the requirement that the entries of  $M$  be non-negative is dropped, the requirement that the column sums be equal to 1 is retained, and the condition on  $M^p$  is replaced by something completely different. However, the conclusions (a), (b), (c) hold true. Our proof is significantly different from all proofs of Theorem 1.1 that we are aware of.

Here is an example:

$$M = \frac{1}{5} \begin{bmatrix} 0 & 2 & -4 \\ -1 & -1 & 0 \\ 6 & 4 & 9 \end{bmatrix}.$$

Examining the first ten powers of  $M$  with a computer strongly suggests that the powers of  $M$  converge. Indeed, Theorem 4.2 applies here. For this, one must examine  $M^2$ ; see Example 6.1. The limit is found, as in the case of a Markov matrix, by determining an eigenvector of  $M$ . It turns out that  $E = \frac{1}{3} [-6 \ 1 \ 8]^T$ , and

$$\lim_{k \rightarrow +\infty} M^k = \frac{1}{3} \begin{bmatrix} -6 & -6 & -6 \\ 1 & 1 & 1 \\ 8 & 8 & 8 \end{bmatrix}.$$

Since  $M$  is only a  $3 \times 3$  matrix, we could show convergence by looking at the eigenvalues, which are  $1, 2/5, 1/5$ .

Theorem 1.1 is often presented as an application of the Perron-Frobenius Theorem. In [6], the authors give a version of the Perron-Frobenius Theorem for matrices with some negative entries, but their results do not seem to be related to ours.

## 2 Definitions.

All numbers in this article are real, except in Example 6.4. We study  $m \times n$  matrices with real entries. The elements of  $\mathbb{R}^n$  will be identified with column matrices, that is, with  $n \times 1$  matrices. By  $J$  we denote any row matrix with all entries equal to 1.

Let  $X \in \mathbb{R}^n$  with entries  $x_1, \dots, x_n$ . Set

$$\text{sum } X := \sum_{j=1}^n x_j, \quad \|X\| := \sum_{j=1}^n |x_j|.$$

Notice that  $\text{sum } X = JX$  and that  $\|\cdot\|$  is the  $\ell_1$ -norm.

For an  $m \times n$  matrix  $A$  with columns  $A_1, \dots, A_n$  the *variation* (or *column variation*) of  $A$  is defined by:

$$\text{var } A := \frac{1}{2} \max_{1 \leq j, k \leq n} \|A_j - A_k\|.$$

If the column sums of a matrix  $A$  are all equal to  $a$ , that is if  $JA = aJ$ , we say that  $A$  is of *type*  $a$ .

## 3 Column variation and matrix type.

In this section we establish the properties of the column variation and the matrix type that are needed for the proof of our main result. We point out that the restriction to real numbers in the theorem below is essential, as Example 6.4 shows.

**Theorem 3.1.** *Let  $A$  be an  $m \times n$  matrix and  $X \in \mathbb{R}^n$ . If  $\text{sum } X = 0$ , then*

$$\|AX\| \leq (\text{var } A)\|X\|.$$

*Proof.* In this proof, for any real number  $t$  we put  $t^+ := \max\{t, 0\}$  and  $t^- := \max\{-t, 0\}$ . Clearly  $t^+, t^- \geq 0$  and  $t = t^+ - t^-$ .

Let  $A_1, \dots, A_n$  be the columns of  $A$ . Assume  $\text{sum } X = 0$ . The conclusion is obvious if  $X = 0$ .

Assume that  $X \neq 0$ . Then, by scaling, we can also assume that  $\|X\| = 2$ . Let  $x_1, \dots, x_n$  be the entries of  $X$ . Then

$$\sum_{k=1}^n x_k^+ = \sum_{k=1}^n x_k^- = 1.$$

Consequently

$$AX = \sum_{k=1}^n x_k A_k = \sum_{k=1}^n x_k^+ A_k - \sum_{j=1}^n x_j^- A_j.$$

Now we notice that  $AX$  is the difference of two convex combinations of the columns of  $A$ . From this, the inequality in the theorem seems geometrically obvious. However, we continue with an algebraic argument:

$$\begin{aligned} AX &= \sum_{k=1}^n \sum_{j=1}^n x_j^- x_k^+ A_k - \sum_{j=1}^n \sum_{k=1}^n x_k^+ x_j^- A_j \\ &= \sum_{k=1}^n \sum_{j=1}^n x_j^- x_k^+ (A_k - A_j). \end{aligned}$$

Consequently,

$$\begin{aligned} \|AX\| &\leq \sum_{k=1}^n \sum_{j=1}^n x_j^- x_k^+ \|A_k - A_j\| \\ &\leq 2(\text{var } A) \sum_{k=1}^n x_k^+ \sum_{j=1}^n x_j^- \\ &= (\text{var } A) \|X\|. \quad \square \end{aligned}$$

**Proposition 3.2.** *Let  $A$  and  $B$  be matrices such that  $AB$  is defined. If  $B$  is of type  $b$ , then*

$$\text{var}(AB) \leq (\text{var } A)(\text{var } B).$$

*Proof.* Assume that  $B$  is of type  $b$  and let  $B_1, \dots, B_l$  be the columns of  $B$ . Then  $AB_1, \dots, AB_l$  are the columns of  $AB$ . Since  $B$  is of type  $b$ , for all  $j, k \in \{1, \dots, l\}$  we have  $\text{sum}(B_j - B_k) = 0$ . Therefore, by Theorem 3.1,

$$\|AB_j - AB_k\| \leq (\text{var } A) \|B_j - B_k\|$$

for all  $j, k \in \{1, \dots, l\}$ . Hence,

$$\begin{aligned} \text{var}(AB) &= \frac{1}{2} \max_{1 \leq j, k \leq l} \|AB_j - AB_k\| \\ &\leq (\text{var } A) \frac{1}{2} \max_{1 \leq j, k \leq l} \|B_j - B_k\| \\ &= (\text{var } A)(\text{var } B). \end{aligned} \quad \square$$

**Proposition 3.3.** *Let  $A$  and  $B$  be matrices such that  $AB$  is defined. If  $A$  is of type  $a$  and  $B$  is of type  $b$ , then  $AB$  is of type  $ab$ .*

*Proof.* If  $JA = aJ$  and  $JB = bJ$ , then  $J(AB) = (JA)B = aJB = abJ$ .  $\square$

## 4 Square matrices.

In the previous section we considered rectangular matrices. Next we study square matrices. With one more property of matrix type, we shall be ready to prove our main result, Theorem 4.2.

**Proposition 4.1.** *If  $M$  is a square matrix of type  $c$ , then  $c$  is an eigenvalue of  $M$ .*

*Proof.* Assume that  $JM = cJ$ . Then  $J(M - cI) = 0$ . That is, the sum of the rows of  $M - cI$  is 0 and so the rows of  $M - cI$  are linearly dependent. Hence,  $M - cI$  is a singular matrix.  $\square$

**Theorem 4.2.** *Let  $M$  be an  $n \times n$  matrix. Suppose that  $M$  is of type 1 and that there exists  $p \in \mathbb{N}$  such that  $\text{var}(M^p) < 1$ . Then the following statements are true.*

(a) *There exists a unique  $E \in \mathbb{R}^n$  such that*

$$ME = E \quad \text{and} \quad \text{sum } E = 1.$$

(b) *Let  $P$  be the square matrix each of whose columns is equal to  $E$ . Then  $P$  is a projection and  $PX = (\text{sum } X)E$  for each  $X \in \mathbb{R}^n$ .*

(c) *The powers  $M^k$  tend to  $P$  as  $k$  tends to  $+\infty$ .*

*Proof.* Assume that  $M$  is of type 1 and that there exists  $p \in \mathbb{N}$  such that  $\text{var}(M^p) < 1$ . By Proposition 4.1, there exists a nonzero  $Y \in \mathbb{R}^n$  such that  $MY = Y$ .

Clearly  $M^p Y = Y$ . If  $\text{sum } Y = 0$ , then, since  $Y \neq 0$ , Theorem 3.1 yields

$$\|Y\| = \|M^p Y\| \leq \text{var}(M^p) \|Y\| < \|Y\|,$$

a contradiction. Setting  $E = (1/\text{sum } Y)Y$  provides a vector whose existence is claimed in (a). To verify uniqueness, let  $F$  be another such vector. Then  $\text{sum}(E - F) = 0$ ,  $M^p(E - F) = E - F$ , and

$$\|E - F\| = \|M^p(E - F)\| \leq \text{var}(M^p) \|E - F\|.$$

Consequently,  $E - F = 0$ , since  $\text{var}(M^p) < 1$ .

By the definition of  $P$  in (b),  $P = EJ$ . Therefore,  $P^2 = (EJ)(EJ) = E(JE)J = E[1]J = EJ = P$ . To complete the proof of (b), we calculate:  $PX = E(JX) = (\text{sum } X)E$ .

Let  $k \in \mathbb{N}$ . Proposition 3.3 implies that  $M^k$  is of type 1. By the division algorithm there exist unique  $q, r \in \mathbb{Z}$  such that  $k = pq + r$  and  $0 \leq r < p$ . Here  $q = \lfloor k/p \rfloor$  is the floor of  $k/p$ . By Proposition 3.2,

$$\text{var}(M^k) \leq (\text{var } M)^r (\text{var}(M^p))^q \leq \left( \max_{0 \leq r < p} (\text{var } M)^r \right) (\text{var}(M^p))^{\lfloor k/p \rfloor}. \quad (4.1)$$

Let  $X \in \mathbb{R}^n$  be such that  $\text{sum } X = 1$ . Then  $\text{sum}(X - E) = 0$  and Theorem 3.1 implies that

$$\|M^k X - E\| = \|M^k(X - E)\| \leq \text{var}(M^k) \|X - E\|. \quad (4.2)$$

Now, since  $\text{var}(M^p) < 1$ , (4.1) implies that

$$\lim_{k \rightarrow +\infty} \text{var}(M^k) = 0,$$

and letting  $X$  in (4.2) run through the vectors in the standard basis of  $\mathbb{R}^n$  proves (c).  $\square$

**Remark 4.3.** If  $M$  is an  $n \times n$  matrix of type 1, and the statements (a), (b), and (c) are true, then there exists  $p \in \mathbb{N}$  such that  $\text{var}(M^p) < 1$ . This follows from the fact that the variation function is continuous on the space of all  $n \times n$  matrices.

## 5 Non-negative matrices.

The propositions in this section are useful for showing that the variation of a Markov matrix is less than 1. They are used to deduce Theorem 1.1 from Theorem 4.2 and also, repeatedly, in Example 6.3.

**Proposition 5.1.** *Let  $A$  be an  $m \times n$  matrix of type  $a$  with non-negative entries. Then  $\text{var } A \leq a$ .*

*Proof.* Let  $A_1, \dots, A_n$  be the columns of  $A$ . Since the entries of  $A$  are non-negative,  $\|A_j\| = a$  for all  $j \in \{1, \dots, n\}$ . Therefore,

$$\|A_j - A_k\| \leq \|A_j\| + \|A_k\| = 2a$$

for all  $j, k \in \{1, \dots, n\}$ , and the proposition follows.  $\square$

**Proposition 5.2.** *Let  $a > 0$ . Let  $A$  be an  $m \times n$  matrix of type  $a$  with non-negative entries. Then the following two statements are equivalent.*

- (i) *The strict inequality  $\text{var } A < a$  holds.*

(ii) For each  $k, l \in \{1, \dots, n\}$  there exists  $j \in \{1, \dots, m\}$  such that the  $j$ -th entries of the  $k$ -th and  $l$ -th columns of  $A$  are both positive.

*Proof.* Let  $A = [a_{jk}]$  and let  $A_1, \dots, A_n$  be the columns of  $A$ . To prove that (i) and (ii) are equivalent we consider their negations. By Proposition 5.1 and the definition,  $\text{var } A = a$  if and only if there exist  $k_0, l_0 \in \{1, \dots, n\}$  such that

$$\|A_{k_0} - A_{l_0}\| = 2a.$$

This is equivalent to

$$\sum_{j=1}^m |a_{jk_0} - a_{jl_0}| = \sum_{j=1}^m (a_{jk_0} + a_{jl_0}). \quad (5.1)$$

Since all the terms in the last equality are non-negative, (5.1) is equivalent to  $a_{jk_0}$  or  $a_{jl_0}$  being 0 for all  $j \in \{1, \dots, m\}$ .

Hence,  $\text{var } A = a$  if and only if there exist  $k_0, l_0 \in \{1, \dots, n\}$  such that for all  $j \in \{1, \dots, m\}$  we have  $a_{jk_0} a_{jl_0} = 0$ . This proves that (i) and (ii) are equivalent.  $\square$

Now we can give a short proof of Theorem 1.1.

*Proof.* Let  $M$  be a regular Markov matrix and assume that  $M^p$  is positive. By Proposition 5.2,  $\text{var}(M^p) < 1$ . Therefore Theorem 4.2 applies.  $\square$

## 6 Examples.

**Example 6.1.** In the Introduction we used

$$M = \frac{1}{5} \begin{bmatrix} 0 & 2 & -4 \\ -1 & -1 & 0 \\ 6 & 4 & 9 \end{bmatrix}$$

as an example of a matrix for which the powers converge. The largest  $\ell_1$ -distance between two columns is between the second and the third, and is equal to  $12/5$ . Therefore,  $\text{var } M = 6/5 > 1$ . But

$$M^2 = \frac{1}{25} \begin{bmatrix} -26 & -18 & -36 \\ 1 & -1 & 4 \\ 50 & 44 & 57 \end{bmatrix}$$

and  $\text{var}(M^2) = 18/25 < 1$ . Hence, Theorem 4.2 applies.

**Example 6.2.** For  $2 \times 2$  matrices of type 1 it is possible to give a complete analysis. Let  $a, b \in \mathbb{R}$  and set

$$M = \begin{bmatrix} 1-a & b \\ a & 1-b \end{bmatrix} \quad \text{and} \quad c = a+b.$$

Then  $\text{var } M = |1-c|$ . We distinguish the following three cases:

- (i)  $c \neq 0$ . The eigenvalues of  $M$  are 1 and  $1 - c$ , and the corresponding eigenvectors are  $\begin{bmatrix} b \\ a \end{bmatrix}$  and  $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ . If  $0 < c < 2$ , then  $\text{var } M < 1$ . Consequently,  $E = \frac{1}{c} \begin{bmatrix} b \\ a \end{bmatrix}$  and  $M^k$  converges. Otherwise,  $M^k$  diverges.
- (ii)  $c = 0, a \neq 0$ . In this case,  $\text{var } M = 1$  and 1 is an eigenvalue of multiplicity 2. It can be shown by induction that

$$M^k = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + k a \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}.$$

So  $M^k$  diverges.

- (iii)  $c = a = b = 0$ . So,  $M = I$ .

Thus, for a  $2 \times 2$  matrix  $M$  of type 1 which is not the identity matrix,  $M^k$  converges if and only if  $\text{var } M < 1$ . Regular  $2 \times 2$  Markov matrices were studied in [4].

**Example 6.3.** Consider the following three kinds of Markov matrices:

$$K = \begin{bmatrix} 1 & + & + \\ 0 & + & + \\ 0 & 0 & + \end{bmatrix}, \quad L = \begin{bmatrix} + & + & 0 \\ + & 0 & + \\ 0 & + & + \end{bmatrix}, \quad M = \begin{bmatrix} 0 & + & 0 \\ 0 & 0 & 1 \\ 1 & + & 0 \end{bmatrix}.$$

Here we use  $+$  for positive numbers. All claims below about the variation rely on Propositions 5.1 and 5.2.

The matrix  $K$  is not regular, but  $\text{var } K < 1$ . Also,  $E = [1 \ 0 \ 0]^T$ .

The matrix  $L$  is not positive, but  $\text{var } L < 1$ . Also, Theorem 1.1 applies since  $L^2$  is positive.

The first five powers of  $M$  are:

$$\begin{bmatrix} 0 & + & 0 \\ 0 & 0 & 1 \\ 1 & + & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & + \\ 1 & + & 0 \\ 0 & + & + \end{bmatrix}, \quad \begin{bmatrix} + & + & 0 \\ 0 & + & + \\ + & + & + \end{bmatrix}, \quad \begin{bmatrix} 0 & + & + \\ + & + & + \\ + & + & + \end{bmatrix}, \quad \begin{bmatrix} + & + & + \\ + & + & + \\ + & + & + \end{bmatrix}.$$

The variation of the first two matrices is 1, while  $\text{var}(M^3) < 1$ . The first positive power of  $M$  is  $M^5$ .

In fact, the following general statement holds. For a  $3 \times 3$  Markov matrix  $M$ , the sequence  $M^k, k \in \mathbb{N}$ , converges to a projection of rank 1 if and only if  $\text{var}(M^3) < 1$ . This was verified by examining all possible cases; see [1].

**Example 6.4.** In this example we consider matrices with complex entries. Let  $\alpha = (-1 + i\sqrt{3})/2$ . Then 1,  $\alpha$ , and  $\bar{\alpha}$  are the cube roots of unity. Notice that  $1 + \alpha + \bar{\alpha} = 0$ ,  $\alpha^2 = \bar{\alpha}$ , and  $\bar{\alpha}^2 = \alpha$ .

The examples below were suggested by the following orthogonal basis for the complex inner product space  $\mathbb{C}^3$ :

$$U = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad V = \begin{bmatrix} 1 \\ \alpha \\ \bar{\alpha} \end{bmatrix}, \quad W = \begin{bmatrix} 1 \\ \bar{\alpha} \\ \alpha \end{bmatrix}.$$

We first give an example which shows that the conclusion of Theorem 3.1 does not hold for matrices with complex entries. Set  $A = \begin{bmatrix} 1 & \bar{\alpha} & \alpha \end{bmatrix}$ . Then  $AV = [3]$ ,  $\text{sum } V = 0$ ,  $\text{var } A = \sqrt{3}/2 < 1$ , and

$$\|AV\| > (\text{var } A)\|V\|.$$

Next we give an example showing that the restriction to real numbers cannot be dropped in Theorem 4.2. Consider the matrices

$$P = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad Q = \frac{1}{3} \begin{bmatrix} 1 & \bar{\alpha} & \alpha \\ \alpha & 1 & \bar{\alpha} \\ \bar{\alpha} & \alpha & 1 \end{bmatrix}.$$

Notice that  $P$  is the orthogonal projection onto the span of  $U$  and  $Q$  is the orthogonal projection onto the span of  $V$ . Let  $c \in \mathbb{R}$  and set

$$M = P + cQ.$$

Then  $PV = 0$  and  $QV = V$ . Therefore  $MV = cV$ , showing that  $c$  is an eigenvalue of  $M$ .

The matrix  $P$  is of type 1 with variation 0, while  $Q$  is of type 0 with variation  $\sqrt{3}/2$ . Hence,  $M$  is of type 1 and

$$\text{var } M = \text{var}(cQ) = |c|\sqrt{3}/2.$$

Therefore, if  $1 < c < 2/\sqrt{3}$ , then  $\text{var } M < 1$ , but  $M^k$  diverges.

## 7 The variation as a norm.

The first proposition below shows that the variation function is a pseudo-norm on the space of all  $m \times n$  matrices. The remaining propositions identify the variation of a matrix as the norm of a related linear transformation.

**Proposition 7.1.** *Let  $A$  be  $m \times n$  matrix.*

- (a) *If  $c \in \mathbb{R}$ , then  $\text{var}(cA) = |c| \text{var } A$ .*
- (b) *All columns of  $A$  are identical if and only if  $\text{var } A = 0$ .*
- (c) *If  $B$  is another  $m \times n$  matrix, then  $\text{var}(A + B) \leq \text{var } A + \text{var } B$ .*

*Proof.* The proofs of (a) and (b) are straightforward. To prove (c), let  $A_1, \dots, A_n$  be the columns of  $A$  and let  $B_1, \dots, B_n$  be the columns of  $B$ . Then  $A_1 + B_1, \dots, A_n + B_n$  are the columns of  $A + B$ , and for all  $j, k \in \{1, \dots, n\}$ ,

$$\|(A_j + B_j) - (A_k + B_k)\| \leq \|A_j - A_k\| + \|B_j - B_k\|. \quad \square$$

**Proposition 7.2.** *Let  $A$  be an  $m \times n$  matrix with more than one column. Then*

$$\text{var } A = \max \left\{ \|AX\| : X \in \mathbb{R}^n, \|X\| = 1, \text{sum } X = 0 \right\}. \quad (7.1)$$

*Proof.* It follows from Theorem 3.1 that the set on the right-hand side of (7.1) is bounded by  $\text{var } A$ . To prove that  $\text{var } A$  is the maximum, let  $A_{j_0}$  and  $A_{k_0}$  be columns of  $A$  such that  $\|A_{j_0} - A_{k_0}\| = 2 \text{var } A$ . Choose  $X_0 \in \mathbb{R}^n$  such that its  $j_0$ -th entry is  $1/2$ , its  $k_0$ -th entry is  $-1/2$  and all other entries are 0. Then  $\|AX_0\| = \text{var } A$ ,  $\|X_0\| = 1$  and  $\text{sum}(X_0) = 0$ .  $\square$

**Remark 7.3.** Let  $\mathcal{V}_n$  denote the set of all vectors  $X \in \mathbb{R}^n$  such that  $\text{sum } X = 0$ . This is a subspace of  $\mathbb{R}^n$ . An  $m \times n$  matrix  $A$  determines a linear transformation from  $\mathcal{V}_n$  to  $\mathbb{R}^m$ . The previous proposition tells us that the norm of this transformation is  $\text{var } A$ .

**Proposition 7.4.** *Let  $B$  be an  $m \times n$  matrix of type  $b$  with more than one row. Then*

$$\text{var } B = \max\left\{\text{var}(ZB) : Z^T \in \mathbb{R}^m, \text{var } Z = 1\right\}. \quad (7.2)$$

*Proof.* If  $\text{var } B = 0$  the statement follows from Proposition 3.2. So, assume  $\text{var } B > 0$ . By Proposition 3.2, the set on the right-hand side of (7.2) is bounded by  $\text{var } B$ . Let  $B = [b_{jk}]$  and let  $B_{k_0}$  and  $B_{l_0}$  be columns of  $B$  such that  $\|B_{k_0} - B_{l_0}\| = 2 \text{var } B$ . Let  $Z_0$  be the row with entries defined by:

$$z_j = \begin{cases} 1 & \text{if } b_{jk_0} > b_{jl_0}, \\ -1 & \text{if } b_{jk_0} \leq b_{jl_0}. \end{cases}$$

Since  $\|B_{k_0} - B_{l_0}\| > 0$  and  $\text{sum}(B_{k_0}) = \text{sum}(B_{l_0}) = 0$ , there exist at least one positive and at least one negative entry in  $Z_0$ . Therefore,  $\text{var}(Z_0) = 1$ . By the definition of  $Z_0$  the difference between  $k_0$ -th and  $l_0$ -th entry in  $Z_0B$  is  $\|B_{k_0} - B_{l_0}\|/2 = \text{var } B$ . Notice that if  $Z$  is a row matrix, then  $\text{var } Z = \frac{1}{2}(\max Z - \min Z)$ . Therefore,  $\text{var}(Z_0B) \geq \text{var } B$ . This proves (7.2).  $\square$

**Remark 7.5.** Let  $\mathbb{R}_m$  denote the set of all row vectors with  $m$  entries. Denote by  $\mathbb{J}_m$  the subspace of all scalar multiples of  $J$ . An  $m \times n$  matrix  $B$  of type  $b$  determines a linear transformation from the factor space  $\mathbb{R}_m/\mathbb{J}_m$  to the factor space  $\mathbb{R}_n/\mathbb{J}_n$  in the following way:

$$Z + \mathbb{J}_m \mapsto ZB + \mathbb{J}_n, \quad Z \in \mathbb{R}_m.$$

It is easy to verify that this is a well-defined linear transformation. By Proposition 7.4, the norm of this transformation is exactly  $\text{var } B$ .

## References

- [1] B. urgus and R. I. Jewett, On the variation of  $3 \times 3$  stochastic matrices, (August 2007), available at <http://myweb.facstaff.wvu.edu/curgus/papers.html>.
- [2] J. L. Doob, *Stochastic Processes*, John Wiley & Sons, 1990; reprint of the 1953 original.

- [3] S. Friedberg, A. Insel, L. Spence, *Linear Algebra*, 4th ed., Prentice Hall, 2002.
- [4] N. J. Rose, On regular Markov chains, this MONTHLY, **92** (1985), 146.
- [5] E. Seneta, *Non-Negative Matrices and Markov Chains*, Springer, 2006.
- [6] P. Tarazaga, M. Raydan, and A. Hurman, Perron-Frobenius theorem for matrices with some negative entries. *Linear Algebra Appl.* **328** (2001), 57–68.

**Branko Ćurgus**

*Department of Mathematics, Western Washington University,  
Bellingham, Washington 98225*  
curgus@cc.wvu.edu

**Robert I. Jewett**

*Department of Mathematics, Western Washington University,  
Bellingham, Washington 98225*