On the second eigenvalues of matrices associated with TCP

Abraham Berman* Thomas Laffey† Arie Leizarowitz ‡ Robert Shorten§

July 28, 2005

Abstract

We consider a convex combination of matrices that arise in the study of communication networks and the corresponding convex combination of Kronecker squares of these matrices. We show that the spectrum of the first convex combination is contained in the spectrum of the second set and that the second largest eigenvalues coincide.

Key Words : *Second eigenvalue of column stochastic matrices; Network congestion control; Communication networks; Kronecker products*

1 Introduction

Let $\alpha_1, ..., \alpha_n$, and $\beta_1, ..., \beta_n$, be positive numbers smaller than 1. In studying non-negative matrix models for TCP one considers the following sets of matrices [SWL05]:

$$A(k) = \begin{bmatrix} \beta_1(k) & 0 & \cdots & 0 \\ 0 & \beta_2(k) & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \beta_n(k) \end{bmatrix} + \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_n \end{bmatrix} [(1-\beta_1(k)), \ldots, (1-\beta_n(k))],$$

*Email: berman@tx.technion.ac.il
†Email: thomas.laffey@ucd.ie
‡Email: la@tx.technion.ac.il
§Corresponding author; Email: robert.shorten@may.ie

where $\beta_i(k)$ is either 1 or β_i and $\sum_{i=1}^{n} \alpha_i = 1$. The non-negative matrices $A_2, .., A_m$ are constructed by taking the matrix A_1,

$$A_1 = \begin{bmatrix} \beta_1 & 0 & \cdots & 0 \\ 0 & \beta_2 & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \beta_n \end{bmatrix} + \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_n \end{bmatrix} \begin{bmatrix} (1-\beta_1), & \ldots, & (1-\beta_n) \end{bmatrix}$$

and setting some, but not all, of the β_i to 1. This gives rise to $m = 2^n - 1$ matrices. We denote the set of these matrices by $\mathcal{A}$, refer to matrices of the form of A_1 as TCP matrices, and say that the other $A_i \in \mathcal{A}$ are generated from A_1. In the context of TCP one also considers the following convex combination of these matrices:

$$M = \sum_{i=1}^{m} \rho_i A_i; \quad (1)$$

$$\hat{M} = \Big(\sum_{i=1}^{m} \rho_i A_i \bigotimes A_i; \Big), \quad (2)$$

where $A_i \in \mathcal{A}$. Under certain statistical assumptions Equation (1) arises when studying the first moment of the stochastic process underlying communication networks employing the TCP algorithm, and Equation(2) arises when studying the second moments of this process. From a practical perspective, one is interested in the Perron eigenvectors of both of these matrices and in their second largest eigenvalues. The Perron eigenvectors of these matrices give the asymptotic values of the first and second moments, and the second largest eigenvalues determine the rate of convergence to these asymptotes. In this paper we show that the second largest eigenvalues of these matrices coincide and provide a necessary condition for a positive column stochastic matrix to be a TCP matrix.

2 Inclusion and equality

We start with the following result.

Theorem 2.1 *Let $B_1, ..., B_m$ be a family of $n \times n$ real matrices of the form:*

$$B_i = D_i + v t_i^T, \quad (3)$$

where v is a common left eigenvector of all the B_i with

$$B_i^T v = \lambda_i v. \quad (4)$$

Then, there exists an orthogonal matrix U such that $U^T B_i U$ are block triangular matrices,

$$U^T B_i U = \begin{bmatrix} \lambda_i & | & 0 & \ldots & 0 \\ - & - & - & - & - \\ & | & & & \\ c_i & | & & S_i & \\ & | & & & \\ & | & & & \end{bmatrix}, \tag{5}$$

where all of the S_i are symmetric.

Proof : Let U be an orthogonal matrix whose first column is $\frac{v}{\|v\|}$. Then, it follows that all the matrices $U^T B_i U$ are block triangular. To show that the matrices S_i are symmetric we observe that $U^T D_i U$ are symmetric, and that all the entries of $U^T v t_i^T U$, except in the first row, are zero. □

Corollary 2.1 *Let $A_1, .., A_m$ be a family of matrices generated by a TCP matrix. Then there exists a non-singular matrix P such that $P^{-1} A_i P$ is of the form (5) with $\lambda_i = 1$ where the matrices S_i are positive definite and $\rho(S_i) \leq 1$.*

Proof : Suppose that A_1 is a TCP matrix. Then, $A_i = D_i + bc_i^T$, $A_i^T e = e$, for all i, where D_i is a diagonal matrix, and b, c_i are strictly positive vectors. To prove the assertion it is enough to show that the matrices A_i are simultaneously similar to $\left\{\tilde{A}_1, ...\tilde{A}_n\right\}$ where $\tilde{A}_i = \tilde{D}_i + \tilde{b}\tilde{c}_i^T$, where $\tilde{D}_i$ is again a diagonal matrix, and $\tilde{b}, \tilde{c}_i$, are vectors. To see that, let $E = diag\left\{\sqrt{b_1}, ..., \sqrt{b_n}\right\}$. Note that E is well defined as the vector b is positive. It is easily seen that the matrices $E^{-1} A_i E$ are of the form in the previous theorem. We can therefore choose $P = EU$. The fact that the S_i have positive real eigenvalues that are not greater than one follows from a slight variation of Theorem 3.2 in [BSL04] (by allowing some of the β_i's to be equal to 1). □

Theorem 2.2 *Consider the matrices M and $\hat{M}$ defined in Equations (1) and (2). Then:*

(i) the eigenvalues of M are eigenvalues of $\hat{M}$;

(ii) all the eigenvalues of M which are different from 1 have multiplicity at least two;

(iii) the second eigenvalue of M is equal to the second eigenvalue of $\hat{M}$.

Proof : We use some properties of the Kronecker product [LT85]. First note that the matrix M is similar to

$$\sum_{i=1}^{m} \rho_i \begin{bmatrix} 1 & 0 \\ c_i & S_i \end{bmatrix}, \tag{6}$$

and that matrix $\hat{M} = \Sigma_{i=1}^{m} \rho_i A_i \bigotimes A_i$ is similar to

$$\begin{bmatrix} \begin{bmatrix} 1 & 0 \\ \sum_{1=1}^{m} \rho_i c_i & \sum_{1=1}^{m} \rho_i S_i \end{bmatrix} & 0 \\ \sum_{1=1}^{m} \rho_i c_i \bigotimes \begin{bmatrix} 1 & 0 \\ c_i & S_i \end{bmatrix} & \sum_{1=1}^{m} \rho_i S_i \bigotimes \begin{bmatrix} 1 & 0 \\ c_i & S_i \end{bmatrix} \end{bmatrix}. \tag{7}$$

Note also that the latter matrix is permutationally similar to a block triangular matrix with diagonal blocks 1, $\Sigma_{i=1}^{m} \rho_i S_i$, $\Sigma_{i=1}^{m} \rho_i S_i$, and $\Sigma_{i=1}^{m} \rho_i S_i \bigotimes S_i$. The assertions of part (i) and (ii) of the theorem follow from this observation. To prove (iii) we need to show that the maximum eigenvalue of $\Sigma_{i=1}^{m} \rho_i S_i \bigotimes S_i$ is less than or equal to the maximum eigenvalue $\Sigma_{i=1}^{m} \rho_i S_i$.

Let μ be the largest eigenvalue of $\Sigma_{i=1}^{m} \rho_i S_i$ (i.e. the second largest eigenvalue of M) and ν be the largest eigenvalue of $\Sigma_{i=1}^{m} \rho_i S_i \bigotimes S_i$. To prove (iii) we have to show that $\mu \geq \nu$. To this end we make use of the fact that the spectrum of $\Sigma_{i=1}^{m} \rho_i S_i$ is the same as the spectrum of $I \bigotimes \left\{ \Sigma_{i=1}^{m} \rho_i S_i \right\} = \left\{ \Sigma_{i=1}^{m} \rho_i I \bigotimes S_i \right\}$. For every $z \in \mathbb{R}^{n^2}$ we have that

$$\begin{aligned} z^T \Big\{ \sum_{i=1}^{m} \rho_i I \bigotimes S_i - \sum_{i=1}^{m} \rho_i S_i \bigotimes S_i \Big\} z &= z^T \Big\{ \sum_{i=1}^{m} \rho_i (I - S_i) \bigotimes S_i \Big\} z && (8) \\ &\geq 0. && (9) \end{aligned}$$

since the S_i are positive definite and the $(I - S_i)$ positive semi-definite. In particular, by Rayleigh - Ritz theorem [HJ85],

$$\begin{aligned} \mu &= max_{\|z\|=1}\, z^T \Big\{ \sum_{i=1}^{m} \rho_i I \bigotimes S_i \Big\} z && (10) \\ \nu &= max_{\|z\|=1}\, z^T \Big\{ \sum_{i=1}^{m} \rho_i S_i \bigotimes S_i \Big\} z && (11) \end{aligned}$$

and $\mu \geq \nu$ which completes the proof. □

Remark 2.1: If the matrices $\{B_1, ..., B_m\}$ in Theorem 2.2 satisfy (4) but not (3), then (5) holds but the matrices S_i need not be symmetric. This implies that parts (i) and (ii) of Theorem 2.2 hold for convex combinations of any column stochastic matrices. However, for part (iii) of the theorem the symmetry and the positive definiteness of the S_i's is important.

Remark 2.2: One may extend the above theorem to consider convex combinations of higher order Kronecker products.

3 TCP matrices

One can generate the family $A_1, .., A_m$ from any column stochastic matrix by replacing some of its columns by the corresponding columns of the identity. A natural question is whether Theorem 2.2 remains true also in this case. Parts (i) and (ii) of Theorem 2.2 follow immediately from Remark 2.1 follow. However, part (iii) is not true as the following example demonstrates.

Example 3.1 *Let*

$$A = \begin{bmatrix} 0.1 & 0.9 \\ 0.9 & 0.1 \end{bmatrix}.$$

With $\rho_1 = \rho_2 = \rho_3 = \frac{1}{3}$ *the second largest eigenvalue of* M *is* -0.2 *and of* $\hat{M}$ *is* 0.22*.*

Remark 3.1 : The fact that the eigenvalues of the matrix A_1 are real and positive plays a central role in the proof of Theorem 2.2. Given this fact, it is natural to ask whether this condition alone is enough to prove the assertions of our theorem. Unfortunately, this is not the case as the following example shows.

Example 3.2 *Let*

$$A_1 = \begin{bmatrix} 0.5799 & 0.3093 & 0.0858 \\ 0.0569 & 0.3515 & 0.4635 \\ 0.3632 & 0.3393 & 0.4507 \end{bmatrix}.$$

and

$$A_2 = \begin{bmatrix} 0.5799 & 0.0000 & 0.0000 \\ 0.0569 & 1.0000 & 0.0000 \\ 0.3632 & 0.0000 & 1.0000 \end{bmatrix}.$$

A_2 *is generated from* A_1*. It is readily shown that the second eigenvalues of* $M = 0.4450A_1 + 0.5550A_2$ *and* $\hat{M} = 0.4450(A_1 \bigotimes A_1) + 0.5550(A_2 \bigotimes A_2)$ *do not coincide. In fact the non-Perron eigenvalues of* M *are complex.*

A TCP matrix is a column stochastic matrix. However, as the previous examples show, not every column stochastic matrix is a TCP matrix. In this section we characterise the matrices that are. We begin with the case of 2×2 matrices.

Theorem 3.1 *The following conditions on a 2×2 column stochastic matrix A are equivalent.*

(a) A is a TCP.

(b) The eigenvalues of A are positive.

(c) Trace$(A) > 1$.

Proof : (a) implies (b) by Theorem 3.2 in [BSL04]. (b) implies (c) since A has two positive eigenvalues and one of them is 1. (c) implies (a) as follows. Let

$$A = \begin{bmatrix} a & 1-b \\ 1-a & b \end{bmatrix}$$

where $a, b \in (0,1)$ and $a + b > 1$. We have to find α, β_1, β_2 such that the matrix A is TCP. Since $Trace(A) > 1$ it follows that $a > 1 - b$. Choose $\alpha \in (1-b, a)$. This interval is a subinterval of $(0,1)$ and it follows that $\alpha \in (0,1)$. One may choose β_1 and β_2 that satisfy $\beta_1 = \frac{a-\alpha}{1-\alpha}$ and $\beta_2 = \frac{b-(1-\alpha)}{\alpha}$. It is easily verified that β_1 and β_2 are both in the interval $(0,1)$. □

We continue with necessary conditions when $n \geq 3$. Since a TCP matrix is the sum of a diagonal matrix and a rank-1 matrix, it follows that for every $i \neq j$,

$$rank A\Big[\{i;j\}; < n > /\{i;j\}\Big] = 1, \qquad (12)$$

where $A\Big[\alpha, \beta\Big]$ denotes the submatrix of A based on the rows indexed by α and positive columns indexed by β, and $< n > = \{1, 2, ..., n\}$. This means that for all $k \notin \{i, j\}$, the ratios $r_{ij} = \frac{a_{ik}}{a_{jk}}$ are the same. Observe also that

$$\begin{aligned} a_{ik} &= \alpha_i(1-\beta_k) \\ a_{jk} &= \alpha_j(1-\beta_k) \end{aligned}$$

where the α's and the β's are as in A_1 in Section 1. It therefore follows that

$$\alpha_i = r_{ij}\alpha_j. \qquad (13)$$

Define $r_{ii} = 1; i = 1, ..., n$, and observe that $\alpha_i = r_{ik}\alpha_k = r_{ik}r_{kj}\alpha_j$. Let R $= < r_{ij} >$. From this we get another necessary condition for the matrix A to be TCP:

$$r_{ij} = r_{ik}r_{kj}, \quad \forall i, j, k \quad \in < n >$$

This corresponds to

$$Rank(R) = 1. \qquad (14)$$

To obtain another necessary condition we denote by m_i the maximal non-diagonal entry in the i'th row of A and define

$$m = \sum_{i=1}^{n} m_i.$$

From $a_{ij} = \alpha_i(1-\beta_j),\ i \neq j$, and the fact that α_i and $(1-\beta_j)$ are between 0 and 1 it follows that $\alpha_i > m_i;\ i = 1,..,n$. We now use the fact that $\sum_{i=1}^{n} \alpha_i = 1$ to obtain,

$$\alpha_i < 1 - m + m_i.$$

Hence,

$$m_i < \alpha_i < 1 - m + m_i. \tag{15}$$

In particular, a necessary condition for a positive column stochastic matrix A to be TCP is

$$m < 1. \tag{16}$$

Remark 3.2: Observe that this implies that $Trace(A) > 1$. This also follows from the fact that all eigenvalues of A are positive.

We summarise the above discussion with the following proposition.

Proposition 3.1 If a positive column stochastic matrix is TCP, then it must satisfy conditions (12), (14) and (16).

Theorem 3.2 *A positive column stochastic matrix A is TCP if and only if it satisfies (12), (14) and (16), and in addition it satisfies*

$$m_k < \frac{r_{k1}}{\sum_{i=1}^{n} r_{i1}} < 1 - m + m_k,\ k = 1, ..., n. \tag{17}$$

Proof : Given the matrix A we want to find α's and β's in $(0,1)$ such that $\sum_{i=1}^{n} \alpha_i = 1$ and

$$A = \begin{bmatrix} \beta_1 & 0 & \cdots & 0 \\ 0 & \beta_2 & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \beta_n \end{bmatrix} + \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_n \end{bmatrix} [(1-\beta_1), \ldots, (1-\beta_n)],$$

It follows from (13) that $\alpha_k = r_{k1}\alpha_1;\ k = 1, ..., n$. Since the sum of the α_i's is 1 it follows that

$$\alpha_k = \frac{r_{k1}}{\sum_{i=1}^{n} r_{i1}}.$$

Such α_i's exist if (15) holds for $i = 1, ..., n$. But this is precisely condition (17). In this case we can choose $\beta_j = \frac{a_{jj}-\alpha_j}{1-\alpha_j} < 1$, so $\beta_j \in (0, 1)$ as is needed. $\square$

The following example shows that the necessity conditions (12), (14) and (16) are not sufficient.

Example 3.3 *Let*

$$A = \begin{bmatrix} 0.7 & 0.4 & 0.3 \\ 0.2 & 0.4 & 0.3 \\ 0.1 & 0.2 & 0.4 \end{bmatrix}.$$

Here (12), (14) and (16) all hold. However, $\frac{r_{11}}{r_{11}+r_{21}+r_{31}} = 0.4 \notin (0.4, 0.5)$. *Hence, A cannot be TCP.*

The following example shows that a Kronecker product of TCP matrices need not be TCP.

Example 3.4 *Let*

$$A = \begin{bmatrix} 0.9 & 0.8 \\ 0.2 & 0.2 \end{bmatrix},$$

A is a TCP matrix since its trace is greater than one; However, $A \bigotimes A$ *is*

$$B = \begin{bmatrix} 0.800 & 0.7200 & 0.7200 & 0.6400 \\ 0.0900 & 0.1800 & 0.0800 & 0.1600 \\ 0.0900 & 0.0800 & 0.1800 & 0.1600 \\ 0.0100 & 0.0200 & 0.0200 & 0.0400 \end{bmatrix}.$$

Since $b_{14} + b_{24} + b_{34} > 1$ *it follows that* B *cannot be TCP.*

Remark 3.3 : The matrices in Examples 3.2 and 3.4 (the matrix B) have a positive spectrum but are not TCP since they do not satisfy the condition (16). The matrix in Example 3.3 has a positive spectrum and satisfies (16) but is not TCP.

4 Equality for general column stochastic matrices

In the previous sections we showed that $\mu(\hat{M}) = \mu(M)$ when M and $\hat{M}$ are generated from a TCP matrix and where $\mu(X)$ is the absolute value of the second largest eigenvalue of a matrix X, and also saw examples of matrices M and $\hat{M}$ that are generated from a positive stochastic matrix where $\mu(\hat{M}) > \mu(M)$.

In this section we study the question of when does $\mu(\hat{M}) = \mu(M)$ where M is a convex combination $\Sigma_{i=1}^{m}\rho_i A_i$ of general column stochastic matrices $\{A_1, ..., A_m\}$, and $\hat{M}$ is the corresponding convex combination $\Sigma_{i=1}^{m}\rho_i A_i \bigotimes A_i$. Recall that by Remark 2.1, the spectrum of M is contained in the spectrum of $\hat{M}$.

The matrix $\hat{M}$ represents a linear operator on $C^{n\times n}$, $\Phi(X) = \Sigma_{i=1}^{m}\rho_i A_i X A_i^T$, so we want to relate the spectrum of M to the spectrum of Φ.

Lemma 4.1 For every X in $C^{n\times n}$,

$\Phi(X)e = MXe$.

Proof: $\Phi(X)e = \Sigma_{i=1}^{m}\rho_i A_i X A_i^T e = \Sigma_{i=1}^{m}\rho_i A_i e$, since A_i^T is stochastic. Hence,
$\Phi(X)e = MXe$.

Corollary 4.1: The $n^2 - n$ dimensional subspace Z of all the matrices in $C^{n\times n}$ whose row sums are zero, Z is all $X \in C^{n\times n} : Xe = 0$. This is Φ-invariant.

Theorem 4.1 : Let $X_1, X_2, ..., X_{n^2-n}, X_{n^2-n+1}, ..., X_{n^2}$ be linearly independent generalized eigenvectors of Φ corresponding to the (not necessarily distinct) eigenvalues $\lambda_1, ..., \lambda_{n^2}$, where $X_1, ..., X_{n^2-n}$ are in Z (and thus are a basis of Z). Then:
(a). $\lambda_{n^2-n+1}, ..., \lambda_{n^2}$ are the eigenvalues of M;
(b). $\mu(\hat{M}) = \mu(M)$ iff $\mu(M) \geq \rho(\Phi_z)$ where $\rho(X)$ denotes the spectral radius of X and Φ_Z is the reduction of Φ to Z.

Proof : For $k > n^2 - n$, $X_k e \neq 0$, and since X_k is a generalized eigenvector of Φ, $\Phi(X_k) = \lambda_k X_k$ or
$\lambda_k X_k + X_l$, $l > n^2 - n$ or
$\lambda_k X_k + X_l$, $l \leq n^2 - n$.

By the lemma, $MX_k e = \Phi(X_k)e = \lambda X_k e$, or
$\lambda_k X_k e + X_l e$, $l > n^2 - n$, or,
$\lambda_k X_k e$ if $l \leq n^2 - n$.

In the first and third cases $X_k e$ is an eigenvector of M corresponding to λ_k and in all cases it is a generalized eigenvector corresponding to λ_k. Thus $\lambda_{n^2-n+1}, ..., \lambda_{n^2}$ are all the eigenvalues of M and $\mu(M) = \mu(\Phi)$ iff no eiganvalue of Φ_Z is greater than $\mu(M)$.

We conclude the paper with a 2×2 example demonstrating the theorem. Consider the convex combinations M and $\hat{M}$ generated from a column stochastic matrix

$$A_1 = \begin{bmatrix} a & 1-b \\ 1-a & b \end{bmatrix} \tag{18}$$

The eigenvalues of M are 1 and $\mu = \mu(M) = Trace(M) - 1 = \rho_1(a+b-1) + \rho_2 b + \rho_3 a$. Computing the restriction of Φ to Z we find that the eienvalues of Φ_Z are μ and $\mu_1 = \rho_1(a+b-1)^2 + \rho_2 b^2 + \rho_3 a^2$. This also follows from the facts that $Trace(\hat{M}) = \rho_1(a+b)^2 + \rho_2(1+b)^2 + \rho_3(1+a)^2$, that 1 is an eigenvalue of $\hat{M}$ and that μ is a multiple eigenvalue of $\hat{M}$. Thus $\rho(\Phi_Z) = max\{\mu, \mu_1\}$ so $\mu(M) = \mu(\hat{M})$ iff $\mu \geq \mu_1$. Thus we have the following necessary and sufficient condition for $\mu(\hat{M}) = \mu(M)$.

Theorem 4.2 :

(a). If $Trace(A) \geq 1$ then the second eigenvalues of M and $\hat{M}$ are equal.

(b). If $Trace(A) = 0$ then the second eigenvalues have the same absolute values and their sum is 0.

(c). If $0 < trace(A) < 1$ then for some convex combinations $\mu(\hat{M}) = \mu(M)$ and for other combinations $\mu(\hat{M}) > \mu(M)$.

Final remark : Recall that the matrix (18) is TCP iff its trace is greater than 1.

Acknowledgments

This work was supported by Science Foundation Ireland grant 00/PI.1/C067 (Shorten, Berman) and the New-York Metropolitan Fund for Research at the Technion (Berman). The authors gratefully thank Rade Stanojevic, Douglas Leith and Helena Smigoc.

References

[BSL04] A. Berman, R. Shorten, and D. Leith. Positive matrices associated with synchronised communication networks. Linear Algebra and its Applications, vl 373, pp. 47-54, 2004.

[HJ85] R. Horn and C. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

[LT85] P. Lancaster and H. Tismenetsky. *The Theory of Matrices*. Academic Press, 1985.

[SWL05] R. Shorten, F. Wirth, and D Leith. Positive matrices and communication networks. Accepted for publication in IEEE Transactions on Networking, 2005.