Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

ELA

STATIONARY VECTORS OF STOCHASTIC MATRICES SUBJECT TO COMBINATORIAL CONSTRAINTS*

JANE BREEN† AND STEPHEN KIRKLAND†

Abstract. Given a strongly connected directed graph D, let S_D denote the set of all stochastic matrices whose directed graph is a spanning subgraph of D. We consider the problem of completely describing the set of stationary vectors of irreducible members of S_D. Results from the area of convex polytopes and an association of each matrix with an undirected bipartite graph are used to derive conditions which must be satisfied by a positive probability vector x in order for it to be admissible as a stationary vector of some matrix in S_D. Given some admissible vector x, the set of matrices in S_D that possess x as a stationary vector is also characterised.

This paper is dedicated to the memory of David Gregory.

Key words. Stochastic matrix; Stationary vector; Directed graph.

AMS subject classifications. 15B51, 15A18, 05C50.

1. Introduction. Let A be an $n \times n$ irreducible stochastic matrix, and let $D(A)$ denote the *directed graph* of A; that is, the graph on vertices labelled $1, \ldots, n$, such that (i, j) is an arc in $D(A)$ if and only if $a_{ij} > 0$. By the well-known Perron-Frobenius theorem (see [20, Section 1.1]), there exists a strictly positive left eigenvector of A corresponding to the Perron value 1, which, when normalised so that the entries sum to 1, is referred to as the *stationary vector* of A.

It is well-known that the structure of the directed graph $D(A)$ can provide qualitative information about the associated matrix; for example, A is an irreducible matrix if and only if $D(A)$ is *strongly connected* (i.e. for any pair of distinct vertices i, j there exists a path from i to j in $D(A)$). It is a frequently seen strategy within the fields of linear algebra and matrix theory to consider the effect or influence of a combinatorial property of a matrix on other analytic or algebraic properties of the matrix. In this paper, we investigate the influence of the directed graph on the stationary vector of a matrix.

*Received by the editors on February 13, 2014. Accepted for publication on April 12, 2015. Handling Editor: Bryan Shader.

†Department of Mathematics, University of Manitoba, Winnipeg, MB, R3T 2N2 (breenj3@myumanitoba.ca; stephen.kirkland@umanitoba.ca). Supported in part by the University of Manitoba under grant number 315729-352500-2000 and the University of Manitoba Graduate Fellowship, by Science Foundation Ireland under grant number 11/RFP.1/MTH/3185, and by NSERC under grant number RGPIN 06123-14.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

In particular, we are interested in the following problem: given a strongly connected directed graph D, let S_D denote the set of all stochastic matrices whose directed graphs are spanning subgraphs of D; that is,

$$S_D = \{A \in \mathbb{R}^{n\times n} \mid A \geq 0,\, A\mathbf{1} = \mathbf{1},\text{ and } D(A) \subseteq D\}.$$

Can we describe the set of all possible stationary vectors of irreducible matrices in S_D? Furthermore, supposing that x is some admissible stationary vector of an irreducible matrix in S_D, we aim to characterise the set of matrices in S_D possessing x as a stationary vector. Our principal goal in this paper is to provide a theoretical characterisation, while an exploration of possible implementations is carried out in Section 5. Though we indicate that these methods can be computationally costly for large or dense directed graphs, the example discussed in Section 6 emphasises the fact that there are many small-scale cases of practical interest that can be analysed using the results in this paper.

Stochastic matrices are central to the theory of finite homogeneous Markov chains, which are used to model a wide range of dynamic systems, such as levels of vehicle traffic or pollution in urban road networks (see [10] and [9], respectively), and population management in mathematical ecology ([6]). In each of these applications, the stationary vector represents some key feature of the system by cataloguing its long-term behaviour. In the case that the transition matrix A is primitive, the iterates of the Markov chain converge to the stationary distribution vector, independent of the initial distribution. In this way, the i^{th} entry of the stationary vector represents the long-term probability that the Markov chain is in the i^{th} state.

The main motivation for this work is the fact that many real-world systems are governed by an underlying directed graph that dictates which transitions between states are permitted. For example, the transition matrix of a Markov chain modelling vehicle traffic is constrained by the given road network, which determines the transitions between states (road segments) that are possible in one time-step. The set S_D represents the set of all possible transition matrices of a Markov chain modelling such a system, and so a solution to the problem posed above will lend itself to the design of a Markov chain which simultaneously respects this given directed graph, and achieves some desirable stationary distribution. This, then, provides an indication of how to control or influence the modelled system so that it has some desirable long-term behaviour.

There is an existing body of work motivated by this observation that real-world systems are often constrained by a given directed graph. This research centres around finding the range of possible values of some parameter of an associated Markov chain, as we range over the matrices in S_D. For example, [16] investigates the minimum rate of convergence of a Markov chain with directed graph D (as measured by a certain

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

ELA

coefficient of ergodicity) where the minimum is taken over all irreducible matrices in S_D, and [15] investigates the minimum value of the Kemeny constant over the set S_D. More relevant to our work in this paper is [17], which examines the minimum value of

$$||x||_\infty := \max\{x_1, \ldots, x_n\},$$

where $x = [x_1 \ \ldots \ x_n]^T$ denotes the stationary vector as we range over the matrices in S_D. In the words of the author, "this problem can be thought of as a 'load balancing' problem, in the sense that we seek to make the maximum entry in the stationary distribution vector as small as possible, subject to the constraint that the directed graph of the corresponding transition matrix is a spanning subgraph of D." This provides another way of viewing the long-run behaviour of a dynamical system and influencing the system in some way so that some desirable stationary distribution is achieved.

There are many other approaches to the analysis of admissible stationary vectors of matrices whose entries are permitted to vary to some degree, and it is a widely researched area because of the potential applications. One notable branch of such research is the concept of condition numbers (see [7]), which provide a measure of the change in the stationary vector when the transition matrix is perturbed. While technically related to our work, this approach lends itself more to applications in which we want to quantify the influence of small errors in the transition matrix on results found via the stationary vector.

Finally, we recall the Markov chain matrix tree theorem [11], which provides solutions for the entries of the stationary vector based on combinatorial properties of the directed graph of the matrix in question. While this result is certainly aligned with the approach we wish to take, we see no obvious way that this result may be implemented in such a way as to account for arbitrary changes in the positive entries of the transition matrix in question.

We remark that the results in this paper have applications beyond stochastic matrices and the Markov chains they represent. We observe that to any nonnegative irreducible matrix M we may associate an irreducible stochastic matrix as follows: Let $\rho(M)$ denote the Perron value of A, and let u and v be right and left Perron vectors of M, respectively, normalised so that $v^T u = 1$. Letting U denote a diagonal matrix whose i^{th} diagonal entry is u_i, we see that

$$A := \frac{1}{\rho(M)} U^{-1} M U \tag{1.1}$$

is both irreducible and stochastic, and, most importantly, the directed graphs of A and M coincide. Furthermore, the stationary vector x of A is equal to Uv, so that

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

ELA

$x_i = u_i v_i$, for $i = 1, \ldots, n$. It is known that for each i, $u_i v_i$ is the derivative of $\rho(M)$ with respect to the i^{th} diagonal entry of M. We will demonstrate this kind of application in Section 6, where we use our results to examine the sustainability of the North Atlantic right whale population.

In the following sections, we freely use common concepts from the areas of combinatorial matrix theory and finite Markov chains. We refer the interested reader to [4] and [20] for background on these topics.

2. Preliminary Observations. Let $A \in S_D$, and suppose that $x^T A = x^T$, for some $x \in \mathbb{R}^n_+$, such that $x^T \mathbf{1} = 1$, where $\mathbf{1}$ denotes the vector of all ones. Set $X := \text{diag}(x)$, the diagonal matrix with the entries of x along the diagonal, and consider the matrix $B := XA$. Notice that

$$B\mathbf{1} = x \quad \text{and} \quad \mathbf{1}^T B = x^T. \tag{2.1}$$

In this way, we see that a positive probability vector x is admissible as a stationary vector for an irreducible $A \in S_D$ if and only if there exists a nonnegative irreducible matrix B, respecting the directed graph D, such that (2.1) holds.

Our approach will be to assume that $x = [x_1 \ \ldots \ x_n]^T$ is some admissible vector (i.e. that there exists some fixed $A \in S_D$ such that $x^T A = x^T$) and we consider the equations in (2.1) as a linear system in the entries of B that are nonzero. These are known from the zero pattern of the matrix A, and we note that $b_{ij} > 0$ only if (i, j) is an arc in D. We will show that in the case that the *bipartite graph* of B is acyclic, solutions may be found to these 'variables' in terms of the entries of x – i.e. each b_{ij} may be written as an expression in $x_1, \ldots, x_n$. Then, since $b_{ij} > 0$, we achieve an inequality condition on x which must be satisfied in order for x to be admissible. Notice that this approach also allows us to construct the matrix B satisfying (2.1) for our chosen x, and hence a matrix $A \in S_D$ possessing x as a stationary vector.

It is evident from (2.1) that the matrix B is a member of the symmetric transportation polytope $\mathcal{T}(x)$ – defined as the set of nonnegative matrices whose row and column sum vectors are both equal to x. Note that the action $A \to XA$ was observed in [13] to act as a bijection between the convex set

$$\mathcal{S}_n(x) := \{A \in \mathbb{R}^{n \times n} \mid A \geq 0, A\mathbf{1} = \mathbf{1}, \text{ and } x^T A = x^T\}$$

and $\mathcal{T}(x)$, when $x \in \mathbb{R}^n_+$. We will use techniques from the areas of convex sets and transportation polytopes to generalise our approach beyond matrices whose bipartite graphs are acyclic. Note that similar techniques were used in [12] to determine the class of stochastic matrices having a common left fixed vector.

Due to the preliminary description of our approach as determining information about the class of matrices B with a "given" row and column sum vector x, it is

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

ELA

not surprising that some of the results in the next section are similar to those in the literature on transportation problems. However, we emphasise that our question, as stated, is motivated in the opposite way to those problems discussed in this linear programming question. In particular, the traditional approach there has been to choose certain row and column sum vectors, and attempt to describe the combinatorial properties of the class of matrices with those sum vectors. Our aim, however, is to investigate how a combinatorial property of a matrix (in particular, the zero pattern, or associated directed graph) affects the range of possible vectors that can be both the row and column sum vector. While very different in spirit, some of the mechanics of dealing with these questions at the small-scale remain the same, and we refer the interested reader to [2] and [8].

Remark 2.1. We note that in the definition of $\mathcal{S}_n(x)$ above, we have not excluded the possibility that a matrix A satisfying $x^TA = x^T$ may be reducible. In general, since our set S_D contains both reducible and irreducible matrices, we will make no distinction between x being a stationary vector of a matrix A, or simply a left fixed vector, which will allow the possibility that A is reducible. Relaxing the constraint in this way will allow us to compute the conditions that are required for x to be the left fixed vector of any matrix in S_D, and in Section 4 we discuss the stricter conditions under which this matrix will be irreducible.

3. The Bipartite Graph.

Let $A \in S_D$ and let x be some positive probability vector, and suppose that $x^TA = x^T$. As before, set $X := \text{diag}(x)$ and $B := XA$. In this section we describe the role of the bipartite graph of B in examining the linear system in (2.1).

Definition 3.1. Let $A = [a_{ij}]$ be an $n \times n$ matrix. The *bipartite graph of* A, denoted $\mathcal{B}(A)$, is the undirected graph with vertex set $\{r_1, \dots, r_n\} \cup \{c_1, \dots, c_n\}$, where r_i is adjacent to c_j if and only if $a_{ij} \neq 0$. The r_i are called the *row vertices* and the c_j the *column vertices*. For a directed graph D, we let $\mathcal{B}(D)$ denote the bipartite graph of the adjacency matrix of D.

Definition 3.2. Let G be an undirected graph with vertices $v_1, \dots, v_n$ and edges $e_1, \dots, e_m$. The *vertex-edge incidence matrix* of G is the $n \times m$ matrix M such that

$$m_{ij} = \begin{cases} 1, & \text{if } v_i \text{ is incident to } e_j; \\ 0, & \text{otherwise.} \end{cases}$$

Consider the bipartite graph $\mathcal{B}(B)$ (which is equal to $\mathcal{B}(A)$) with row vertices $r_1, \dots, r_n$ and column vertices $c_1, \dots, c_n$. Let $\omega(\cdot)$ denote the *weight* of an edge or a

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

vertex, and set

$$\omega(r_i c_j) := b_{ij}.$$

Suppose we also weight the row vertices so that $\omega(r_i) := x_i$, for $i = 1, \ldots, n$. Then the row sums of B (given by $B\mathbf{1} = x$) are evident from the bigraph in that each equation in $B\mathbf{1} = x$,

$$x_i = \sum_{j=1}^{n} b_{ij},$$

is given by

$$\omega(r_i) = \sum_{j=1}^{n} \omega(r_i c_j). \tag{3.1}$$

Similarly, by weighting the column vertices, the column sums $\mathbf{1}^T B = x^T$ are also immediately evident from the bipartite graph. This results in the following proposition.

Proposition 3.3. *Let B be an $n \times n$ matrix with bigraph $\mathcal{B}(B)$, and let $x \in \mathbb{R}^n_+$. The coefficient matrix of the linear system obtained from the equations*

$$B\mathbf{1} = x, \quad \mathbf{1}^T B = x^T$$

is equal to the vertex-edge incidence matrix of $\mathcal{B}(B)$.

This correspondence will be used to determine the conditions on x that ensure x is admissible as a left fixed vector of some matrix $A \in S_D$. We will also see that it is useful when the bipartite graph $\mathcal{B}(A)$ is a forest (that is, acyclic); in particular, the incidence matrix has full column rank when $\mathcal{B}(A)$ is a forest, while this is not true in general. However, we show now that it is in fact enough to consider only cases in which the bipartite graph is a forest.

Proposition 3.4. *Suppose that $x \in \mathbb{R}^n_+$ and $x^T A = x^T$ for some $A \in S_D$. Then there exists $\widetilde{A} \in S_D$ such that $x^T \widetilde{A} = x^T$, and the bipartite graph of $\widetilde{A}$ is a forest with no isolated vertices.*

Proof. Let $x \in \mathbb{R}^n_+$, and suppose $A \in S_D$ such that $x^T A = x^T$. Then $A \in \mathcal{S}_n(x)$ (defined in Section 2), which is a convex polytope, and thus A may be written as a convex combination of the extreme points of $\mathcal{S}_n(x)$:

$$A = \lambda_1 A_1 + \lambda_2 A_2 + \ldots + \lambda_m A_m, \quad \text{some } m,$$

where $\lambda_i > 0$ and $\sum_i \lambda_i = 1$. Of course, $x^T A_i = x^T$ for each i, and $\mathcal{B}(A_i) \subseteq \mathcal{B}(A)$, so $A_i \in S_D$ for each i.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

It is proven in [18] that a matrix $M \in \mathcal{T}(x)$ is an extreme point if and only if the bipartite graph of M is a forest with no isolated vertices. Since the bijection

$$\mathcal{S}_n(x) \longleftrightarrow \mathcal{T}(x)$$
$$A \longleftrightarrow XA$$

described in [13] (and referenced above) preserves the extreme points of the polytope and does not affect the bipartite graph, we conclude that $\mathcal{B}(A_i)$ is a forest with no isolated vertices, for each i. $\square$

Remark 3.5. For the remainder of this section, we operate under the assumption that the bipartite graph $\mathcal{B}(D)$ is a forest. If not, then it follows from the above that we may consider each spanning subgraph F of $\mathcal{B}(D)$ which is a forest, and find the corresponding conditions that ensure x is a left fixed vector of some matrix $\widehat{A} \in S_D$ whose bigraph is this F. Then for any $A \in S_D$ such that

$$A = \lambda_1 A_1 + \ldots + \lambda_m A_m, \tag{3.2}$$

where $\mathcal{B}(A_i) = F_i$, some spanning forest $F_i \subset \mathcal{B}(D)$, $x^T A = x^T$ if and only if x satisfies the conditions of *each* F_i.

Note that with regard to the characterisation of the set of matrices possessing x as a stationary vector, this result implies that for every forest $F_i \subset \mathcal{B}(D)$ of which x satisfies the conditions, we can construct the matrix A_i such that $x^T A_i = x^T$, and $\mathcal{B}(A_i) = F_i$. These matrices A_i are then the extreme points of the convex polytope of matrices possessing x as a left fixed vector, and every such matrix is of the form (3.2).

We now consider the rank of the linear system (2.1) through the correspondence with the incidence matrix of the bipartite graph.

Proposition 3.6. *Given a tree T, there exists an ordering of the vertices and edges of T such that when the vertex-edge incidence matrix M is obtained with respect to this order, $m_{ij} = 0$ whenever $j > i$, and $m_{ij} = 1$ for $i = j$.*

Proof. We use induction on the order of T. For the base case, suppose that $|T| = 2$. The reader may easily satisfy himself that the hypothesis holds for the incidence matrix of such a tree.

Suppose now that the induction hypothesis holds for all trees T, $|T| < m$, and consider a tree T with m vertices, of which k are pendent vertices. Label the k pendent vertices as $v_1, \ldots, v_k$ in any order, and let $e(i)$ denote the pendent edge incident to v_i. Then delete these edges and vertices. By induction there exists some method of ordering the remaining vertices so that according to this order, the vertex-edge

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

incidence matrix is

$$M = \left[\begin{array}{c|c} I_k & O \\ \hline M_{21} & M' \end{array}\right]$$

where M_{21} is some $(0,1)$-matrix and O is the zero matrix, both of appropriate dimension, while I_k is the $k \times k$ identity matrix, and M' is an $(m-k)\times(m-k-1)$ incidence matrix of the smaller tree T' obtained by the removal of the pendent vertices and edges from T. By induction, M' satisfies the condition stated in the proposition and thus the proposition holds for all trees. ∎

Corollary 3.7. *The vertex-edge incidence matrix of a forest F with no isolated vertices has full column rank.*

Proof. For a forest with r components, we may order the vertices and edges of each component according to Proposition 3.6, and list these sequentially so that the incidence matrix M of F is a block matrix with rectangular blocks representing the components of F, and zeros elsewhere; i.e.

$$M = \left[\begin{array}{c|c|c|c} C_1 & O & \dots & O \\ \hline O & C_2 & \dots & O \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline O & O & \dots & C_r \end{array}\right]. \tag{3.3}$$

Since each submatrix C_i has full column rank by Proposition 3.6, it follows that M has full column rank. ∎

Note that Proposition 3.6 and Corollary 3.7 are well-known results in the literature; see for example, [1, Lemma 2.17]. These results, coupled with Proposition 3.4, demonstrate that when $\mathcal{B}(B)$ is a forest, there are unique solutions for the weights b_{ij} in terms of the vector entries x_k, assuming (as we may) that the linear system is consistent. In order to find these solutions, and the conditions required on x that give a nonnegative solution to the linear system, we bring the augmented matrix for the linear system to row echelon form. The entries in the augmented column will then be linear combinations of $x_1, \dots, x_n$, while there will be some nonzero rows and some zero rows in the coefficient matrix of the system. The augmented entries corresponding to the nonzero rows will determine the inequality conditions described before, and the entries in the augmented column that correspond to zero rows produce extra conditions on x. There will be one zero row in each rectangular block of the matrix (3.3), indicating that these extra conditions are derived from the component structure of $\mathcal{B}(A)$. The following proposition describes these conditions, which we will refer to

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

as *component conditions.*

PROPOSITION 3.8. *Given a directed graph D, let $A \in S_D$ and let $x \in \mathbb{R}^n_+$ such that $x^T A = x^T$. For each component C of $\mathcal{B}(A)$,*

$$\sum_{r_i \in C} x_i = \sum_{c_j \in C} x_j. \tag{3.4}$$

Proof. It is easily seen that if v is a left null vector of an incidence matrix M, then if vertex i is adjacent to vertex j,

$$v_i + v_j = 0.$$

From this, it is not difficult to show that the left null space of M is spanned by vectors v_C, where, for each component C of the graph represented by M:

$$[v_C]_j = \begin{cases} 1, & \text{if } j \text{ corresponds to a row vertex in } C; \\ -1, & \text{if } j \text{ corresponds to a column vertex in } C; \\ 0, & \text{otherwise.} \end{cases}$$

Letting M be the incidence matrix of $\mathcal{B}(D)$ and considering the vectors v_C which span its null space, we have, for each component C

$$v_C^T \hat{x} = 0$$

where $\hat{x}$ is the (doubled) vector of vertex weights in the appropriate order. The component conditions (3.4) follow. □

We now consider a method for finding the unique solutions to the unknown matrix entries b_{ij}.

From the linear system, we have that

$$x_i = \sum_j b_{ij} \quad \text{and} \quad x_i = \sum_k b_{ki}.$$

Suppose we are looking for the weight b_{ij} of a particular edge $r_i c_j$ in terms of the x_i. Then

$$\begin{aligned} b_{ij} &= x_i - \sum_{k \neq j} b_{ik} \\ &= x_j - \sum_{k \neq i} b_{kj} \end{aligned} \tag{3.5}$$

Thus an expression can be given of the weight of an edge e in terms of the x_i when the expressions of the weights of all of the other edges incident to either the row vertex or the column vertex of e are known.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

Now, the weight of a pendent edge, say r_ic_j, is immediately determined by the pendent vertex to which it is incident; that is, $b_{ij} = x_i$ if $\deg(r_i) = 1$ (similarly, $b_{ij} = x_j$ if $\deg(c_j) = 1$). Referring to the edge ordering in Proposition 3.6, note that the weight of the edge e_k may be written in terms of the edges e_i, where $i < k$. Thus using (3.5) we may solve inductively for the weights of all edges of F.

We can also give a formula for computing the weight of any edge r_ic_j of F.

PROPOSITION 3.9. *Let F be a forest, $x \in \mathbb{R}^n_+$, and suppose that B is an $n \times n$ matrix such that $\mathcal{B}(B) = F$ and $B\mathbf{1} = x$, $\mathbf{1}^T B = x^T$. Then the solution for the unknown entry $b_{ij} := \omega(r_ic_j)$ may be determined directly from F as follows:*

Let C_r denote the component of $F \setminus \{r_ic_j\}$ containing r_i, and C_c the component of $F \setminus \{r_ic_j\}$ containing c_j, and let $V(C_r), V(C_c)$ denote the vertex sets of C_r and C_c, respectively. Then

$$\omega(r_ic_j) = \sum_{r_k \in V(C_r)} x_k - \sum_{c_l \in V(C_r)} x_l, \tag{3.6}$$

or, equivalently,

$$\omega(r_ic_j) = \sum_{c_k \in V(C_c)} x_k - \sum_{r_l \in V(C_c)} x_l. \tag{3.7}$$

Proof. First of all, note that the equivalence of these two expressions follows from Proposition 3.8, since $V(C_r) \cup V(C_c)$ determines a single component of F (the one containing the edge r_ic_j), and equating these expressions gives us the corresponding component condition as in (3.4).

To prove these solutions, we will use the induction alluded to above. For the base case, suppose that r_ic_j is a pendent edge, and, without loss of generality, suppose that r_i is the pendent vertex. Then $V(C_r) = \{r_i\}$, and

$$\omega(r_ic_j) = x_i,$$

and the hypothesis holds.

Now fix $r_ic_j \in F$. We assume that the weight of every other edge incident to r_i is known, and is computed according to the induction hypothesis. That is, let $r_ic_{j_1}, r_ic_{j_2}, \ldots r_ic_{j_s}$ denote the relevant edges, and let C_α denote the component of $F \setminus \{r_ic_{j_\alpha}\}$ that contains c_{j_α}, for each $\alpha = 1, \ldots, s$. Then

$$\omega(r_ic_{j_\alpha}) = \sum_{c_k \in V(C_\alpha)} x_k - \sum_{r_l \in V(C_\alpha)} x_l.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

ELA

Notice that the vertex set of the component C_r of $F \setminus \{r_i c_j\}$ that contains r_i is equal to

$$V(C_1) \cup V(C_2) \cup \cdots \cup V(C_s) \cup \{r_i\}.$$

From (3.5),

$$\begin{aligned}
\omega(r_i c_j) &= x_i - \sum_{k \neq j} b_{ik} \\
&= x_i - \sum_{\alpha=1}^{s} \omega(r_i c_{j_\alpha}) \\
&= x_i - \sum_{\alpha=1}^{s} \left(\sum_{c_k \in V(C_\alpha)} x_k - \sum_{r_l \in V(C_\alpha)} x_l \right) \\
&= \sum_{r_l \in V(C_r)} x_l - \sum_{c_k \in V(C_r)} x_k. \qquad \square
\end{aligned}$$

We note that the inductive method is more straightforward for finding the weights of every edge in the bipartite graph, while Proposition 3.9 gives a concise formula which may be more appropriate when the weight of a single edge is required.

We summarise the results of this section with the following theorem:

THEOREM 3.10. *Let D be a strongly connected directed graph, $\mathcal{B}(D)$ its bipartite graph, and let $x \in \mathbb{R}^n$. Then x is a left fixed vector of some $A \in S_D$ if and only if there is a spanning forest F of $\mathcal{B}(D)$ with no isolated vertices, such that:*

(a) the component condition (3.4) holds for each component C of F;
(b) for each edge $r_i c_j$ of F, the weight $w(r_i c_j)$ as computed in (3.6) and (3.7) is positive.

REMARK 3.11. We emphasise that in the statement of the above theorem, we say that given *any* vector in $\mathbb{R}^n$ (not necessarily positive) that satisfies the conditions in Propositions 3.8 and 3.9, it must then be a positive left fixed vector of a matrix in S_D. This is as a result of the inductive method of determining the weights - if the weight of a pendent edge is positive, this ensures the positivity of a single entry of x, and using (3.5), positivity of the whole vector follows.

If we are only interested in probability vectors, then we must include the condition that $x_1 + x_2 + \ldots + x_n = 1$.

REMARK 3.12. An alternative formulation of the conditions in Theorem 3.10 can be found in a paper of Brualdi [3]. In this article, the author discusses the more general set $\mathfrak{P}(R, S)$ of matrices with a given zero pattern $\mathfrak{P}$, row sum vector R, and

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

ELA

column sum vector S. What we have named component conditions in this special case are accounted for in the author's assumption of nondecomposability, and our edge-weight conditions can be seen to be equivalent to his characterisation of a necessary and sufficient condition for $\mathfrak{P}(R,S)$ to be nonempty [3, Theorem 2.1], when $\mathfrak{P}$ is the zero pattern of a matrix whose bipartite graph is a forest.

Notable differences between the approaches include Brualdi's investigation of the more general transportation polytope, while we focus on the symmetric case (i.e. $R = S$) since the basis of our discussion is concerned with stationary vectors of stochastic matrices. We observe that the strength of our approach rests on Remark 3.5, as we succeed in completely characterising the set of all matrices with a given row and column sum vector, while [3] gives a construction of one matrix in the class on the condition that the row and column sums are rational, along with a method for constructing a new matrix in the class from one that is known. Additionally, by focusing only on matrices in the class whose bipartite graphs are forests, we reduce the complexity of the graph before performing the analysis, meaning that we can present fewer, and more explicit conditions.

Finally, we note that Brualdi's result gives a condition for the existence of a matrix with precisely the given graph and desired row/column sums, while our approach allows for solutions with edges of the original graph being absent. The difference in intention is key here; our original problem is posed because of the applications of Markov chain theory, and it is natural in some of these settings, e.g. the closing of some roads in a road network-type example.

Example 3.13.

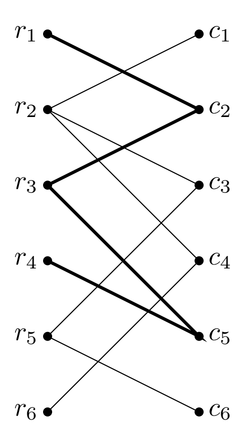


Fig. 3.1: $\mathcal{B}(D)$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

Consider the bipartite graph $\mathcal{B}(D)$ shown in Figure 3.1. We first note that it is a forest of two components, imposing the condition

$$x_1 + x_3 + x_4 = x_2 + x_5 \tag{3.8}$$

on any left fixed vector x of a matrix in S_D. We refer to the component with darker edges in Fig. 3.1 as Component 1, and the other as Component 2.

The conditions derived as the weights of the edges are as follows:

- Component 1:

$$\begin{aligned}\omega(r_1c_2) &= x_1\\ \omega(r_4c_5) &= x_4\\ \omega(r_3c_2) &= x_2 - x_1\\ \omega(r_3c_5) &= x_5 - x_4\end{aligned}$$

- Component 2:

$$\begin{aligned}\omega(r_2c_1) &= x_1\\ \omega(r_5c_6) &= x_6\\ \omega(r_6c_4) &= x_6\\ \omega(r_5c_3) &= x_5 - x_6\\ \omega(r_2c_4) &= x_4 - x_6\\ \omega(r_2c_3) &= x_2 - x_4 + x_6 - x_1\end{aligned}$$

We have presented these in the order described in Proposition 3.6, so that the reader may see more clearly the inductive method described above. For example, the weight of r_2c_3 is determined by assuming that those before it in the order are known, as follows:

$$\begin{aligned}\omega(r_2) &= \omega(r_2c_1) + \omega(r_2c_3) + \omega(r_2c_4)\\ \Rightarrow \quad \omega(r_2c_3) &= \omega(r_2) - \omega(r_2c_1) - \omega(r_2c_4)\\ &= x_2 - x_1 - (x_4 - x_6). \end{aligned} \tag{3.9}$$

To illustrate the equivalence of the method in Proposition 3.9, suppose we remove r_2c_3 from $\mathcal{B}(D)$. The corresponding "row component" C_r is the induced subgraph with vertex set $\{r_2, r_6, c_1, c_4\}$, and the "column component" C_c is induced by the vertex set $\{r_5, c_3, c_6\}$. Thus

$$\begin{aligned}&\omega(r_2c_3) = x_2 + x_6 - x_1 - x_4\\ \text{or} \qquad &\omega(r_2c_3) = x_3 + x_6 - x_5\end{aligned}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

which are equivalent to each other by (3.8), and also to the expression obtained in (3.9).

Consider

$$x^T = \left[\begin{array}{cccccc} 0.1 & 0.2 & 0.225 & 0.125 & 0.25 & 0.1 \end{array} \right].$$

It may be verified that x satisfies the conditions derived from $\mathcal{B}(D)$, and thus there exists $A \in S_D$ such that $x^T A = x^T$. But since $b_{ij} = \omega(r_i c_j)$, we can compute the matrix B to be

$$B = \begin{bmatrix} 0 & 0.1 & 0 & 0 & 0 & 0 \\ 0.1 & 0 & 0.075 & 0.025 & 0 & 0 \\ 0 & 0.1 & 0 & 0 & 0.125 & 0 \\ 0 & 0 & 0 & 0 & 0.125 & 0 \\ 0 & 0 & 0.15 & 0 & 0 & 0.1 \\ 0 & 0 & 0 & 0.1 & 0 & 0 \end{bmatrix}$$

and thus

$$\begin{aligned} A &= X^{-1}B \\ &= \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0.375 & 0.125 & 0 & 0 \\ 0 & 0.444 & 0 & 0 & 0.556 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0.6 & 0 & 0 & 0.4 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Remark 3.14. While Proposition 3.9 is an interesting result, and reinforces the importance of the bipartite graph in this problem, there are certainly computationally simpler methods by which the edge-weight conditions may be found. In particular, we may do the following: the incidence matrix of the bigraph (or coefficient matrix of the system) may be decomposed according to the components of $\mathcal{B}(B)$, so that

$$M = \left[\begin{array}{c|c|c|c} C_1 & O & \dots & O \\ \hline O & C_2 & \dots & O \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline O & O & \dots & C_r \end{array} \right].$$

As remarked before, this matrix in row echelon form has r zero rows, while the rest are nonzero. Deleting those zero rows results in an invertible matrix, $\widehat{M}$. Suppose a

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

ELA

$$
\begin{array}{cc}
 & \begin{array}{cccc|cccccc} r_1c_2 & r_4c_5 & r_3c_2 & r_3c_5 & r_2c_1 & r_5c_6 & r_6c_4 & r_5c_3 & r_2c_4 & r_2c_3 \end{array} \\
\begin{array}{c} r_1 \\ r_4 \\ c_2 \\ c_5 \\ r_3 \\ c_1 \\ c_6 \\ r_6 \\ r_5 \\ c_4 \\ c_3 \\ r_2 \end{array} &
\left[\begin{array}{cccc|cccccc}
1 & 0 & 0 & 0 & & & & & & \\
0 & 1 & 0 & 0 & & & & & & \\
1 & 0 & 1 & 0 & & & & & & \\
0 & 1 & 0 & 1 & & & & & & \\
0 & 0 & 1 & 1 & & & & & & \\
\hline
 & & & & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & 0 & 1 & 0 & 0 & 0 & 0 \\
 & & & & 0 & 0 & 1 & 0 & 0 & 0 \\
 & & & & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & 0 & 0 & 1 & 0 & 1 & 0 \\
 & & & & 0 & 0 & 0 & 1 & 0 & 1 \\
 & & & & 1 & 0 & 0 & 0 & 1 & 1
\end{array}\right]
\end{array}
$$

(a) The incidence matrix M of $\mathcal{B}(D)$ from Example 3.13.

$$
\begin{array}{cc}
 & \begin{array}{cccc|cccccc} r_1 & r_4 & c_2 & c_5 & c_1 & c_6 & r_6 & r_5 & c_4 & c_3 \end{array} \\
\begin{array}{c} r_1c_2 \\ r_4c_5 \\ r_3c_2 \\ r_3c_5 \\ r_2c_1 \\ r_5c_6 \\ r_6c_4 \\ r_5c_3 \\ r_2c_4 \\ r_2c_3 \end{array} &
\left[\begin{array}{cccc|cccccc}
1 & 0 & 0 & 0 & & & & & & \\
0 & 1 & 0 & 0 & & & & & & \\
-1 & 0 & 1 & 0 & & & & & & \\
0 & -1 & 0 & 1 & & & & & & \\
\hline
 & & & & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & 0 & 1 & 0 & 0 & 0 & 0 \\
 & & & & 0 & 0 & 1 & 0 & 0 & 0 \\
 & & & & 0 & -1 & 0 & 1 & 0 & 0 \\
 & & & & 0 & 0 & -1 & 0 & 1 & 0 \\
 & & & & 0 & 1 & 0 & -1 & 0 & 1
\end{array}\right]
\end{array}
$$

(b) The inverse of the truncated incidence matrix, $\widehat{M}^{-1}$.

$$
\widehat{M}^{-1} \cdot \left[\begin{array}{c} x_1 \\ x_4 \\ x_2 \\ x_5 \\ \hline x_1 \\ x_6 \\ x_6 \\ x_5 \\ x_4 \\ x_3 \end{array}\right]
= \left[\begin{array}{c} x_1 \\ x_4 \\ x_2 - x_1 \\ x_5 - x_4 \\ \hline x_1 \\ x_6 \\ x_6 \\ x_5 - x_6 \\ x_4 - x_6 \\ x_3 - x_5 + x_6 \end{array}\right]
= \left[\begin{array}{c} \omega(r_1c_2) \\ \omega(r_4c_5) \\ \omega(r_3c_2) \\ \omega(r_3c_5) \\ \hline \omega(r_2c_1) \\ \omega(r_5c_6) \\ \omega(r_6c_4) \\ \omega(r_5c_3) \\ \omega(r_2c_4) \\ \omega(r_2c_3) \end{array}\right]
$$

(c) The vector $\widehat{M}^{-1} \cdot \hat{x}$ displaying the edge-weight conditions as they would be produced using Prop. 3.9.

Fig. 3.2: Application of Remark 3.14 to Example 3.13

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

ELA

new vector $\hat{x}$ of length $2n-r$ is constructed from x so that the weights x_i appear in an order corresponding to the vertex order with respect to which M is written, and entries that correspond to deleted rows are deleted. It is evident that the product of the i^{th} row of $\widehat{M}^{-1}$ with $\hat{x}$ is the weight of the i^{th} edge. Thus the inverse matrix $\widehat{M}^{-1}$ is all that is necessary to determine the weights of the edges, b_{ij}.

Let us illustrate using Example 3.13. The incidence matrix for the bipartite graph in Fig. 3.1, written with respect to the order described in Prop. 3.6, is the one shown in Fig. 3.2a (where the blank entries in the off-diagonal blocks are zero). If we delete the rows corresponding to the vertices r_3 and r_2 (one from each component), and invert the resulting matrix, we obtain the matrix $\widehat{M}^{-1}$ shown in Fig. 3.2b. Constructing the vector $\hat{x}$ according to our description above and multiplying by $\widehat{M}^{-1}$ results in the vector of edge-weight conditions displayed in Fig. 3.2c, which is identical to the set of conditions derived from the bipartite graph $\mathcal{B}(D)$ shown in Fig. 3.1.

We remark that this example also demonstrates that when computing the edge weights it is sufficient to consider each component (and its vertex-edge incidence matrix) individually - an observation we will make use of when discussing the implementation of our results.

Remark 3.15. We note with regard to Example 3.13 that once the weight of each edge in $\mathcal{B}(D)$ was found, the set of conditions derived from each spanning forest $F \subset \mathcal{B}(D)$ would be easily achieved, since if $e \in \mathcal{B}(D)$ is not an edge of F, we simply have $\omega(e) = 0$. Although we observed early in this section that when $\mathcal{B}(D)$ is not acyclic it is enough to find conditions from any spanning forest of $\mathcal{B}(D)$, we now conclude that we may consider only *edge-maximal* forests, and relax the positivity constraint of the edge weights, insisting only that for each i, there exists some j such that $\omega(r_i c_j) > 0$. This is to ensure that the bipartite graph has no isolated vertices, and hence the resulting matrix – of which x is a left fixed vector – has no zero rows or columns.

4. Determining stationary vectors of irreducible matrices. Thus far, we have derived conditions by which we may determine all left fixed vectors of matrices in S_D, which may be reducible. Recall that a matrix A is *reducible* if there exists a permutation matrix P such that

$$PAP^T = \left[\begin{array}{c|c} A_1 & O \\ \hline A_{21} & A_2 \end{array}\right]. \tag{4.1}$$

In this section we determine the conditions under which x is the stationary vector of an irreducible matrix in S_D, and the conditions under which it is a left fixed vector of only reducible members of S_D. First, we determine a relationship between a reducible

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

matrix and its bipartite graph.

Definition 4.1. Let G be a bipartite graph with vertex set $\{r_1, \ldots, r_n\} \cup \{c_1, \ldots, c_n\}$, and let $\widehat{G}$ be an induced subgraph of G. We say that $\widehat{G}$ is a *balanced subgraph* of G if $\widehat{G}$ has no isolated vertices, and $r_i \in \widehat{G} \Leftrightarrow c_i \in \widehat{G}$.

Given this definition, we have the following characterisation of reducible matrices in terms of their bipartite graphs: an $n \times n$ matrix A is reducible if and only if there exists a balanced subgraph $\widehat{\mathcal{B}} \subset \mathcal{B}(A)$ such that each row vertex of $\widehat{\mathcal{B}}$ is adjacent only to column vertices of $\widehat{\mathcal{B}}$; that is, there is no edge incident to a row vertex of $\widehat{\mathcal{B}}$ that is not in the edge set of $\widehat{\mathcal{B}}$.

We now observe the following relationship between certain edges in the bipartite graph, or certain entries of the matrix $B = XA$.

Proposition 4.2. *Let B be an $n \times n$ matrix and $x \in \mathbb{R}^n_+$ such that $\mathbf{1}^T B = x^T$ and $B\mathbf{1} = x$. Then if for some permutation matrix P,*

$$PBP^T = \left[\begin{array}{c|c} B_1 & B_{12} \\ \hline B_{21} & B_2 \end{array}\right]$$

where B_1, B_2 are square submatrices, the sum of the entries in the submatrix B_{12} is equal to the sum of entries in B_{21}.

Proof. Suppose that B_1 is a $k \times k$ submatrix. Since the sum of the first k rows and the sum of the first k columns of PBP^T are equal, it follows that the sums of the entries in B_{12} and the entries in B_{21} must be equal. ◻

Corollary 4.3. *Given a directed graph D, let x be a positive vector such that $x^T A = x^T$ for some $A \in S_D$. Then A is reducible if and only if there exists a permutation matrix P such that*

$$PAP^T = \left[\begin{array}{c|c} A_1 & O \\ \hline O & A_2 \end{array}\right].$$

Equivalently, A is reducible if and only if $\mathcal{B}(A)$ is the disjoint union of two balanced subgraphs of $\mathcal{B}(D)$.

In this way, we have established that x is the left fixed vector of a reducible matrix in S_D if and only if it satisfies the component and edge-weight conditions derived from a spanning forest of $\mathcal{B}(D)$ that is a disjoint union of two balanced subgraphs.

Example 4.4. The bipartite graph $\mathcal{B}(D)$ of Example 3.13 (see Fig. 3.1) admits the two subgraphs shown in Fig. 4.1, each of which is a disjoint union of two balanced subgraphs. We note that $\mathcal{B}(D)$ does not admit any other such subgraphs – i.e. any

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

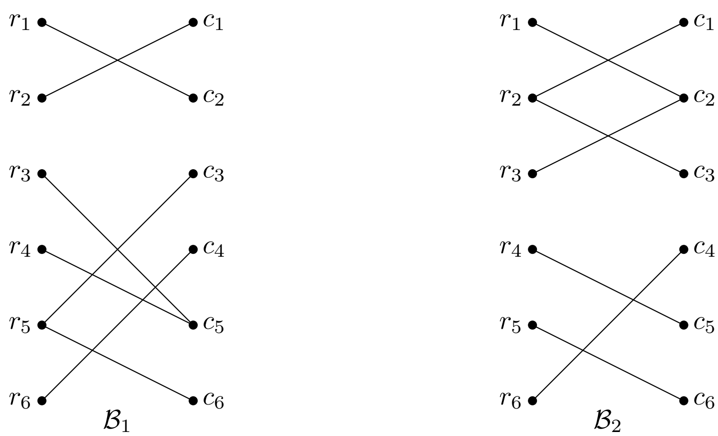


Fig. 4.1: Two spanning subgraphs of $\mathcal{B}(D)$ representing reducible matrices in S_D

reducible matrix in S_D with a positive left fixed vector must have one of these graphs in Fig. 4.1 as its bipartite graph.

In the case of $\mathcal{B}_1$, reducibility is achieved when r_2c_3, r_2c_4 and r_3c_2 are absent from the graph of $\mathcal{B}(D)$ – i.e. $\omega(r_2c_3) = \omega(r_2c_4) = \omega(r_3c_2) = 0$. Similarly, a matrix $A \in S_D$ with left fixed vector x has bipartite graph $\mathcal{B}_2$ if and only if $\omega(r_2c_4) = \omega(r_3c_5) = \omega(r_5c_3) = 0$.

Suppose now that $x \in \mathbb{R}^n_+$ is a left fixed vector of some matrix $A \in S_D$. Then A is reducible if and only if x satisfies each of

$$\begin{aligned} x_2 - x_1 &= 0 \\ x_3 - x_5 + x_6 &= 0 \\ x_4 - x_6 &= 0 \end{aligned}$$

or each of

$$\begin{aligned} x_4 - x_6 &= 0 \\ x_5 - x_4 &= 0 \\ x_5 - x_6 &= 0. \end{aligned}$$

Given a directed graph D such that $\mathcal{B}(D)$ is as in Fig. 3.1, and some x that is a left fixed vector of a matrix in S_D (i.e. x is known to satisfy the conditions of

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

Theorem 3.10), we can say that x is in fact a stationary vector if and only if at least one expression from each group above is positive. Alternatively (in this case), x is a stationary vector if the forest whose conditions are satisfied by x as per Theorem 3.10 is $\mathcal{B}(D)$ itself, since $\mathcal{B}(D)$ contains no other spanning subgraphs with no isolated vertices that are not equivalent to those shown in Fig. 4.1 (via Proposition 4.2).

Remark 4.5. As the preceding example suggests, a precise description of the set of positive vectors that serve as the stationary vectors for some irreducible member of S_D will, in general, be quite involved, particularly when $\mathcal{B}(D)$ is not acyclic. Consequently, we will not pursue that problem further here. However, as the following result shows, any stationary vector associated with an irreducible member of S_D can be approximated arbitrarily closely by vectors satisfying the conditions of Theorem 3.10.

Proposition 4.6. *Given a strongly connected directed graph D, the set of all positive left fixed vectors of matrices in S_D is the topological closure of the set of stationary distributions of irreducible members of S_D.*

Proof.

Let $A \in S_D$ be reducible, and suppose, without loss of generality, that

$$A = \left[\begin{array}{c|c} A_1 & O \\ \hline O & A_2 \end{array}\right], \tag{4.2}$$

with A_1, A_2 irreducible, and that A has a positive left fixed probability vector

$$x^T = [\beta_1 x_1^T \mid \beta_2 x_2^T],$$

where x_1^T, x_2^T are the stationary vectors of A_1, A_2 respectively, and $0 < \beta_1, \beta_2 < 1$, $\beta_1 + \beta_2 = 1$.

Consider the family of matrices defined as follows:

$$A'(\varepsilon,\delta) = \left[\begin{array}{c|c} A_1 - \varepsilon e_i e_j^T & \varepsilon e_i e_k^T \\ \hline \delta e_r e_s^T & A_2 - \delta e_r e_t^T \end{array}\right], \tag{4.3}$$

where $\varepsilon, \delta > 0$, and where indices i, j, k, r, s, t are chosen appropriately; that is, in such a way that the (i,j) entry of A_1 and the (r,t) entry of A_2 are both nonzero, and the (i,k) or the (r,s) entry of the relevant off-diagonal block is one that corresponds to an edge in D. In other words, this family of matrices $A'(\varepsilon,\delta)$ represent perturbations of the matrix A that result in an irreducible matrix in S_D.

We will show that there exists a sequence of matrices $A'(\varepsilon,\delta)$ such that the stationary vectors converge to the left fixed vector x of A as $\varepsilon \to 0$. To do this,

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

we will describe the stationary vector of a matrix of the form (4.3) using stochastic complementation, and then express δ in terms of ε in such a way as to achieve the result. We give a brief overview of the theory of stochastic complementation here (see [19] for a more in-depth discussion and rigorous proofs).

Suppose that we have an irreducible stochastic matrix A with stationary vector x, so that

$$A = \left[\begin{array}{c|c} A_1 & A_{12} \\ \hline A_{21} & A_2 \end{array}\right].$$

Define

$$\begin{aligned} S_1 &:= A_1 + A_{12}(I - A_2)^{-1}A_{21}; \\ S_2 &:= A_2 + A_{21}(I - A_1)^{-1}A_{12}. \end{aligned}$$

These are irreducible and stochastic. Let x_i be the stationary distribution of S_i; then

$$x = \left[\begin{array}{c} a_1 x_1 \\ \hline a_2 x_2 \end{array}\right]$$

where $[a_1\, a_2]$ is the stationary vector of the 2×2 matrix

$$C = \left[\begin{array}{c|c} x_1^T A_1 \mathbf{1} & x_1^T A_{12} \mathbf{1} \\ \hline x_2^T A_{21} \mathbf{1} & x_2^T A_2 \mathbf{1} \end{array}\right]. \tag{4.4}$$

Applying this approach to the matrix in (4.3), we consider the stochastic complement

$$\begin{aligned} S_1 &= (A_1 - \varepsilon e_i e_j^T) + \varepsilon e_i e_k^T (I - A_2 + \delta e_r e_t^T)^{-1} \delta e_r e_s^T \\ &= A_1 - \varepsilon e_i e_j^T + \varepsilon e_i e_s^T, \end{aligned}$$

since S_1 must be stochastic. We denote the stationary vector of S_1 by z_1. Similarly, we compute the stochastic complement

$$S_2 = A_2 - \delta e_r e_t^T + \delta e_r e_k^T,$$

and denote its stationary vector by z_2. Note that $z_1 \to x_1$ as $\varepsilon \to 0$, and $z_2 \to x_2$ as $\delta \to 0$, since A_1 and A_2 are both irreducible.

Now, the stationary vector of $A'(\varepsilon, \delta)$ may be written as

$$z(\varepsilon, \delta)^T = [\alpha_1 z_1^T \mid \alpha_2 z_2^T]$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

where $[\alpha_1 \, \alpha_2]$ is the stationary vector of the 2×2 matrix computed as in (4.4); that is,

$$\alpha_1 = \frac{\delta z_2(r)}{\varepsilon z_1(i) + \delta z_2(r)};$$
$$\alpha_2 = \frac{\varepsilon z_1(i)}{\varepsilon z_1(i) + \delta z_2(r)},$$

where $z_j(k)$ denotes the k^{th} entry of the vector z_j. Choosing

$$\delta := \frac{\varepsilon \beta_1 x_1(i)}{\beta_2 x_2(r)}$$

ensures that $\alpha_1 \to \beta_1$ and $\alpha_2 \to \beta_2$ as $\varepsilon \to 0$, and also that $z_2 \to x_2$ as $\varepsilon \to 0$.

Thus given some reducible matrix $A \in S_D$ of the form (4.2), there exists an irreducible $A' \in S_D$, dependent on some $\varepsilon > 0$, such that the stationary vector of A' converges to the left fixed vector of A as $\varepsilon \to 0$.

Now suppose that A has k strong components, i.e.

$$A = \left[\begin{array}{c|c|c|c} A_1 & O & \dots & O \\ \hline O & A_2 & \dots & O \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline O & O & \dots & A_k \end{array}\right], \qquad \text{some } k.$$

The argument presented above may be used as a technique to show that if x is a left fixed vector of the matrix A with k strong components, then given $\varepsilon > 0$, there is a matrix $\widehat{A} \in S_D$ having $k-1$ strong components and a left fixed vector $\hat{x}$ such that $||x - \hat{x}|| < \varepsilon$. Iterating this argument yields an irreducible matrix $A' \in S_D$ having stationary vector y, such that

$$||x - y|| < (k-1)\varepsilon.$$

Thus the set of all left fixed probability vectors of matrices in S_D is the topological closure of the set of all stationary vectors of irreducible matrices in S_D.
$\square$

5. Implementation. Given a directed graph D, our implementation centres around finding the edge-maximal subforests of $\mathcal{B}(D)$ and the corresponding conditions described in Propositions 3.8 and 3.9. Both tasks would require extensive time to do by hand, and so it is preferable to use mathematical programming software such as MATLAB instead. Since graphs are more difficult to deal with in such software, we

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

ELA

rely on the linear algebraic approach outlined in Remark 3.14, where it was observed that the edge-weight conditions can be found by inverting the truncated incidence matrices of the maximal subforests.

The edge-maximal acyclic subgraphs of a connected component C of $\mathcal{B}(D)$ are found by removing a single edge from each cycle in C. Any two trees obtained in this way will differ from each other only in the choice of edge removed from each cycle. This suggests that a possible update scheme may be useful, which deduces the edge-weight conditions of one tree from a similar one.

An algorithm is provided in [21] that returns all spanning trees of a connected graph. It achieves this by giving a 'root tree' and a series of edge exchanges, and the results are ordered so that two adjacent trees in the sequence differ by the exchange of exactly one edge. We can then use a formula found in [14] to compute the inverse incidence matrix of one tree from that of another, when they differ by one edge. In particular, for trees T_1, T_2 with M_1, M_2 denoting the truncated incidence matrices (obtained with respect to the same order), we can write

$$M_2 = M_1 + ye_k^T,$$

where y is equal to the k^{th} column of $M_2 - M_1$. This amounts to replacing the k^{th} edge of T_1 by some other edge. By [14, Section 0.7.4], if $1 + e_k^T M_1^{-1} y \neq 0$, then

$$M_2^{-1} = M_1^{-1} - \frac{1}{1 + e_k^T M_1^{-1} y} M_1^{-1} y e_k^T M_1^{-1}. \tag{5.1}$$

Observation 5.1. We claim that the denominator of this fraction in (5.1), $1 + e_k^T M_1^{-1} y$, is equal to either ± 1, simplifying the above expression. Our reasoning for this is as follows:

- Each of M_1, M_2 can, by an appropriate permutation of rows and columns, be brought to lower triangular form with 1s on the diagonal. Hence $det(M_1)$ is either $+1$ or -1 (depending on the signs of the row and and column permutations that bring it to triangular form), and similarly $det(M_2)$ is either $+1$ or -1.
- Thus $det(M_1^{-1} M_2)$ is also either $+1$ or -1.
- We have $M_1^{-1} M_2 = I + M_1^{-1} y e_k^T$, and it's a straightforward exercise to show that if $u, v \in R^n$, then $det(I + uv^T) = 1 + v^T u$. Hence $det(I + M_1^{-1} y e_k^T) = 1 + e_k^T M_1^{-1} y$.
- Consequently we have

$$\begin{aligned} 1 + e_k^T M_1^{-1} y &= det(I + M_1^{-1} y e_k^T) \\ &= det(M_1^{-1} M_2) \\ &= \pm 1. \end{aligned}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

In this way, (5.1) may be simplified to read:

$$M_2^{-1} = M_1^{-1} \pm M_1^{-1} y e_k^T M_1^{-1}. \tag{5.2}$$

We may determine the circumstances under which we get $+1$ or -1 by determining the signs of the row and column permutations that bring M_1 and M_2 to lower triangular form. Without loss of generality, assume that M_1 is in lower triangular form, and that M_2 is formed by removing a column of M_1 and replacing it with some new column vector that represents the new edge. Assuming the resulting matrix is no longer in lower triangular form, we know that it can be achieved by permuting the rows and columns in such a way that the corresponding vertices and edges of T_2 have been re-ordered appropriately.

Recall, however, that a row (corresponding to some vertex v) has been deleted from each of M_1 and M_2. For this reason, we describe a new method of ordering the vertices and edges of T_2, dependent on v, so that v arrives last in the vertex order and the resulting incidence matrix satsifies the conditions of Proposition 3.6. This is easily extended from the inductive method described in the proof of Proposition 3.6 by simply passing over v at each inductive step; i.e. if at some stage v is a pendent vertex of T', we do not label it among the others that we add to the order and then delete. The edge order is deduced from the vertex order by setting v as the *source vertex*, and orienting (assigning direction to) each edge so that it points away from v. Then $e(i)$ is defined to be the edge pointing towards the i^{th} vertex in the vertex order.

To determine the signs of the row and column permutations required to bring M_2 to lower triangular form, we determine the number of edges in T_2 that have a different orientation than in T_1. If no edges change orientation, then a simultaneous permutation of rows and columns is required to bring M_2 to lower triangular form, which has sign $+1$. If some edges change orientation, then we perform a simultaneous permutation of the rows and columns so that the vertex order is achieved, and we then require some further re-ordering of the columns to achieve the edge order.

Suppose the new edge e' added to create T_2 becomes oriented to some vertex v_{i_1}. Then the edge $e(i_1)$ that was oriented to v_{i_1} in T_1 has had its position in the ordering usurped by e', and we must swap these two columns of M_2. Now, $e(i_1)$ must now be oriented to some other vertex v_{i_2}, and so this enforces another exchange, of the columns representing $e(i_1)$ and $e(i_2)$. This process will continue until it reaches the vertex v_{i_k} to which the removed edge e was oriented. It is clear, then, that the permutation of the columns is a product of transpositions, the number of which is equal to the number of edges in the path from v_{i_1} to v_{i_k}. However, note that if e and e' are oriented in the same direction (i.e. right-to-left or left-to-right) then the number must be even, as v_{i_1} and v_{i_k} are both row or both column vertices. If e and e' are

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

ELA

oriented in opposite directions, the number of transpositions will be odd. Hence the sign in the update formula (5.2) will be -1 if we exchange an edge for one oriented in the same direction, and $+1$ for one oriented in the opposite direction.

We illustrate the process with an example.

Example 5.2. We refer back to the bipartite graph shown in Fig. 3.1 and examined in Example 3.13 and Remark 3.14. Suppose that we want to find the edge-weight conditions of the bipartite graph obtained from $\mathcal{B}(D)$ by exchanging the edge r_2c_3 for r_6c_6. The conditions derived from Component 1 will not be affected, so we examine only Component 2, which will be our T_1. The last vertex in the ordering of Component 2 is r_2, and we show the orientation of the edges in Fig. 5.1a, along with the truncated incidence matrix M_1, labelled with the vertex and edge orderings of T_1.

Fig. 5.1b displays the bipartite graph obtained by removing r_2c_3 and adding the edge r_6c_6, along with the matrix obtained from M_1 by replacing the column corresponding to r_2c_3 by one representing r_6c_6, and then performing a simultaneous permutation so that the rows are ordered according to the new vertex order of T_2. Evidently, further re-ordering of the columns is necessary to bring it to lower triangular form. To further re-order the columns in M_2, we see that we must swap r_6c_6 in the order for the edge previously oriented to c_6, r_5c_6, which is now oriented to r_5. We then swap r_5c_6 for r_5c_3, which is now oriented to c_3, the vertex to which r_2c_3 was oriented. Thus the permutation that brings M_2 to lower triangular form has sign $+1$, and so to find the edge-weight conditions for T_2, we compute M_2^{-1} using the update formula (5.2) with a -1 as opposed to a $+1$. This could be told immediately by comparing the orientation of r_6c_6 in T_2 with the orientation of r_2c_3 in T_1: since they are ordered the same way, this implies the permuation will have sign $+1$, and the update formula (5.2) will use a -1.

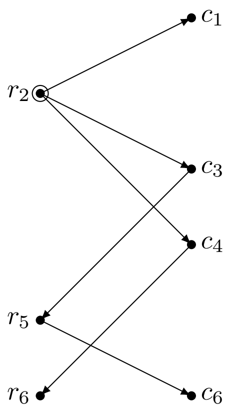
Suppose instead that we were to exchange r_2c_4 for r_6c_6, as shown in Fig. 5.1c. Since these edges have opposite orientation to each other, this means we would use the update formula with a $+1$.

Remark 5.3. We have described the process of determining the sign in the update formula using the properties of the bipartite graph, which as we have observed, are more difficult to deal with in the software we wish to use to implement our solutions. For this reason, we give the following alternative method for determining the sign in (5.2).

We have defined $y := (M_2 - M_1)e_k$, and from (5.2) we deduce that

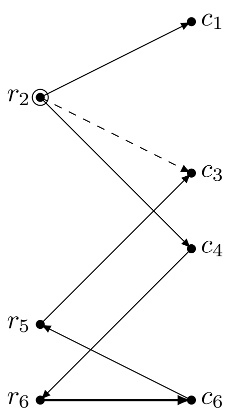
$$M_1 - M_2 = \pm M_2 M_1^{-1} y e_k^T$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015



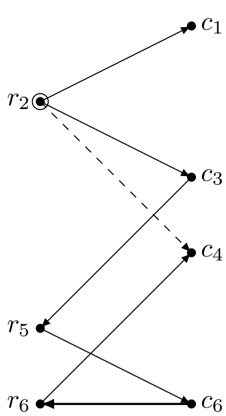
$$\begin{array}{c} \\ c_1 \\ c_6 \\ r_6 \\ r_5 \\ c_4 \\ c_3 \end{array} \begin{array}{c} \begin{array}{cccccc} r_2c_1 & r_5c_6 & r_6c_4 & r_5c_3 & r_2c_4 & r_2c_3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \end{array}$$

(a) T_1 and the truncated incidence matrix M_1



$$\begin{array}{c} \\ c_1 \\ c_3 \\ r_5 \\ c_6 \\ r_6 \\ c_4 \end{array} \begin{array}{c} \begin{array}{cccccc} r_2c_1 & \mathbf{r_6c_6} & r_5c_3 & r_5c_6 & r_6c_4 & r_2c_4 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \end{array}$$

(b) T_2 and the permuted incidence matrix M_2, before the edge re-ordering



$$\begin{array}{c} \\ c_1 \\ c_4 \\ r_6 \\ c_6 \\ r_5 \\ c_3 \end{array} \begin{array}{c} \begin{array}{cccccc} r_2c_1 & \mathbf{r_6c_6} & r_6c_4 & r_5c_6 & r_5c_3 & r_2c_3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \end{array}$$

(c) T_3 and the permuted incidence matrix M_3, before the edge re-ordering

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

and hence

$$-y = \pm M_2(M_1^{-1}y).$$

Compute $M_1^{-1}y$ (in $\mathcal{O}(n)$ operations), and then select a nonzero entry of y, say y_a. We can determine the sign from the equation

$$-y_a = \pm e_a^T M_2 M_1^{-1} y$$

in another $\mathcal{O}(n)$ operations.

Both the algorithm in [21] and the process of obtaining the inverse incidence matrices may be implemented in the case $\mathcal{B}(D)$ is a forest by applying the methods to the connected components of $\mathcal{B}(D)$. We note, however, that the time complexity of the algorithm in [21] is $\mathcal{O}(V + E + N)$, where V is the number of vertices, E the number of edges, and N the number of spanning trees of the graph, and that this is optimal. This implementation of our results suffers, then, for large or particularly dense graphs.

Despite the issues with time complexity, we note that there are many applications involving systems that are governed by relatively small or sparse directed graphs. We present the following example to demonstrate the usefulness of our results in such cases.

6. Example: North Atlantic right whale population. We recall that to any nonnegative irreducible matrix M, we can associate a stochastic matrix A by the calculation

$$A = \frac{1}{\rho(M)} U^{-1} M U,$$

as discussed in the introduction, and the i^{th} entry of the stationary vector of A then represents the derivative of $\rho(M)$ with respect to the i^{th} diagonal entry of the matrix M.

We can apply this approach to the population projection matrix of a *stage-classified matrix model* of a population (detailed in [5, Ch. 4]). Since the Perron value of this projection matrix represents the asymptotic growth of the population, information regarding the stationary vector of the corresponding stochastic matrix A will provide some insight into the sensitivity of this growth rate. This could indicate better strategies for managing a population, a common aim in examining these models.

To demonstrate how our results in this paper may be useful in such a situation, we consider the specific example of the female North Atlantic right whale population.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

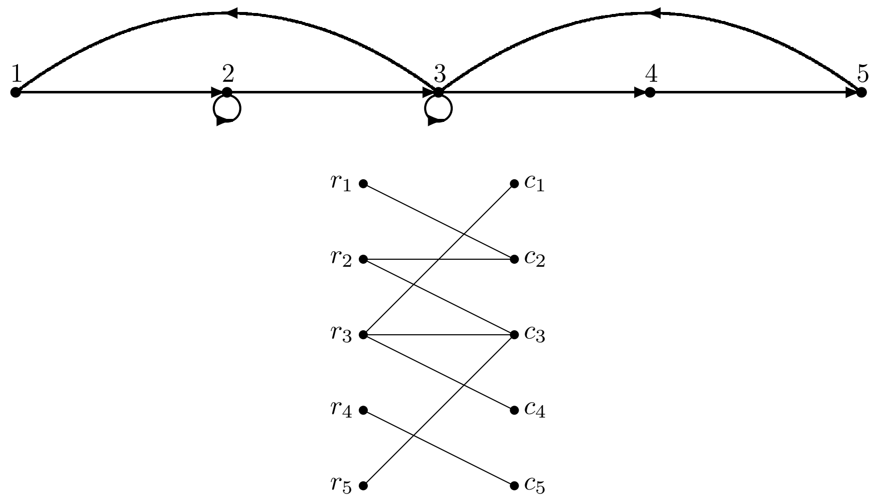


Figure 6: Life cycle graph for the North American right whale (female) and the corresponding bipartite graph

The projection matrix for this population has directed graph indicated in Fig. 6, according to [6] - which is also referred to as a life cycle graph. Each vertex represents a stage in the life cycle of the female whale (calf, immature, mature, mother, post-breeding), while arcs indicate a contribution by one class to another within a single projection interval (one year, in this case), either via reproduction or by leaving one stage and entering another. Note that the loops at each vertex in the graph represent the proportion of the population in a single class that survive and remain in the same class after the interval has passed.

Letting D denote the directed graph in Fig. 6, we consider the bipartite graph $\mathcal{B}(D)$, also shown in Fig. 6. The component condition given by this graph is

$$x_4 - x_5 = 0,$$

and the edge-weights are:

$$\begin{aligned}
\omega(r_1c_2) &= x_1\\
\omega(r_3c_1) &= x_1\\
\omega(r_3c_4) &= x_4\\
\omega(r_5c_3) &= x_5\\
\omega(r_2c_2) &= x_2 - x_1\\
\omega(r_3c_3) &= x_3 - x_1 - x_4\\
\omega(r_2c_3) &= x_1\\
\omega(r_4c_5) &= x_4.
\end{aligned}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

ELA

Thus the only non-trivial conditions on the stationary vector x of a stochastic matrix with this bipartite graph are as follows:

$$\begin{aligned} x_4 &= x_5, \\ x_2 &\geq x_1, \\ x_3 &\geq x_1 + x_4. \end{aligned} \tag{6.1}$$

From this, we see that the given directed graph D dictates some unexpected relationships between the derivatives of the Perron value with respect to the diagonal entries of the transition matrix. In this example, we have determined that the derivative with respect to either the second or third diagonal entry is the largest, regardless of how the projection matrix varies. In particular, this may be interpreted to mean that the growth rate of the population is most sensitive to increases in the proportion of members of the immature or mature (fertile) classes that survive and remain in the class within one projection interval. Though one may intuitively expect this, we note that these conditions (6.1) produce quantitative conclusions as well as the above qualitative conclusion - for example, x_3 is not only larger than x_1 and x_4, it is larger than the sum of these quantities. In general, these relationships may help to inform conservation techniques for a population modelled in this way.

Acknowledgement: The authors would like to express their thanks to the referee, whose careful reading and constructive comments helped to improve this paper.

REFERENCES

[1] Ravindra Bapat. *Graphs and Matrices*. Springer, 2010.
[2] Ethan D Bolker. Transportation polytopes. *Journal of Combinatorial Theory, Series B*, 13(3):251–262, 1972.
[3] Richard Brualdi. Convex sets of non-negative matrices. *Canad. J. Math*, 20:144–157, 1968.
[4] Richard Brualdi and Herbert Ryser. *Combinatorial Matrix Theory*. Cambridge University Press, Cambridge, 1991.
[5] Hal Caswell. *Matrix Population Models*. Sinauer Associates, Inc., Sunderland, second edition, 2001.
[6] Hal Caswell. Stage, age and individual stochasticity in demography. *Oikos*, 118(12):1763–1782, 2009.
[7] Grace Cho and Carl Meyer. Comparison of perturbation bounds for the stationary distribution of a Markov chain. *Linear Algebra and its Applications*, 335(1):137–150, 2001.
[8] Vasek Chvatal. *Linear programming*. Macmillan, 1983.
[9] Emanuele Crisostomi, Stephen Kirkland, Arieh Schlote, and Robert Shorten. Markov Chain Based Emissions Models: A Precursor for Green Control. In *Green IT: Technologies and Applications*, pages 381–400. Springer, 2011.
[10] Emanuele Crisostomi, Stephen Kirkland, and Robert Shorten. A Google-like model of road network dynamics and its application to regulation and control. *International Journal of Control*, 84(3):633–651, 2011.
[11] Mark Freĭdlin and Alexander Wentzell. *Perturbations of Stochastic Dynamic Systems*. Springer-Verlag, 1984.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 28, pp. 25-53, April 2015

[12] David Gregory, Stephen Kirkland, and Norman Pullman. Row-stochastic matrices with a common left fixed vector. *Linear Algebra and its Applications*, 169:131–149, 1992.

[13] Darald Hartfiel. A study of convex sets of stochastic matrices induced by probability vectors. *Pacific J. Math*, 52:405–418, 1974.

[14] Roger Horn and Charles Johnson. *Matrix Analysis.* Cambridge University Press, 2nd edition, 2013.

[15] Stephen Kirkland. Fastest expected time to mixing for a Markov chain on a directed graph. *Linear Algebra and its Applications*, 433(11):1988–1996, 2010.

[16] Stephen Kirkland. The minimum coefficient of ergodicity for a Markov chain with a given directed graph. *Linear Algebra and its Applications*, 2013.

[17] Steve Kirkland. Load balancing for markov chains with a specified directed graph. *Linear and Multilinear Algebra*, 62(11):1491–1508, 2014.

[18] Victor Klee and Christoph Witzgall. Facets and vertices of transportation polytopes. *Mathematics of the Decision Sciences*, 3:257–282, 1968.

[19] Carl Meyer. Stochastic complementation, uncoupling Markov chains, and the theory of nearly reducible systems. *SIAM Review*, 31(2):240–272, 1989.

[20] Eugene Seneta. *Non-negative matrices and Markov chains.* Springer-Verlag, New York, 1981.

[21] Akiyoshi Shioura and Akihisa Tamura. Efficiently scanning all spanning trees of an undirected graph. *J. Operation Research Society Japan*, 38:331–344, 1993.