A NOTE ON THE STOCHASTIC ROOTS OF STOCHASTIC MATRICES

Qi-Ming HE and **Eldon Gunn**

Department of Industrial Engineering
Dalhousie University, Halifax, Canada B3J 2X4
Qi-Ming.He@dal.ca and *Eldon.Gunn@dal.ca*

Abstract:

In this paper, we study the stochastic root matrices of stochastic matrices. All stochastic roots of 2×2 stochastic matrices are found explicitly. A method based on characteristic polynomial of matrix is developed to find all real root matrices that are functions of the original 3×3 matrix, including all possible (function) stochastic root matrices. In addition, we comment on some numerical methods for computing stochastic root matrices of stochastic matrices.

Keyword: Markov chain, stochastic matrix, matrix root, linear algebra.

---

1. Introduction

Let $A = (a_{i,j})$ be an $m\times m$ stochastic matrix, i.e., A is a nonnegative matrix satisfying $A\mathbf{e} = \mathbf{e}$, where m is a positive integer and $\mathbf{e}$ is the column vector with all elements being one. Let $\mathcal{A}_m$ be the set of all $m\times m$ stochastic matrices. The problem of interest in this paper can be stated as follows: for a positive integer n and $A\in\mathcal{A}_m$, find a matrix B such that

$$A = B^n, \quad B \geq 0, B\mathbf{e} = \mathbf{e}. \tag{1.1}$$

If there exists a stochastic matrix B to equation (1.1), we call B an nth stochastic root matrix of A (or an nth stochastic root). The objective of this paper is to find stochastic roots of a stochastic matrix of order 2 or 3 and to discuss some numerical methods for more general cases.

Finding a stochastic root for a stochastic matrix has many applications in stochastic modelling (see Egilsson [6], Neuts [14], and Seneta [15]). Quite often in practice, the use of available sample data for estimating the transition matrix of a Markov chain is limited. For instance, consider a Markov chain that describes the state of weather condition in an airport (Egilsson [6]). For certain purpose, it is required that the state of weather condition must be predicted on a quarter hour basis (every fifteen minutes). However, information about weather condition in the airport was only collected and recorded hourly. Using the existing data, the transition matrix A of the Markov chain for the state of the airport at the end of each hour can be estimated. Then we need to find a transition matrix B such that $A = B^4$ for a Markov chain to track the state of the airport on a quarter hour basis. Another example arises in risk management of portfolio. In that case, company's credit ratings are recorded yearly. Thus, when the process of changes in credit ratings is modelled as a Markov chain, the transition matrix for yearly changes can be estimated. However, investment horizon is shorter than a year, usually a month or a quarter of a year. Therefore, a quarterly (monthly) transition matrix is needed to estimate potential quarterly (monthly) loss resulting from companies in the portfolios being downgraded to a lower rating. Again, we need to find stochastic roots of stochastic matrices. These are just some examples of a common situation in which practitioners find that data has been collected on a time basis that is different from that which is necessary for modelling. Thus it is of great interest to practitioners to find stochastic roots of stochastic matrices.

Although the stochastic root problem seems a problem of broad interest, the study of the problem is limited. A related problem, known as Elfving's problem, is to find an infinitesimal generator T such that $A = \exp\{T\}$ for a stochastic matrix A (Elfving [7]). If there is a solution to Elfving's problem, then A has an nth stochastic root $B = \exp\{T/n\}$ for any positive n. It has been proved in Kingman [11] that Elfving's problem has a solution if and only if A has an nth stochastic root for all positive integer n. Therefore, Elfving's problem can be considered as a special case of the stochastic root problem (1.1). Since the existence conditions for stochastic roots are different for different n, the problem of interest in this paper is more complicated than Elfving's problem.

Finding a root matrix of a matrix appears quite often in the literature of matrix theory and its applications (Bhatia [3], Higham [10], Lancaster and Tismenetsky [12], Minc [13]). It is known that a stochastic solution to Problem (1.1) does not always exist and, if it exists, may not be unique. On the other hand, for some special classes of matrices, root matrices can be found explicitly. For instance, for a symmetric matrix, root matrices can be found. Bjorck and Hammarling [4] proposed a Schur method to find square roots of a matrix. Denman [5] and Higham [8, 9] developed methods to find roots of real matrices or real roots of real matrices. There are also algorithms developed for computing a particular root (Higham [9, 10]). However, there is no method developed to find all solutions for any m and n (Problem (1.1)). The problem becomes more complicated when we restrict the solutions to the set of stochastic matrices. Furthermore, it is widely known that there may not be a uniform and effective approach towards the matrix root problem. Thus, it is quite possible that we have to deal with the stochastic root problem (1.1) on a case-by-case basis. Therefore, there is a need to study special cases, such as the case with m=2 or m=3.

While it is possible to find all or some roots of matrices of order 2 or 3, it is quite challenging to find just one stochastic root if the order is large. Various approaches were utilized in the past to solve the matrix root problem (Bjorck and Hammarling [3], Higham [8, 10], etc.), including Taylor expansion, Newton's method, nonlinear programs, Pade approximation, etc. However, these methods cannot always find a stochastic root. In this paper, we briefly discuss three numerical methods that may find a solution to Problem (1.1) for any n and m.

The spectrum inverse problem in matrix theory is closely related to the stochastic root problem (e.g., Bapat and Ragharan [1] and Chapter 7 in Minc [13]). Useful results are obtained only for doubly stochastic matrices and positive matrices. Since the solution (if it exists) for the spectrum inverse problem is not unique and an algorithm for finding all the solutions has not been found yet (Bapat and Ragharan [1]), this approach may not be able to produce the expected solution (even if a solution can be found).

The rest of the paper is organized as follows. In Section 2, we find all stochastic root matrices of 2×2 stochastic matrices explicitly. It seems that the 2×2 case is the only one that all solutions can be found explicitly. In Section 3, all real roots of 3×3 stochastic matrices that are functions of the original stochastic matrix are found. The stochastic roots that are functions of the original stochastic matrix are among these roots. In Section 4, we comment on the Taylor expansion method, a nonlinear programming method, and a Newton's method for computing stochastic roots.

2. Explicit Solutions for the 2×2 Case

In this section, we find explicit solutions to Problem (1.1) when m=2. The idea is to explore a relationship between the diagonal elements of A and the existence of a solution to Problem (1.1). Denote by Tr(A) = $a_{1,1}$ + $a_{2,2}$ the trace of A and det(A) the determinant of A (Lancaster and Tismenetsky [12]). Numerical examples show that, in order to have a stochastic root B, we must have either Tr(A) ≥ 1 or, if Tr(A) < 1, $a_{1,1}$ and $a_{2,2}$ are close to each other. These observations lead to the following main theorem for m=2.

Theorem 2.1. *Assume that n is a positive integer and A is a* 2×2 *stochastic matrix. If n is odd, a stochastic root matrix B satisfying* $A = B^n$ *exists if and only if*

$$\mathrm{Tr}(A)+\min\left\{\frac{(1-a_{1,1})^n}{(1-a_{2,2})^n},\ \frac{(1-a_{2,2})^n}{(1-a_{1,1})^n}\right\}\geq 1. \quad (2.1)$$

If it exists, the stochastic solution $B = (b_{i,j})$ *is unique and is given as follows. If* $a_{11} + a_{22} = 2$, *then* $B = I$, *where I is the identity matrix. If* $a_{11} + a_{22} < 2$, *B is given by:* $b_{1,2}=1-b_{1,1}$, $b_{2,1}=1-b_{2,2}$,

$$b_{1,1} = \frac{1-a_{2,2}+(1-a_{1,1})\left(a_{1,1}+a_{2,2}-1\right)^{1/n}}{2-a_{1,1}-a_{2,2}}, \quad (2.2)$$

$$b_{2,2} = \frac{1-a_{1,1}+(1-a_{2,2})\left(a_{1,1}+a_{2,2}-1\right)^{1/n}}{2-a_{1,1}-a_{2,2}}.$$

If n is even, a stochastic root matrix B satisfying $A = B^n$ *exists if and only if Tr(A)* ≥ 1, *or equivalently, det(A)* ≥ 0. *For this case, there are at most two stochastic roots. If* $a_{1,1} + a_{2,2}$ = 2, *there are two stochastic roots:*

$$B = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (2.3)$$

If $a_{1,1} + a_{2,2} < 2$ *and the following condition is satisfied*

$$\min\left\{\frac{(1-a_{1,1})^n}{(1-a_{2,2})^n},\ \frac{(1-a_{2,2})^n}{(1-a_{1,1})^n}\right\} \geq \mathrm{Tr}(A)-1\geq 0, \quad (2.4)$$

there are exactly two stochastic roots B: one is given by equation (2.2) *and the other is given by:* $b_{1,2}=1-b_{1,1}$, $b_{2,1}=1-b_{2,2}$,

$$b_{1,1} = \frac{1-a_{2,2}-(1-a_{1,1})(a_{1,1}+a_{2,2}-1)^{1/n}}{2-a_{1,1}-a_{2,2}}, \quad (2.5)$$

$$b_{2,2} = \frac{1-a_{1,1}-(1-a_{2,2})(a_{1,1}+a_{2,2}-1)^{1/n}}{2-a_{1,1}-a_{2,2}}.$$

If condition (2.4) *is not satisfied, there is exactly one stochastic root B, which is given by equation* (2.2).

Proof. The proof of Theorem 2.1 is based on the Jordan canonical form of matrix. Details are given in Appendix A.

Note 2.1: According to Kingman [11], if m=2, the equation $A = \exp\{T\}$ has an infinitesimal generator solution T if and only if $\mathrm{Tr}(A) \geq 1$. According to Theorem 2.1, $\mathrm{Tr}(A) \geq 1$ implies that a stochastic solution to Problem (1.1) exists for all positive integer n, which is consistent with the result in Kingman [11]. However, the condition $\mathrm{Tr}(A) \geq 1$ is not necessary for the existence of a stochastic solution to Problem (1.1) if n is odd. For instance, let

$$A = \begin{pmatrix} 0.3 & 0.7 \\ 0.6 & 0.4 \end{pmatrix}, \quad B = \begin{pmatrix} 0.038306 & 0.961694 \\ 0.824309 & 0.175691 \end{pmatrix}. \quad (2.6)$$

Then the matrix A and B satisfy the equation $A = B^5$. Since $\mathrm{Tr}(A) = 0.7 < 1$, by Theorem 2.1, there is no solution to $A = B^n$ if n is even. Therefore, there is no infinitesimal generator T satisfying $A = \exp\{T\}$.

Theorem 2.1 implies that there are only a finite number of stochastic roots for any stochastic matrix if m=2. That is not true if m≥3 (see the example in equation (3.2)). Theorem 2.1 shows that all stochastic roots are functions of A, except when $A = I$. That is no longer true for m≥3 (equation (3.2)). In fact, the solutions to Problem (1.1) with m≥3 are much more complicated than that of the m=2 case. Thus, we do not intend to find all stochastic roots if m≥3. Instead, we shall focus on stochastic roots that are functions of A. Further, the method used for solving the m=2 case cannot be applied to more general case directly. Therefore, another method based on the characteristic polynomial of matrix is developed in Section 3 to solve the m=3 case.

3. Real Roots for the 3×3 Case

A matrix B is a *function* of the matrix A if $B = \Sigma_i d_i A^i$ for some complex numbers $\{d_0, d_1, \ldots\}$ (Lancaster and Tismenetsky [12]). If m=3, as was pointed out in Section 2, there can be infinite number of solutions to Problem (1.1) and many of the roots are not functions of A. For instance, the identity matrix I has stochastic roots such as I itself,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}. \quad (3.1)$$

Apparently, the last two stochastic roots are not functions of I. Another example is

$$A = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}, \quad (3.2)$$
$$B(\varepsilon) = A + \varepsilon \begin{pmatrix} 0 & 0 & 0 \\ -2 & 1 & 1 \\ 2 & -1 & -1 \end{pmatrix}.$$

It is easy to verify that $A = (B(\varepsilon))^n$ for any n≥2 and $B(\varepsilon)$ is stochastic for all small enough ε. However, the matrix $B(\varepsilon)$ is not a function of A (non function root) It was shown in Higham [9] that non function roots exists only if the algebraic multiplicities of some eigenvalues of A are more than one. For lower order matrix A, it is possible to find when A has eigenvalues with algebraic multiplicities that are larger than one. According to Theorem 4 in Higham [9], the number of real roots that are functions of A is finite. Consequently, it is possible to find all real roots that are functions of A. Thus, we relax the constraints in Problem (1.1) and try to find all real roots with row sums being one, i.e.,

$$A = B^n, \quad B\mathbf{e} = \mathbf{e}, \tag{3.3}$$

and B is a function of A. It is easy to see that the solutions to Problem (1.1) that are functions of A are included in the solutions to Problem (3.3).

The method to solve Problem (3.3) is based on the characteristic polynomial of the solution B. Let $\{\mu_1, \mu_2, \mu_3\}$ be the eigenvalues of the matrix B with $\mu_1 = 1$. Denote by $f_B(x) = \det(xI-B)$, the characteristic polynomial of B. It is well known that

$$f_B(x) = \prod_{i=1}^{3}(x-\mu_i) = \sum_{i=0}^{3} b_i x^i, \tag{3.4}$$

where

$$b_0 = -\mu_2\mu_3, \quad b_1 = \mu_2\mu_3 + \mu_2 + \mu_3, \quad b_2 = -(1+\mu_2+\mu_3), \quad b_3 = 1. \tag{3.5}$$

Using the well known fact $f_B(B) = 0$, we express A and A^2 as functions of B.

Lemma 3.1 If B is a solution to Problem (3.3), we have, for n=2,

$$AB + b_2 A + b_1 B + b_0 I = 0, \quad A^2 + b_2 AB + b_1 A + b_0 B = 0; \tag{3.6}$$

for $n \ge 3$,

$$A = B^n = c_{n,0} I + c_{n,1} B + c_{n,2} B^2; \quad A^2 = B^{2n} = c_{2n,0} I + c_{2n,1} B + c_{2n,2} B^2, \tag{3.7}$$

where

$$c_{3,0} = -b_0, \quad c_{3,1} = -b_1, \quad c_{3,2} = -b_2; \quad c_{k+1,0} = -b_0 c_{k,2}, \quad k \ge 3; \quad c_{k+1,i} = -b_i c_{k,2} + c_{k,i-1}, \quad 1 \le i \le 2, \ k \ge 3. \tag{3.8}$$

Furthermore, we have $c_{k,0}+c_{k,1}+c_{k,2} = 1$ for $k\ge 3$.

Proof. The proof is based on $f_B(B) = 0$, i.e., $B^3 = -(b_0I+ b_1B+ b_2B^2)$. Equations (3.6) to (3.8) can be obtained by routine calculations. This completes the proof of Lemma 3.1.

By equations (3.6) and (3.7), we can express B with A and A^2 as follows:

$$B = \frac{-b_0 b_2 I + (b_1 - b_2^2)A + A^2}{b_1 b_2 - b_0}, \quad \text{if } n = 2, \ b_1 b_2 - b_0 \ne 0;$$
$$B = \frac{-(c_{n,0}c_{2n,2} - c_{2n,0}c_{n,2})I + c_{2n,2}A - c_{n,2}A^2}{c_{n,1}c_{2n,2} - c_{2n,1}c_{n,2}}, \quad \text{if } n \ge 3, \ c_{n,1}c_{2n,2} - c_{2n,1}c_{n,2} \ne 0. \tag{3.9}$$

It is easy to verify that the sum of the coefficients of I, A, and A^2 in equation (3.9) is one. Suppose that the *minimum polynomial* of B has a degree of 2 (denote it again by $f_B(x)$), i.e., $f_B(x) = (x-1)(x-\mu_2) = x^2-(1+\mu_2)x+\mu_2$, where μ_2 is the eigenvalue that is different from 1. Similar to equation (3.9), the matrix B can be expressed in terms of A as:

$$B = \frac{-c_{n,0}I + A}{c_{n,1}}, \quad \text{if } n \ge 2, \ c_{n,1} \ne 0, \tag{3.10}$$

where, for $k \ge 2$,

$$c_{2,0} = -\mu_2, \ c_{2,1} = 1 + \mu_2; \quad c_{k+1,0} = -\mu_2 c_{k,1}, \quad c_{k+1,1} = (1+\mu_2)c_{k,1} + c_{k,0}. \tag{3.11}$$

It can be verified that the sum of the coefficients of I and A in equation (3.10) is one. If the minimum polynomial of B has a degree of 1, we must have $B = I = A$.

Next, we find the eigenvalues of B and identify conditions for equations (3.9) and (3.10). Let $\{\lambda_1, \lambda_2, \lambda_3\}$ be the eigenvalues of the matrix A. Since A is stochastic, we set $\lambda_1 = 1$. The following equations can be established for λ_2 and λ_3:

$$\begin{aligned} \mathrm{Tr}(A) &= 1 + \lambda_2 + \lambda_3, \\ \det(A) &= \lambda_2 \lambda_3. \end{aligned} \tag{3.12}$$

By equation (3.12), it is easy to obtain

$$\lambda_{2,3} = \frac{\mathrm{Tr}(A) - 1 \pm \sqrt{(\mathrm{Tr}(A)-1)^2 - 4\det(A)}}{2}. \tag{3.13}$$

If $(\mathrm{Tr}(A)–1)^2 < 4\det(A)$, denote by $\lambda_{2,3} = \alpha \pm \beta = re^{\pm i\theta}$, where

$$\alpha = \frac{\mathrm{Tr}(A)-1}{2}, \quad \beta = \frac{\sqrt{4\det(A) - (\mathrm{Tr}(A)-1)^2}}{2}; \tag{3.14}$$
$$r = \sqrt{\alpha^2 + \beta^2}, \quad \theta = \cos^{-1}(\alpha / r).$$

The solutions to Problem (3.3) have much to do with the quantity $(\mathrm{Tr}(A)–1)^2–4\det(A)$. According to the sign of $(\mathrm{Tr}(A)–1)^2–4\det(A)$, we give all the solutions to Problem (3.3) that are functions of A in the following theorem.

Theorem 3.2 Assume that n is a positive integer and A is a 3×3 stochastic matrix. All solutions to Problem (1.1) that are functions of A are solutions to Problem (3.3). All solutions to Problem (3.3) that are functions of A can be computed as follows.

If $(\mathrm{Tr}(A)–1)^2 < 4\det(A)$, there are total n solutions to Problem (3.3). Each solution corresponds to a set of eigenvalues: for $0 \le k \le n-1$,

$$\begin{aligned} &\mu_1 = 1, \quad \mu_2 = r^{1/n} \exp\{i(\theta + 2k\pi)/n\}, \\ &\mu_3 = \overline{\mu}_2. \end{aligned} \tag{3.15}$$

The corresponding solution can be calculated by equation (3.9).

If $(\mathrm{Tr}(A)–1)^2 > 4\det(A)$ and n is odd, the solution to Problem (3.3) is unique. The eigenvalues of the solution are $\mu_1 = 1$, $\mu_2 = (\lambda_2)^{1/n}$, and $\mu_3 = (\lambda_3)^{1/n}$. If $(\mathrm{Tr}(A)–1)^2 > 4\det(A)$, n is even, and $\det(A) < 0$, there is no solution to Problem (3.1). If $(\mathrm{Tr}(A)–1)^2 > 4\det(A)$, n is even, $\det(A) \ge 0$, and $\mathrm{Tr}(A) < 1$, there is no solution to Problem (3.3). If $(\mathrm{Tr}(A)–1)^2 > 4\det(A)$, n is even, $\det(A) \ge 0$, and $\mathrm{Tr}(A) \ge 1$, there are four solutions to Problem (3.3). The four sets of eigenvalues are $\{1, \mu_2 = (\lambda_2)^{1/n}, \mu_3 = (\lambda_3)^{1/n}\}$, $\{1, \mu_2 = –(\lambda_2)^{1/n}, \mu_3 = (\lambda_3)^{1/n}\}$, $\{1, \mu_2 = (\lambda_2)^{1/n}, \mu_3 = –(\lambda_3)^{1/n}\}$, and $\{1, \mu_2 = –(\lambda_2)^{1/n}, \mu_3 = –(\lambda_3)^{1/n}\}$. For all the cases having solutions, the corresponding solutions can be calculated by equation (3.9).

If $(\mathrm{Tr}(A)–1)^2 > 4\det(A)$ and $\lambda_2 = 1$, then A is reducible. Problem (3.3) is reduced to solve a Problem (3.3) for a 2×2 substochastic matrix.

Assume $(\mathrm{Tr}(A)–1)^2 = 4\det(A)$. If $\lambda_1 = \lambda_2 = \lambda_3 = 1$, the only solution to Problem (3.3) that is a function of A is $A=I$ itself. There can be other solutions that are not functions of A (see equation (3.1) for an example). If $1 = \lambda_1 > \lambda_2 = \lambda_3 = \alpha$, $A^2–(1+\alpha)A + \alpha I \neq 0$, n is odd, there is a unique solution B that can be calculated by either equation (3.9) or equation (3.10), whichever is valid. If $1 = \lambda_1 > \lambda_2 = \lambda_3 = \alpha > 0$, $A^2–(1+\alpha)A + \alpha I \neq 0$, and n is even, there are possibly two solutions to Problem (3.3) with eigenvalues $\{\mu_1=1, \mu_2 = \mu_3 = |\alpha|^{1/n}\}$ and $\{\mu_1=1, \mu_2 = \mu_3 = –|\alpha|^{1/n}\}$, which can be computed by equation (3.9). If $1 = \lambda_1 > \lambda_2 = \lambda_3 = \alpha$, $A^2–$

$(1+\alpha)A + \alpha I \neq 0$, n is even and $\alpha<0$, there is no solution to Problem (3.3). If $1 = \lambda_1 > \lambda_2 = \lambda_3 = \alpha$, $A^2-(1+\alpha)A + \alpha I \neq 0$, $\alpha = 0$, there is no solution to Problem (3.3). If $1 = \lambda_1 > \lambda_2 = \lambda_3 = \alpha$, $A^2-(1+\alpha)A + \alpha I = 0$, and $\alpha\neq 0$, there are possibly n sets of eigenvalues of B: for $0\leq k\leq n-1$,

$$\mu_1 = 1, \quad \mu_2 = |\alpha|^{1/n} \exp\{i2k\pi)/n\}, \quad \mu_3 = \bar{\mu}_2. \tag{3.16}$$

For each set of eigenvalues, the matrix B can be computed by using equation (3.10). If $1 = \lambda_1 > \lambda_2 = \lambda_3 = \alpha$, $A^2-(1+\alpha)A + \alpha I = 0$, and $\alpha=0$, there are infinitely many solutions to Problem (3.3) that are not functions of A (see Equation (3.2)). For this case, $B = A$ is a solution to Problem (3.3) with eigenvalues $\{\mu_1 = 1, \mu_2 = \mu_3 = 0\}$.

Proof. See Appendix B.

Note 3.1 The conditions for the existence of a solution to Problem (3.3) can be useful in practice. For instance, if n is even, we must have $\det(A)\geq 0$ in order to have a stochastic root. This condition is apparently true for any order m. In general, finding conditions for the existence of a solution to Problem (1.1) is an interesting research issue. Unfortunately, results in Sections 2 and 3 show that simple necessary and sufficient conditions may not exist for cases with $m>3$.

The method used in this section can be applied to find all stochastic roots that are functions of A. However, to use this method, we must find either all the eigenvalues of B or the coefficients of the characteristic polynomial (minimal polynomial) of B. For lower order matrix A (m=4, 5, …), it might be possible to do so. Unfortunately, if m is not small ($m>5$), it can be difficult. Furthermore, for $m>5$, cases with multiple eigenvalues are extremely difficult to deal with, since there are many possibilities and these can be dealt with only on a case by case basis. Therefore, there is a need to explore other methods to find a solution to Problem (1.1) if m is not small.

4. Comments on Computational Methods

As was pointed out at the end of Sections 2 and 3, the methods used to solve the lower order cases cannot be applied to higher order case effectively. Therefore, there is a need to look for computational methods that can find one or several solutions to Problem (1.1). Based on our numerical experimentations, we briefly discuss three numerical methods.

Taylor Expansion Algorithm This method is based on the Taylor expansion of the function $z^{1/n}$ for complex number z and positive real number n:

$$z^{1/n} = (1-(1-z))^{1/n} = 1 - \frac{1}{n}\sum_{k=1}^{\infty} b_k (1-z)^k, \tag{4.1}$$

where

$$b_k = \frac{1}{k}\prod_{j=1}^{k-1}\left(1-\frac{1}{jn}\right), \quad k \geq 1. \tag{4.2}$$

Based on equation (4.1), the following algorithm can be used for computing a matrix sequence $\{B[k], k\geq 0\}$ for any matrix A. Let $B[0] = I$ and $b_1 = 1$. For $k\geq 1$, let

$$B[k] = B[k-1] - \frac{1}{n} b_k (I-A)^k; \quad b_{k+1} = b_k\left(\frac{k}{k+1}\right)\left(1-\frac{1}{kn}\right). \tag{4.3}$$

Let $\{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ be all the eigenvalues of A, it is well-known that $\{B[k], k\geq 0\}$ converges if and only if the eigenvalues $\{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ satisfy $|\lambda_i - 1| < 1$ for $1 \le i \le m$. Denote the limit of $\{B[k], k \ge 0\}$ as B. It is easy to verify that $B\mathbf{e} = \mathbf{e}$ and B is an nth root of the matrix A.

The Taylor expansion method can be easily implemented to find a root of a stochastic matrix. However, this method has several limitations. First, the convergence region is too small. Second, the limit B may not be a stochastic matrix, even when the matrix A has an nth stochastic root. Third, the convergence rate of the algorithm is slow and the algorithm is instable.

A Nonlinear Programming Approach The Problem (1.1) can be transformed into nonlinear programming problems. For instance, for a positive integer n and $A \in \mathscr{S}_m$, we define a nonlinear program:

$$\begin{aligned} \varsigma^* &= \min \left\{ \| B^n - A \|_2 \right\} \\ s.t. \quad & B \in \mathscr{S}_m, \end{aligned} \tag{4.4}$$

where $\|\cdot\|_2$ defines an l_2 norm on the set of all real $m \times m$ matrices (see Bhatia [3]). Apparently, we have $0 \le \varsigma^* \le m^2$. It is readily seen that Problem (1.1) has a stochastic solution if and only if the optimal solution to Problem (4.4) has $\varsigma^* = 0$.

Probem (4.4) is a nonlinear program with a convex feasible region. Using existing methods for solving nonlinear programming problems (Bertsekas [2]), we can solve Problem (4.4) for any n and $A \in \mathscr{S}_m$. For instance, the well-known *feasible decent direction method* can be used to solve Problem (4.4). However, solving Problem (4.4) can be complicated and time consuming, especially when n or m is large. In fact, it is quite often that the solutions given by using existing methods are not even close to a real solution of Problem (1.1). The failure can be, at least partially, explained as follows. According to Perron-Frobenius theory (Minc [13]), an irreducible stochastic matrix B can be approximated by $\mathbf{eu} + O(\rho)$, where the row vector $\mathbf{u}$ is a left eigenvector of B corresponding to eigenvalue 1, and ρ is the absolute value of the eigenvalue of B with the second largest modulus. Then B^n can be approximated by $\mathbf{eu} + O(\rho^n)$. Suppose that stochastic matrices B_1 and B_2 have approximations $\mathbf{eu} + O(\rho_1)$ and $\mathbf{eu} + O(\rho_2)$, respectively. If $\rho_1 < 1$ and $\rho_2 < 1$, then the difference between B_1^n and B_2^n is asymptotically $O(\rho_1^n - \rho_2^n)$, which can be significantly smaller than the difference between B_1 and B_2: $O(\rho_1 - \rho_2)$. Thus, it is possible that both B_1^n and B_2^n are close to A, but only one of them is the solution. Thus, unless an algorithm explicitly takes into account *some* special properties of Problem (1.1), it may perform poorly.

Since for any positive integer n, A^n can be expressed in terms of $\{I, A, A^2, \ldots, A^{m-1}\}$, any stochastic root B as a function of A can be expressed as $B = h_0 I + h_1 A + h_2 A^2 + \ldots + h_{m-1} A^{m-1}$. Then we can restrict our search to the set of stochastic matrices that are function of A so as to reduce the number of variables involved in the nonlinear program (4.4). For instance, Problem (4.4) can be changed to:

$$\begin{aligned} \varsigma^* &= \min \left\{ \left\| \left(\sum_{i=0}^{m-1} h_i A^i \right)^n - A \right\|_2 \right\} \\ s.t. \quad & \sum_{i=0}^{m-1} h_i A^i \ge 0, \quad \sum_{i=0}^{m-1} h_i = 1. \end{aligned} \tag{4.5}$$

The advantage of the nonlinear program (4.5) is that the number of variables is reduced to m, instead of m^2 for (4.4). The feasible region of (4.5) is still convex. The disadvantage of (4.5) is that it may not be able to find a solution to Problem (1.1) if all solutions to Problem (1.1) are not functions of *A*. The performance of this method depends heavily on the algorithm employed to solve the nonlinear programs.

Newton's method Newton's method is widely used to develop algorithms for computing matrix roots (Higham [8] and references therein). For instance, the following algorithm can be used for computing a solution to Problem (1.1). Let $B[0] = 0$ and, for $k \geq 1$,

$$B[k] = B[k-1] + \frac{1}{n}\left(A - \left(B[k-1]\right)^n\right). \quad (4.6)$$

If the sequence $\{B[k], k \geq 0\}$ converges to a matrix *B*, by equation (4.6), we must have $A = B^n$. This approach is easy to be implemented and is quite efficient if the sequence does converge. On the other hand, the limit may not be a stochastic matrix. Furthermore, the convergence region of this method can be small and it depends on *m*, *n*, as well as *A*. Thus, it is difficult to identify the convergence region for this algorithm.

Appendix A. Proof of Theorem 2.1

Since $A \in \mathcal{S}_2$, its Perron-Frobenius eigenvalue (the eigenvalue with the largest modulus) is one. Let ξ be the other eigenvalue of *A*. Then ξ must be a real number and $|\xi| \leq 1$. The Jordan canonical form of *A* can be given as

$$A = Q\begin{pmatrix} 1 & 0 \\ 0 & \xi \end{pmatrix}Q^{-1} \quad \text{or} \quad A = Q\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}Q^{-1}$$

where $Q = (q_{i,j})$ is a non-singular real matrix. Since the trace of a matrix is invariant for similar matrices, we must have $\mathrm{Tr}(A) = a_{1,1} + a_{2,2} = 1 + \xi$. For the second case in equation (A.1), $\xi = 1$. Then we must have $a_{1,1} = a_{2,2} = 1$. Thus, $A = I$ since $A \in \mathcal{S}_2$. Then *A* is not similar to the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Therefore, *A* can only have the first Jordan canonical form given in equation (A.1).

Suppose that *B* satisfying $A = B^n$ is stochastic. Then *B* must have two real eigenvalues $\{1, \mu\}$ with $\xi = \mu^n$. The matrix *B* must have a Jordan canonical form

$$B = P\begin{pmatrix} 1 & 0 \\ 0 & \mu \end{pmatrix}P^{-1},$$

where *P* is a non-singular real matrix. Since $A = B^n$, we have $P\,\mathrm{diag}(1,\xi)P^{-1} = Q\,\mathrm{diag}(1,\xi)Q^{-1}$, which leads to $\mathrm{diag}(1,\xi)P^{-1}Q = P^{-1}Q\,\mathrm{diag}(1,\xi)$. Let $L = P^{-1}Q = (l_{i,j})$. Then we must have $l_{1,2} = l_{1,2}\xi$ and $l_{2,1} = l_{2,1}\xi$. If $\xi \neq 1$, $l_{1,2} = l_{2,1} = 0$, which implies that $l_{1,2} = l_{1,2}\mu$ and $l_{2,1} = l_{2,1}\mu$. Then we have

$$B = P\begin{pmatrix} 1 & 0 \\ 0 & \mu \end{pmatrix}P^{-1} = Q\begin{pmatrix} 1 & 0 \\ 0 & \mu \end{pmatrix}Q^{-1}. \quad \text{(A.3)}$$

If $\xi = 1$, $A = I$. If *n* is odd and $\xi = 1$, $\mu = \xi^{1/n} = 1$. Then equation (A.3) holds. If *n* is even and $\xi = 1$, $\xi^{1/n} = \pm 1$. We first assume that $\mu = 1$. Then equation (A.3) holds. The case with $\mu = -1$ with be dealt with later.

Assume (A.3) holds. Since $a_{1,1} + a_{2,2} - 1 = \xi$, we have $\mu = \pm(a_{1,1} + a_{2,2} - 1)^{1/n}$. (Note: if *n* is odd, there is only one root $\mu = (a_{1,1} + a_{2,2} - 1)^{1/n}$.)

Since *Q* is non-singular, i.e., $\det(Q) = q_{1,1}q_{2,2} - q_{1,2}q_{2,1} \neq 0$, it is easy to obtain

$$Q^{-1} = \begin{pmatrix} q_{2,2} & -q_{1,2} \\ -q_{2,1} & q_{1,1} \end{pmatrix}\frac{1}{(q_{1,1}q_{2,2} - q_{1,2}q_{2,1})}. \quad \text{(A.4)}$$

Expanding the first equation in (A.1), we obtain,

$$a_{1,1}=\frac{q_{1,1}q_{2,2}-\xi q_{1,2}q_{2,1}}{q_{1,1}q_{2,2}-q_{1,2}q_{2,1}};$$

$$a_{1,2}=\frac{-q_{1,1}q_{1,2}+\xi q_{1,2}q_{1,1}}{q_{1,1}q_{2,2}-q_{1,2}q_{2,1}}=\frac{(\xi-1)q_{1,2}q_{1,1}}{(q_{1,1}q_{2,2}-q_{1,2}q_{2,1})}; \quad \text{(A.5)}$$

$$a_{2,1}=\frac{q_{2,1}q_{2,2}-\xi q_{2,1}q_{2,2}}{q_{1,1}q_{2,2}-q_{1,2}q_{2,1}}=\frac{(1-\xi)q_{2,1}q_{2,2}}{(q_{1,1}q_{2,2}-q_{1,2}q_{2,1})};$$

$$a_{2,2}=\frac{-q_{2,1}q_{1,2}+\xi q_{1,1}q_{2,2}}{q_{1,1}q_{2,2}-q_{1,2}q_{2,1}}.$$

It is easy to verify that $a_{1,1}-a_{2,2}=(1-\xi)(q_{1,1}q_{2,2}+q_{1,2}q_{2,1})/(q_{1,1}q_{2,2}-q_{1,2}q_{2,1})$. Similarly, we can expand equation (A.3) to obtain $b_{1,1}+b_{2,2}=1+\mu$ and $b_{1,1}-b_{2,2}=(1-\mu)(q_{1,1}q_{2,2}+q_{1,2}q_{2,1})/(q_{1,1}q_{2,2}-q_{1,2}q_{2,1})$. Combining these results yields, if $\xi<1$,

$$b_{1,1}+b_{2,2}=1+\mu=1\pm\left(a_{1,1}+a_{2,2}-1\right)^{1/n};$$

$$b_{1,1}-b_{2,2}=(a_{1,1}-a_{2,2})\frac{(1-\mu)}{1-\xi} \quad \text{(A.6)}$$

$$=(a_{1,1}-a_{2,2})\frac{\left(1\mp\left(a_{1,1}+a_{2,2}-1\right)^{1/n}\right)}{2-a_{1,1}-a_{2,2}},$$

if $\xi=1$, $A=I$ and $B=I$ (recall the assumption $\mu=1$). Expressions in equation (A.6) lead to the following expressions: (if n is odd, the sign "±" is read as "+".)

$$b_{1,1}=\frac{1-a_{2,2}}{2-a_{1,1}-a_{2,2}}\pm\frac{(1-a_{1,1})}{2-a_{1,1}-a_{2,2}}\left(a_{1,1}+a_{2,2}-1\right)^{1/n}; \quad \text{(A.7)}$$

$$b_{2,2}=\frac{1-a_{1,1}}{2-a_{1,1}-a_{2,2}}\pm\frac{(1-a_{2,2})}{2-a_{1,1}-a_{2,2}}\left(a_{1,1}+a_{2,2}-1\right)^{1/n}.$$

Now, we consider the case with $\mu=-1$. If n is even, $\xi=1$ and $\mu=-1$, then $B=P\,\mathrm{diag}(1,-1)P^{-1}$, which implies that $\mathrm{Tr}(B)=0$. Since B is stochastic, we must have $B=J=\begin{pmatrix}0&1\\1&0\end{pmatrix}$ and $A=I$.

To prove Theorem 2.1, we consider the following three cases: Case #1: $a_{1,1}=a_{2,2}=1$; Case #2: n is even; and Case #3: n is odd.

Case #1. If $a_{1,1}=a_{2,2}=1$, the matrix $A=I$ is the identity matrix. Then $B=I$ is the unique solution if n is odd. If n is even, there are exactly two stochastic roots: I and J. Conditions in a) and b) in Theorem 2.1 are satisfied for all positive n since $\mathrm{Tr}(A)=2$.

Case #2. Suppose that n is even. If $A=B^n$ has a stochastic solution, by equation (A.7), we must have $\mathrm{Tr}(A)=a_{1,1}+a_{2,2}\geq 1$. On the other hand, if $\mathrm{Tr}(A)=a_{1,1}+a_{2,2}\geq 1$, it is easy to see that the two solutions given by equation (A.7) satisfying $A=B^n$ (with $b_{1,2}=1-b_{1,1}$ and $b_{2,1}=1-b_{2,2}$). Now, we need to verify that B given by (A.7) is a stochastic matrix. To do so, we only need to prove that $0\leq b_{1,1}\leq 1$ and $0\leq b_{2,2}\leq 1$.

First, we read "±" as "+" in (A.7). That $b_{1,1}$ and $b_{2,2}$ are nonnegative is obtained immediately from $a_{1,1}+a_{2,2}\geq 1$. Furthermore, that $b_{1,1}\leq 1$ (and $b_{2,2}\leq 1$) is proved as follows:

$$\frac{1-a_{2,2}+(1-a_{1,1})(a_{1,1}+a_{2,2}-1)^{1/n}}{2-a_{1,1}-a_{2,2}}\leq 1 \quad \text{(A.8)}$$

$$\Leftrightarrow 1-a_{2,2}+(1-a_{1,1})\left(a_{1,1}+a_{2,2}-1\right)^{1/n}\leq 2-a_{1,1}-a_{2,2}$$

$$\Leftrightarrow (1-a_{1,1})\left(a_{1,1}+a_{2,2}-1\right)^{1/n}\leq 1-a_{1,1}.$$

The last inequality in (A.8) holds since $1\leq a_{1,1}+a_{2,2}\leq 2$. Therefore, this solution is always stochastic. By equation (A.7), it is easy to see that Tr(B) ≥ 1 (because we choose the nonnegative nth root of $a_{1,1}+a_{2,2}-1$).

Second, we read "±" as "–" in (A.7). By (A.7), it is easy to verify that $b_{1,1} \le 1$ and $b_{2,2} \le 1$. To find when $b_{1,1}\ge 0$ and $b_{2,2}\ge 0$, we have the following calculations:

$$
\begin{aligned}
&\frac{1-a_{2,2}-(1-a_{1,1})\left(a_{1,1}+a_{2,2}-1\right)^{1/n}}{2-a_{1,1}-a_{2,2}} \ge 0 \\
\Leftrightarrow\ & 1-a_{2,2}-(1-a_{1,1})\left(a_{1,1}+a_{2,2}-1\right)^{1/n} \ge 0 \qquad \text{(A.9)} \\
\Leftrightarrow\ & (1-a_{2,2})/(1-a_{1,1}) \ge \left(a_{1,1}+a_{2,2}-1\right)^{1/n},
\end{aligned}
$$

which leads to the condition given in equation (2.4). For this solution, it can be verified that Tr(B) ≤ 1.

Case #3. Suppose that n is an odd number (>2). If a solution to $A = B^n$ exists, then we must have $b_{1,1}\ge 0$ and $b_{2,2}\ge 0$. We read "±" in (A.7) as "+" only. By the first expression in equation (A.7), we must have

$$
\begin{aligned}
&\frac{1-a_{2,2}+(1-a_{1,1})\left(a_{1,1}+a_{2,2}-1\right)^{1/n}}{2-a_{1,1}-a_{2,2}} \ge 0 \qquad \text{(A.10)} \\
\Leftrightarrow\ & 1-a_{2,2}-(1-a_{1,1})\left(1-a_{1,1}-a_{2,2}\right)^{1/n} \ge 0 \\
\Leftrightarrow\ & \left(1-a_{1,1}-a_{2,2}\right)^{1/n} \le (1-a_{2,2})/(1-a_{1,1}) \\
\Leftrightarrow\ & 1-a_{1,1}-a_{2,2} \le \left((1-a_{2,2})/(1-a_{1,1})\right)^{n},
\end{aligned}
$$

which leads to the conditions in (2.1). Apparently, the solution is unique.

On the other hand, suppose that conditions in equation (2.1) are satisfied. Then it can be verified that (A.7) gives a stochastic solution to equation $A = B^n$. This completes the proof of Theorem 2.1.

Appendix B. Proof of Theorem 3.2

In the proof, we consider three cases separately: $(\mathrm{Tr}(A)-1)^2 < 4\det(A)$, $(\mathrm{Tr}(A)-1)^2 > 4\det(A)$, and $(\mathrm{Tr}(A)-1)^2 = 4\det(A)$.

First, we assume $(\mathrm{Tr}(A)-1)^2 < 4\det(A)$. Then A has 3 different eigenvalues: $\lambda_1 = 1$, $\lambda_2 = re^{i\theta}$, and $\lambda_3 = re^{-i\theta}$. The B has three different eigenvalues given in equation (3.15) for some k ($0\le k\le n-1$). By equation (3.5), the coefficient of the characteristic polynomial of B can be calculated. By equation (3.9), the matrix B can be calculated if the denominator in equation (3.9) is not zero. Since all eigenvalues of A are different, the denominator in equation (3.9) cannot be zero. Otherwise, equations (3.7) and (3.9) imply that the minimum polynomial of A has a degree 2 or less, which is impossible since all three eigenvalues of A are different. Therefore, for each k ($0\le k\le n-1$), a solution B to Problem (3.3) can be computed by equation (3.9). The matrix B is real and $B\mathbf{e}=\mathbf{e}$ since the sum of the coefficient of I, A, and A^2 in equation (3.9) is one.

Second, we assume $(\mathrm{Tr}(A)-1)^2 > 4\det(A)$. This case is more complex than the first one. For this case, the matrix A has three real eigenvalues: $\lambda_1=1 \ge \lambda_2 > \lambda_3$. If $\lambda_2 =1$, the matrix A must be reducible. Otherwise, if A is irreducible, the eigenvalue one has an algebraic multiplicity one (Minc [13]). If A is reducible, then

$$
A = \begin{pmatrix} 1 & 0 & 0 \\ * & * & * \\ * & * & * \end{pmatrix}. \qquad \text{(B.2)}
$$

The Problem (3.3) is reduced to a root problem with m=2, which can be solved by applying results from Section 2 (where m=2). We omit the details. Next, we assume $\lambda_1=1 > \lambda_2 > \lambda_3$. It is easy to see that the eigenvalues of B must be different real numbers. There are following four cases to be dealt with.

1) If $(\mathrm{Tr}(A)–1)^2 > 4\det(A)$ and n is odd, the matrix B has three different eigenvalues given by $\mu_1 = 1$, $\mu_2 = (\lambda_2)^{1/n}$, and $\mu_3 = (\lambda_3)^{1/n}$. The matrix B can be computed by equation (3.9), which is the unique solution to Problem (3.3) (Note that A has three different eigenvalues).

2) If $(\mathrm{Tr}(A)–1)^2 > 4\det(A)$, n is even, and $\det(A) < 0$, there is no solution to Problem (3.3). Otherwise, $\det(A) = (\det(B))^n \geq 0$, a contradiction.

3) If $(\mathrm{Tr}(A)–1)^2 > 4\det(A)$, n is even, $\det(A) \geq 0$, and $\mathrm{Tr}(A) < 1$, we must have $\lambda_3 < 0$. Then μ_3 is a complex number since n is even, which contradicts to the fact that μ_3 is real. Therefore, there is no solution to Problem (3.3).

4) If $(\mathrm{Tr}(A)–1)^2 > 4\det(A)$, n is even, $\det(A) \geq 0$, and $\mathrm{Tr}(A) \geq 1$, then $\lambda_1=1 > \lambda_2 > \lambda_3 > 0$. There are four sets of eigenvalues of B: $\{1, \mu_2 = (\lambda_2)^{1/n}, \mu_3 = (\lambda_3)^{1/n}\}$, $\{1, \mu_2 = –(\lambda_2)^{1/n}, \mu_3 = (\lambda_3)^{1/n}\}$, $\{1, \mu_2 = (\lambda_2)^{1/n}, \mu_3 = –(\lambda_3)^{1/n}\}$, and $\{1, \mu_2 = –(\lambda_2)^{1/n}, \mu_3 = –(\lambda_3)^{1/n}\}$. Each set of eigenvalues corresponds to a solution to Problem (3.3). All the solutions can be computed by equation (3.9).

Finally, we assume $(\mathrm{Tr}(A)–1)^2 = 4\det(A)$. For this case, $1 = \lambda_1 \geq \lambda_2 = \lambda_3 = \alpha$. If $\lambda_2 = 1$, we must have $\mathrm{Tr}(A) = 3$, which implies $A = I$. The only solution to Problem (3.3) that is a function of A is $A=I$ itself. But there can be other solutions that are not functions of A (see equation (3.1) for an example). If $1 = \lambda_1 > \lambda_2 = \lambda_3 = \alpha$, we consider two cases based on the form of the Jordan block corresponding to the eigenvalue α. Suppose that the matrix A has a Jordan canonical form

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{pmatrix}. \tag{B.3}$$

If $\alpha \neq 0$, then B must have a Jordan canonical form

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \mu_2 & 1 \\ 0 & 0 & \mu_2 \end{pmatrix}, \tag{B.4}$$

where $\mu_2 = \mu_3$ is a real number. Thus, if n is odd, there is only one solution to Problem (3.3) with $\{\mu_1 = 1, \mu_2 = \mu_3 = \alpha^{1/n}\}$. Since the minimal polynomial of A has a degree of 3, $c_{n,1}c_{2n,2} – c_{n,2}c_{2n,1} \neq 0$ for $n>2$ or $b_1b_2 – b_0 \neq 0$ for $n=2$. Then the matrix B can be calculated by equation (3.9). If n is even and $\alpha>0$, there are possibly two solutions to Problem (3.3) with eigenvalues $\{\mu_1=1, \mu_2 = \mu_3 = |\alpha|^{1/n}\}$ and $\{\mu_1=1, \mu_2 = \mu_3 = –|\alpha|^{1/n}\}$. Again, since the minimal polynomial of A has a degree of 3, we can use equation (3.9) to compute B. If n is even and $\alpha<0$, there is no solution to Problem (3.3).

If equation (B.3) holds with $\alpha = 0$, there is no solution to Problem (3.3) that is function of A since

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{B.5}$$

Suppose that the matrix A has a Jordan canonical form

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix}. \tag{B.6}$$

If $\alpha\neq0$, there are possibly n sets of eigenvalues of B given in (3.16). For each set of eigenvalues, the minimum polynomial of A has a degree of 2. Thus, the matrix B can be computed by using equation (3.10).

If $\alpha=0$, there are infinitely many solutions to Problem (3.3) that are not functions of A. Equation (3.2) presents some examples of such solutions. For this case, $B = A$ is a solution to Problem (3.3) with eigenvalues $\{\mu_1 = 1, \mu_2 = \mu_3 = 0\}$, which can be verified by using equation (3.11). This completes the proof of Theorem 3.2.

Acknowledgements: The authors would like to thank Dr. D. Liu, Dr. M. Neuts, and Dr. Yanhui Wang for their valuable suggestions and comments on the research. This research was supported financially by operating grants from NSERC.

References

[1] Bapat, R.B. and T.E.S. Raghavan, *Nonnegative Matrices and Applications*, Cambridge University Press, Cambridge, 1996.

[2] Bertsekas,D.P., *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, 1982.

[3] Bhatia, R., *Matrix Analysis*, Springer Verlag, New York, 1997.

[4] Bjorck, A. and Hammarling, S., A Schur method for the square root of a matrix, *Linear Algebra and its Applications*, **Vol 52/53**: 127-140, 1983.

[5] Denman, E.D., Roots of real matrices, *Linear Algebra and its Applications*, **Vol 36**, 133-139, 1981.

[6] Egilsson, T., *Effects of Queueing Congestion on Runway Capacity at Airports*, Master Thesis, Department of Industrial Engineering, Dalhousie University, Halifax, Canada, 2001.

[7] Elfving, G., Zur Theorie der Markoffschen Ketten, *Acta Soc. Sci. Fennicae* n. Ser. A 2, No. 8, 2-17, 1937.

[8] Higham, N.J., Newton's method for the matrix square root, *Mathematics of Computation*, **Vol 46**, No. 174, 537-549, 1986.

[9] Higham, N.J., Computing real square roots of a real matrix, *Linear Algebra and its applications*, **Vol 88/89**, 405-430, 1987.

[10] Higham, N.J., *Accuracy and Stability of Numerical Algorithms*, 2nd Edition, SIAM, Philadelphia, USA, 2002.

[11] Kingman, J.F.C., The imbedding problem for finite Markov chains, *Z. Wahrscheinlichkeitstheorie*, **Vol. 1**, 14-24, 1962.

[12] Lancaster, P. and M. Tismenetsky, *The Theory of Matrices*, Academic, New York, 1985.

[13] Minc, H., *Nonnegative Matrices*, Wiley, New York, 1988.

[14] Neuts, M.F., *Probability*, Allyn and Bacon, Boston, Massachusetts, 1973.

[15] Seneta, E., *Non-negative Matrices: An Introduction to Theory and Applications*, John Wiley & Sons, New York, 1973.