ELA

ON THE KEMENY CONSTANT AND STATIONARY DISTRIBUTION VECTOR FOR A MARKOV CHAIN[*]

STEVE KIRKLAND[†]

Abstract. Suppose that A is an irreducible stochastic matrix of order n, and denote its eigenvalues by $1, \lambda_2, \ldots, \lambda_n$. The Kemeny constant, $\mathcal{K}(A)$ for the Markov chain associated with A is defined as $\mathcal{K}(A) = \sum_{j=2}^{n} \frac{1}{1-\lambda_j}$, and can be interpreted as the mean first passage from an unknown initial state to an unknown destination state in the Markov chain. Let w denote the stationary distribution vector for A, and suppose that $w_1 \leq w_2 \leq \cdots \leq w_n$. In this paper, we show that $\mathcal{K}(A) \geq \sum_{j=1}^{n}(j-1)w_j$, and we characterise the matrices yielding equality in that bound. The results are established using techniques from matrix theory and the theory of directed graphs.

Key words. Stochastic matrix, Stationary distribution vector, Kemeny constant.

AMS subject classifications. 15B51, 60J10, 15A42.

1. Introduction and preliminaries. A square entrywise nonnegative matrix A of order n is called *stochastic* if $A\mathbf{1} = \mathbf{1}$, where $\mathbf{1}$ denotes the all–ones vector of the appropriate order. Stochastic matrices are central to the theory of discrete time, time homogeneous Markov chains on a finite state space. For instance, if the stochastic matrix A is *primitive*, that is A^m has all positive entries for some $m \in \mathbb{N}$, then as is well–known, the iterates of a Markov chain with transition matrix A converge to the (unique) left Perron vector w of A, normalised so that $w^T\mathbf{1} = 1$. That eigenvector w, which is known as the *stationary distribution vector* for the Markov chain, thus carries information about the long–term behaviour of the Markov chain associated with A. We remark that in the case that A is irreducible but not primitive (in other words, the directed graph D of A is strongly connected and the greatest common divisor of the lengths of the cycles in D exceeds 1) then A still has a stationary distribution vector w, though the sequence of iterates of the corresponding Markov chain does not converge to w in general. Instead, a weaker conclusion holds, namely that for any

$$\boxed{\textbf{ELA}}$$

nonnegative vector x such that $x^T \mathbf{1} = 1$, we have

$$\lim_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} x^T A^k = w^T.$$

Thus, we see that even in this case, the stationary distribution vector w still carries some long–term information about the associated Markov chain.

If one happens to be interested in the short–term properties of a Markov chain, then the corresponding *mean first passage times* provide a useful collection of quantities for measuring the behaviour of a Markov chain over a shorter time scale. Recall that for a Markov chain with an irreducible transition matrix of order n, the mean first passage time $\mu_{i,j}$ from state i to state j is the expected number of steps necessary for the Markov chain to arrive at state j for the first time, given that it started in state i. Much is known about mean first passage times, and we refer the reader to [7] for a discussion of that topic from matrix–theoretic and graph–theoretic perspectives. In particular, a remarkable result of Kemeny asserts that for each $i = 1, \ldots, n$, the quantity

(1.1) $$\kappa_i \equiv \sum_{j=1,\ldots,n, j \neq i} \mu_{i,j} w_j$$

is independent of the choice of the index i. Indeed, it turns out that if the eigenvalues of our irreducible transition matrix A are given by $1, \lambda_2, \ldots, \lambda_n$, then

(1.2) $$\kappa_i = \sum_{j=2}^{n} \frac{1}{1 - \lambda_j}, \quad i = 1, \ldots, n.$$

The quantity on the right hand side of (1.2) is known as the *Kemeny constant* for the Markov chain associated with A, and throughout this paper we denote it by $\mathcal{K}(A)$.

The Kemeny constant admits several interpretations. From (1.1) and the fact that $\mu_{i,i} = \frac{1}{w_i}$, $i = 1, \ldots, n$, we find that for each $i = 1, \ldots, n$, $\mathcal{K}(A) + 1$ can be seen as the expected number of steps needed to arrive at a randomly chosen destination state, starting from vertex i. Alternatively, it is observed in [8] that $\mathcal{K}(A) = \sum_{i=1}^{n} \sum_{j=1,\ldots,n, j \neq i} w_i \mu_{i,j} w_j$; hence, one may view the Kemeny constant in terms of the expected number of steps in a trip from a randomly chosen initial state to a randomly chosen destination state. Finally, we note that a result of Hunter [5] facilitates an interpretation of the Kemeny constant in terms of the so–called expected time to mixing for the associated Markov chain. These various interpretations of the Kemeny constant have led to its use as an indicator of the efficiency of certain vehicle traffic networks (see [2, 3]), since in those models, low values of the Kemeny constant correspond to low average travel times. In a related vein, results in [7, Section 5.3] show that the Kemeny constant is correlated with the conditioning of the stationary

distribution vector when A is perturbed, with small values of the Kemeny constant corresponding to well–conditioned stationary distribution vectors.

In view of these last observations regarding low values of the Kemeny constant, it is not surprising that there is interest in identifying stochastic matrices A such that $\mathcal{K}(A)$ is small (in some sense). For example it is known that for an irreducible stochastic matrix A of order n, we have $\mathcal{K}(A) \geq \frac{n-1}{2}$ (see [5]), with equality holding if and only if A is the adjacency matrix of a directed cycle of length n (see [6]). In a related vein, in [6] a lower bound on $\mathcal{K}(A)$ is provided in terms of the length of a longest cycle in the directed graph of A, and the matrices yielding equality in that lower bound are characterised.

In this paper, we continue in a similar spirit by investigating how the long–term information carried by the stationary distribution vector for an irreducible stochastic matrix A is reflected in the short–term information embedded in the Kemeny constant. Specifically, in our main result (Theorem 2.2 below), we prove that if A is an $n \times n$ irreducible stochastic matrix with stationary distribution vector w, and if the entries of w are in nondecreasing order, then $\mathcal{K}(A) \geq \sum_{j=1}^{n}(j-1)w_j$. Our second key result (Theorem 3.7 below) explicitly characterises the matrices yielding equality in the bound of Theorem 2.2. We observe here that the equality characterisation given in Theorem 3.7 facilitates the construction of optimal (in terms of the Kemeny constant) transition matrices exhibiting specified long–term properties (in terms of the stationary distribution).

Throughout the sequel, we assume familiarity with basic results on stochastic matrices and Markov chains, as well as on directed graphs. The interested reader is referred to [11] for background on the former and [1] for background on the latter.

2. A lower bound on the Kemeny constant in terms of the stationary distribution. In order to establish Theorem 2.2, we require a few technical observations. To fix ideas, suppose that we have an irreducible stochastic matrix A of order n with stationary distribution vector w. We write A in partitioned form by partitioning off the last row and column of A:

$$A = \left[\begin{array}{c|c} T & (I-T)\mathbf{1} \\ \hline \frac{1}{w_n}\overline{w}^T(I-T) & 1 - \frac{1}{w_n}\overline{w}^T(I-T)\mathbf{1} \end{array} \right],$$

where $\overline{w}$ is formed from w by deleting its last entry. (We note in passing that necessarily both $(I-T)\mathbf{1}$ and $\overline{w}^T(I-T)$ are entrywise nonnegative vectors.) Continuing with this notation, it is well–known that the mean first passage times into vertex n are given by $\mu_{i,n} = e_i^T(I-T)^{-1}\mathbf{1}$, $i = 1,\ldots,n-1$ (see [11]). Further, the matrix

$$S \equiv T + \frac{1}{\overline{w}^T(I-T)\mathbf{1}}(I-T)\mathbf{1}\overline{w}^T(I-T),$$

$$\boxed{\textbf{ELA}}$$

which is called a *stochastic complement* [9], is known to be irreducible and stochastic, with the vector $\frac{1}{1-w_n}\overline{w}$ as its stationary distribution vector. It turns out that $\mathcal{K}(A)$ and $\mathcal{K}(S)$ are connected: from the proof of Theorem 6.5.1 in [7], we find that

$$(2.1) \qquad \mathcal{K}(A) = \mathcal{K}(S) + \frac{w_n \overline{w}^T (I - T)^{-1} \mathbf{1}}{1 - w_n}.$$

The following technical result will be useful in establishing Theorem 2.2. Recall that a square, entrywise nonnegative matrix is *substochastic* if each of its row sums is bounded above by 1.

LEMMA 2.1. *Let T be a substochastic matrix of order k whose spectral radius is less than* 1. *Then*

$$(2.2) \qquad trace(I - T)^{-1} \geq k.$$

Equality holds in (2.2) *if and only if T is nilpotent.*

Proof. We proceed by induction on k, and note that the result is readily established when $k = 1$. Suppose now that the result holds for some $k \in \mathbb{N}$, and that T is of order $k + 1$. We partition out the last row and column of T as

$$\left[\begin{array}{c|c} T_{1,1} & t_{1,2} \\ \hline t_{2,1} & t_{2,2} \end{array} \right].$$

Using the partitioned form of the inverse [4], we find that

$$(I - T)^{-1} =$$
$$\left[\begin{array}{c|c} (I - T_{1,1})^{-1} + \delta(I - T_{1,1})^{-1} t_{1,2} t_{2,1} (I - T_{1,1})^{-1} & \delta(I - T_{1,1})^{-1} t_{1,2} \\ \hline \delta t_{2,1} (I - T_{1,1})^{-1} & \delta \end{array} \right],$$

where $\delta = \frac{1}{1 - t_{2,2} - t_{2,1}(I - T_{1,1})^{-1} t_{1,2}}$. Hence, we have

$$trace(I - T)^{-1} \geq trace((I - T_{1,1})^{-1}) + \frac{1}{1 - t_{2,2} - t_{2,1}(I - T_{1,1})^{-1} t_{1,2}}.$$

Applying the induction hypothesis, we find readily that $trace(I - T)^{-1} \geq k + 1$.

Further, if $trace(I-T)^{-1} = k+1$, then necessarily we have $trace((I-T_{1,1})^{-1}) = k$ and $t_{2,2}+t_{2,1}(I-T_{1,1})^{-1}t_{1,2} = 0$. Again invoking the induction hypothesis, we find that $T_{1,1}$ is nilpotent; that fact, combined with the condition $t_{2,2} + t_{2,1}(I-T_{1,1})^{-1}t_{1,2} = 0$ now readily yields that T must be nilpotent. Finally, if T is nilpotent, then equality must hold in (2.2). □

ELA

Here is one of the main results of this paper.

THEOREM 2.2. *Suppose that A is an irreducible stochastic matrix of order n, that w is the stationary distribution vector of A, and that $w_1 \leq w_2 \leq \cdots \leq w_n$. Then*

$$(2.3) \qquad \mathcal{K}(A) \geq \sum_{j=1}^{n}(j-1)w_j.$$

Denote the leading $(n-1) \times (n-1)$ principal submatrix of A by T, and the leading $(n-1)$–subvector of w by $\overline{w}$. Equality holds in (2.3) if and only the following hold:

i) T is nilpotent; and

ii) $\overline{w}^T(I-T)^{-1}\mathbf{1} = \sum_{j=1}^{n}(n-j)w_j$.

Proof. We proceed by induction on n, and note that the case that $n=2$ is readily established.

Suppose now that the statements hold for some $n-1$ with $n-1 \geq 2$, and that A is of order n. We note that from the hypothesis, A can be written as

$$A = \left[\begin{array}{c|c} T & (I-T)\mathbf{1} \\ \hline \frac{1}{w_n}\overline{w}^T(I-T) & 1 - \frac{1}{w_n}\overline{w}^T(I-T) \end{array}\right].$$

Observe that $I - A$ can be factored as $I - A = XY$, where

$$X = \left[\begin{array}{c} I-T \\ \hline -\frac{1}{w_n}\overline{w}^T(I-T) \end{array}\right] \quad \text{and} \quad Y = \left[\begin{array}{c|c} I & -\mathbf{1} \end{array}\right].$$

Since XY is a full–rank factorisation of A, the eigenvalues of YX coincide with the nonzero eigenvalues of $I - A$; it now follows from (1.2) that $\mathcal{K}(A) = trace((YX)^{-1}) = trace(((I-T) + \frac{1}{w_n}\mathbf{1}\overline{w}^T(I-T))^{-1}) = trace((I-T)^{-1}) - \overline{w}^T(I-T)^{-1}\mathbf{1}$.

Next, we consider the following stochastic complement:

$$S = T + \frac{1}{\overline{w}^T(I-T)\mathbf{1}}(I-T)\mathbf{1}\overline{w}^T(I-T),$$

which is irreducible, stochastic, and has $\frac{1}{1-w_n}\overline{w}$ as its stationary distribution vector. From (2.1), we have $\mathcal{K}(A) = \mathcal{K}(S) + \frac{w_n\overline{w}^T(I-T)^{-1}\mathbf{1}}{1-w_n}$.

We now consider two cases: a) $\overline{w}^T(I-T)^{-1} \leq \sum_{j=1}^{n}(n-j)w_j$; and b) $\overline{w}^T(I-T)^{-1} > \sum_{j=1}^{n}(n-j)w_j$.

a) Since $\mathcal{K}(A) = trace((I-T)^{-1}) - \overline{w}^T(I-T)^{-1}\mathbf{1}$, we find that $\mathcal{K}(A) \geq trace((I-T)^{-1}) - \sum_{j=1}^{n}(n-j)w_j$. Applying Lemma 2.1, we have $\mathcal{K}(A) \geq n-1-\sum_{j=1}^{n}(n-j)w_j = \sum_{j=1}^{n}(j-1)w_j$.

b) Since $\mathcal{K}(A) = \mathcal{K}(S) + \frac{w_n \overline{w}^T (I-T)^{-1} \mathbf{1}}{1-w_n}$, we see that

$$\mathcal{K}(A) > \mathcal{K}(S) + \frac{w_n}{1-w_n} \sum_{j=1}^{n} (n-j)w_j.$$

Applying the induction hypothesis to S, we find that

$$\mathcal{K}(S) \geq \frac{1}{1-w_n} \sum_{j=1}^{n-1} (j-1)w_j = n - 2 - \frac{1}{1-w_n} \sum_{j=1}^{n-1} (n-j-1)w_j.$$

Consequently, we have

$$\mathcal{K}(A) > n - 2 - \frac{1}{1-w_n} \sum_{j=1}^{n-1} (n-j-1)w_j + \frac{w_n}{1-w_n} \sum_{j=1}^{n} (n-j)w_j$$

$$= n - 1 - \sum_{j=1}^{n} (n-j)w_j = \sum_{j=1}^{n} (j-1)w_j.$$

In either case, we find that (2.3) holds.

Next we consider the characterisation of the matrices yielding equality in (2.3). From cases a) and b) above, we find that equality holds in (2.3) if only if $trace((I-T)^{-1}) = n - 1$ and $\overline{w}^T(I-T)^{-1} = \sum_{j=1}^{n}(n-j)w_j$. From Lemma 2.1, we see that $trace((I-T)^{-1}) = n - 1$ precisely when T is nilpotent. The desired characterisation of equality in (2.3) now follows. $\blacksquare$

REMARK 2.3. Suppose that $v \in \mathbb{R}^n$ is a nonnegative vector, that $v^T \mathbf{1} = \mathbf{1}$, and that the entries of v are in nondecreasing order. Suppose further that for some index i_0 with $1 \leq i_0 \leq n-1$, we have $v_{i_0} < v_{i_0+1}$. Form $\tilde{v}$ from v by replacing the entries of v in positions i_0 and $i_0 + 1$ by $\frac{v_{i_0} + v_{i_0+1}}{2}$ in both positions. A straightforward computation reveals that $\sum_{j=1}^{n}(j-1)v_j = \sum_{j=1}^{n}(j-1)\tilde{v}_j + \frac{v_{i_0+1} - v_{i_0}}{2} > \sum_{j=1}^{n}(j-1)\tilde{v}_j$. It now follows that over the class of nonnegative vectors v in $\mathbb{R}^n$ whose entries are in nondecreasing order and sum to 1, the function $f(v) \equiv \sum_{j=1}^{n}(j-1)v_j$ is uniquely minimised when all of the entries in v are equal – i.e., when $v = \frac{1}{n}\mathbf{1}$ – and that the minimum value attained is $f(\frac{1}{n}\mathbf{1}) = \frac{n-1}{2}$. In particular, suppose that A is an irreducible stochastic matrix of order n having stationary vector w with entries in nondecreasing order. Applying Theorem 2.2 and the preceding observation, we find that $\mathcal{K}(A) \geq f(w) \geq \frac{n-1}{2}$, thus yielding an alternate proof of the inequality $\mathcal{K}(A) \geq \frac{n-1}{2}$ established in [5]. Further, this line of reasoning also shows that if $\mathcal{K}(A) = \frac{n-1}{2}$, then necessarily $w = \frac{1}{n}\mathbf{1}$.

$$\boxed{\textbf{ELA}}$$

REMARK 2.4. Suppose that B is an irreducible nonnegative matrix of order n, and let x and y denote right and left Perron vectors of B, respectively, normalised so that $y^T x = 1$. Let the Perron value of B be λ_1, and denote the remaining eigenvalues of B by $\lambda_2, \ldots, \lambda_n$. Letting X denote the diagonal matrix whose diagonal entries are the corresponding entries of x, it is straightforward to verify that the matrix $A = \frac{1}{\lambda_1} X^{-1} B X$ is irreducible, stochastic, and has the vector $\begin{bmatrix} x_1 y_1 & x_2 y_2 & \cdots & x_n y_n \end{bmatrix}$ as its stationary distribution vector. If we suppose that the rows and columns of B have been simultaneously permuted so that $x_1 y_1 \leq x_2 y_2 \leq \cdots \leq x_n y_n$, then applying Theorem 2.2 to A, we obtain the following (modest) generalisation of (2.3):

$$\sum_{j=2}^{n} \frac{\lambda_1}{\lambda_1 - \lambda_j} \geq \sum_{j=1}^{n} (j-1) x_j y_j.$$

3. A characterisation of the equality case in (2.3). While Theorem 2.2 provides a characterisation of the matrices yielding equality (2.3), that characterisation is somewhat opaque, since it is framed in terms of $(I - T)^{-1}$. Evidently, it is far preferable to have a characterisation of the equality case in (2.3) that is expressed directly in terms of the matrix A. We devote the this section to establishing just such a characterisation.

Our next technical result concerns nilpotent substochastic matrices.

LEMMA 3.1. *Let T be a nilpotent substochastic matrix of order r, and let $x \in \mathbb{R}^r$ be a positive vector such that $x_1 \leq x_2 \leq \cdots \leq x_r$. Suppose that $(I - T)^{-1} \mathbf{1} \geq 0$ and $x^T (I - T) \geq 0^T$, let $H = (I - T)^{-1}$, and let Δ denote the directed graph of T. Then for $1 \leq i, j \leq r$, we have*

$$(3.1) \qquad h_{i,j} \leq \begin{cases} 0, & \text{if there is no walk from } i \text{ to } j \text{ in } \Delta \\ 1, & \text{if } i < j \text{ and there is a walk from } i \text{ to } j \text{ in } \Delta \\ 1, & \text{if } i = j \\ \frac{x_j}{x_i}, & \text{if } i > j \text{ and there is a walk from } i \text{ to } j \text{ in } \Delta. \end{cases}$$

Proof. We first claim that for any $1 \leq i, j \leq r, h_{i,j} \leq 1$. To establish the claim, we proceed by induction on r. For the case $r = 2$, we have either

$$T = \begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix},$$

where $a \leq 1$, or

$$T = \begin{bmatrix} 0 & 0 \\ b & 0 \end{bmatrix},$$

where $b \leq \frac{x_1}{x_2}$. In either case, (3.1) follows readily.

ELA

Suppose now that the claim holds for some $r - 1 \geq 2$, and that T is of order r. Evidently each diagonal entry of H is 1, so suppose that we have distinct indices i, j, and consider $h_{i,j}$. If there is no walk from i to j in Δ, then certainly $h_{i,j} = 0$. Next, suppose that there is a walk from i to j in Δ, and that the longest such walk has length at most $r - 2$. Then there is a proper principal submatrix of T, say $\tilde{T}$, such that $h_{i,j}$ is equal to an appropriate entry of $(I - \tilde{T})^{-1}$. In that case, we find that $h_{i,j} \leq 1$ from the induction hypothesis. Finally, suppose that the longest walk from i to j in Δ has length $r - 1$. Then

$$h_{i,j} = \sum_l t_{i,l} e_l^T (I - T)^{-1} e_j,$$

where the sum is taken over indices l such that $t_{i,l} > 0$, and in addition, the longest walk from l to j has length at most $r - 2$. From the cases already considered, we find that $e_l^T (I - T)^{-1} e_j \leq 1$ for each such l, and since $\sum_{l=1}^r t_{i,l} \leq 1$, we deduce that $h_{i,j} \leq 1$. This completes the proof of the claim.

Next, we claim that $h_{i,j} \leq \frac{x_j}{x_i}$ for $i, j = 1, \ldots, r$. To establish this claim, we consider the matrix $U = X^{-1} T^T X$, where $X \equiv diag(x)$ is the diagonal matrix whose diagonal entries are the corresponding entries of x. It is readily verified that U is nilpotent and substochastic, with $(I - U)\mathbf{1} \geq 0$ and $x^T(I - U) \geq 0^T$. Applying our earlier claim to U, we see that $(I - U)_{j,i}^{-1} \leq 1$ for $i, j = 1, \ldots, r$. This last now readily yields that $h_{i,j} \leq \frac{x_j}{x_i}$ for $i, j = 1, \ldots, r$.

Finally, with both claims in hand, it is straightforward to verify that (3.1) holds. $\square$

Recall that a directed graph on n vertices is a *transitive tournament* if its adjacency matrix is permutationally similar to the $n \times n$ $(0, 1)$ matrix with 1s on and above the first superdiagonal, and 0s elsewhere. Much is known about tournaments in general (and transitive tournaments in particular); we refer the interested reader to the classic book of Moon [10] for background material. It is known that each transitive tournament on $n \geq 2$ vertices has a unique Hamilton path, and that in fact each transitive tournament is uniquely identified by its Hamilton path.

Lemma 3.1 enables us to establish the following.

COROLLARY 3.2. *Let T, H and x be as in Lemma 3.1, and let D denote the directed graph of H. If $x^T(I - T)^{-1}\mathbf{1} = \sum_{j=1}^r (r + 1 - j)x_j$, then*

i) D is formed from a transitive tournament by adding a loop at each vertex; and

ii) for all $i, j = 1, \ldots, r$,

(3.2)
$$h_{i,j} = \begin{cases} 0, & \text{if } i \not\rightarrow j \text{ in } D \\ 1, & \text{if } i \leq j \text{ and } i \rightarrow j \text{ in } D \\ \frac{x_j}{x_i}, & \text{if } i > j \text{ and } i \rightarrow j \text{ in } D. \end{cases}$$

ELA

S. Kirkland

Proof. Let B denote the adjacency matrix of D and let U, V denote the strictly upper triangular part and the lower triangular part of B, respectively (observe that V has 1s on the diagonal). Let L denote the lower triangular matrix with 1s on and below the main diagonal, and 0s elsewhere. Since T is nilpotent, D has no cycles of length greater than 1. Hence, we find that $B + B^T \leq L + L^T$. Consequently, we have $(U^T + V) + (U + V^T) \leq L + L^T$. It now follows that the lower triangular matrix $U^T + V$ satisfies $U^T + V \leq L$. From Lemma 3.1, it follows that $H \leq U + X^{-1}VX$, where $X = diag(x)$. Since $V \leq L - U^T$, we thus find that $H \leq U + X^{-1}VX \leq U + X^{-1}(L - U^T)X$.

Observe now that $x^T H \mathbf{1} \leq x^T U \mathbf{1} + \mathbf{1}^T (L - U^T) X \mathbf{1} = \mathbf{1}^T L x = \sum_{j=1}^{r} (r+1-j) x_j$. In particular, since $x^T H \mathbf{1} = \sum_{j=1}^{r} (r+1-j) x_j$, it must be the case that $H = U + X^{-1}VX$ and $V = L - U^T$. Conditions i) and ii) now follow readily. $\square$

Consider a directed graph on r vertices formed from a transitive tournament by adding a loop at each vertex. Recalling that any transitive tournament is uniquely specified by its Hamilton path, we introduce the following notation. Given a permutation $i_1, \ldots, i_r$ of the numbers $1, \ldots, r$, we let $D(i_1, \ldots, i_r)$ denote the directed graph formed from the transitive tournament with Hamilton path $i_1 \rightarrow i_2 \rightarrow \cdots \rightarrow i_r$ by adding a loop at each vertex. Further, given a positive vector $x \in \mathbb{R}^r$ with $x_1 \leq \cdots \leq x_r$, we let $H(x, D(i_1, \ldots, i_r))$ be the $r \times r$ matrix such that for each $p, q = 1, \ldots, r$,

$$h_{p,q} = \begin{cases} 0, & \text{if } p \nrightarrow q \text{ in } D(i_1, \ldots, i_r) \\ \min\{1, \frac{x_q}{x_p}\}, & \text{if } p \rightarrow q \text{ in } D(i_1, \ldots, i_r). \end{cases}$$

Evidently $H(x, D(i_1, \ldots, i_r))$ is a matrix of the type appearing in (3.2).

Our next technical result is straightforward.

LEMMA 3.3. *Suppose that P is an upper triangular matrix of order r with 1s on the main diagonal, and let z be a positive vector in $\mathbb{R}^r$. Suppose further that P is an inverse M–matrix such that $P^{-1}\mathbf{1} \geq 0, z^T P^{-1} \geq 0^T$. Denote P's leading and trailing principal submatrices of order $r-1$ by $\overline{P}$ and $\tilde{P}$, respectively, and denote the subvectors formed from z by deleting its last and first entries by $\overline{z}$ and $\tilde{z}$, respectively. Then $\overline{P}$ and $\tilde{P}$ are inverse M–matrices; further, we have $\overline{P}^{-1}\mathbf{1} \geq 0, \overline{z}^T \overline{P}^{-1} \geq 0^T, \tilde{P}^{-1}\mathbf{1} \geq 0$ and $\tilde{z}^T \tilde{P}^{-1} \geq 0^T$.*

Proof. Write P as

$$P = \left[\begin{array}{c|c} \overline{P} & u \\ \hline 0^T & 1 \end{array} \right],$$

$$\boxed{\textbf{ELA}}$$

so that

$$P^{-1} = \left[\begin{array}{c|c} \overline{P}^{-1} & -\overline{P}^{-1}u \\ \hline 0^T & 1 \end{array} \right].$$

Since P is an inverse M–matrix, so is $\overline{P}$; further, $\overline{P}^{-1}u \geq 0$. Since $z^T P^{-1} \geq 0^T$, we find that $\overline{z}^T \overline{P}^{-1} \geq 0^T$, and since $P^{-1}\mathbf{1} \geq 0$, we have $\overline{P}^{-1}\mathbf{1} \geq \overline{P}^{-1}u \geq 0$. A similar argument establishes the desired conclusions for $\tilde{P}$ and $\tilde{z}$. $\square$

Next, we investigate some useful properties of the matrix $H(x, D(i_1, \ldots, i_r))$.

PROPOSITION 3.4. *Suppose that $r \geq 3$ and that $x \in \mathbb{R}^r$ is a positive vector whose entries are in nondecreasing order. Suppose that $i_1, \ldots, i_r$ is a permutation of the numbers $1, \ldots, r$. Let $H \equiv H(x, D(i_1, \ldots, i_r))$, and suppose that H is an inverse M–matrix such that $H^{-1}\mathbf{1} \geq 0$ and $x^T H^{-1} \geq 0^T$. Then one of the following holds:*

a) $x_{i_1} \leq x_{i_2} \leq \cdots \leq x_{i_r}$;

b) $x_{i_1} \geq x_{i_2} \geq \cdots \geq x_{i_r}$;

c) there is an index k with $2 \leq k \leq r - 1$ such that $x_{i_1} \geq x_{i_2} \geq \cdots \geq x_{i_k}$ and $x_{i_k} < x_{i_{k+1}} \leq \cdots \leq x_{i_r}$.

Proof. We proceed by induction on r, and begin with the case that $r = 3$. Observe that there are six possible cases for the list of integers i_1, i_2, i_3, namely: $1, 2, 3$; $2, 1, 3$; $3, 1, 2$; $3, 2, 1$; $1, 3, 2$; and $2, 3, 1$. Since $x_1 \leq x_2 \leq x_3$, we find that in each of the first four of these cases satisfies one of a), b) and c). For the fifth case, we have

$$H = \left[\begin{array}{ccc} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & \frac{x_2}{x_3} & 1 \end{array} \right].$$

Hence,

$$H^{-1} = \left[\begin{array}{ccc} 1 & -\left(\frac{x_3 - x_2}{x_3}\right) & -1 \\ 0 & 1 & 0 \\ 0 & -\frac{x_2}{x_3} & 1 \end{array} \right].$$

Since $H^{-1}\mathbf{1} \geq 0$, then necessarily $x_2 = x_3$; hence, $x_1 \leq x_3 = x_2$, so that a) is satisfied. Similarly, for the sixth case, we have

$$H = \left[\begin{array}{ccc} 1 & 0 & 0 \\ \frac{x_1}{x_2} & 1 & 1 \\ \frac{x_1}{x_3} & 0 & 1 \end{array} \right],$$

S. Kirkland

so that

$$
H^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{x_1(x_3-x_2)}{x_2 x_3} & 1 & -1 \\ -\frac{x_1}{x_3} & 0 & 1 \end{bmatrix}.
$$

Since $x^T H^{-1} \geq 0^T$, it must be the case that $x_2 = x_3$; hence, $x_2 = x_3 \geq x_1$, so that b) is satisfied. This completes the analysis for the case that $r = 3$.

Henceforth, we suppose that $r \geq 4$ and that the induction hypothesis holds for matrices of order $r - 1$. Form $\overline{x}$ from x by deleting its i_r–th entry, and form $\tilde{x}$ from x by deleting its i_1–th entry. Applying Lemma 3.3, it follows that the matrices $\overline{H} \equiv H(\overline{x}, D(i_1, \ldots, i_{r-1}))$ and $\tilde{H} \equiv H(\tilde{x}, D(i_2, \ldots, i_r))$ both satisfy the hypotheses of our proposition. Thus, the induction hypothesis applies to both $\overline{H}$ and $\tilde{H}$. Applying the induction hypothesis to $\overline{H}$, we see that either i) $x_{i_1} \leq x_{i_2} \leq \cdots \leq x_{i_{r-1}}$; ii) $x_{i_1} \geq x_{i_2} \geq \cdots \geq x_{i_{r-1}}$; or iii) there is an index k with $2 \leq k \leq r - 2$ such that $x_{i_1} \geq x_{i_2} \geq \cdots \geq x_{i_k}$ and $x_{i_k} < x_{i_{k+1}} \leq \cdots \leq x_{i_{r-1}}$.

If i) holds and $x_{i_{r-1}} \leq x_{i_r}$, then a) holds. If i) holds and $x_{i_{r-1}} > x_{i_r}$, then applying the induction hypothesis to $\tilde{H}$, we find that necessarily it is the case that $x_{i_2} = x_{i_3} = \cdots = x_{i_{r-1}}$. But then there is a permutation matrix Q so that

$$
QHQ^T = \left[\begin{array}{ccccc|c} 1 & 1 & 1 & \cdots & 1 & \gamma \\ 0 & 1 & 1 & \cdots & 1 & \frac{x_{i_r}}{x_{i_2}} \\ 0 & 0 & 1 & \cdots & 1 & \frac{x_{i_r}}{x_{i_3}} \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & 1 & \frac{x_{i_r}}{x_{i_{r-1}}} \\ \hline 0 & 0 & 0 & \cdots & 0 & 1 \end{array} \right],
$$

where $\gamma = \min\{1, \frac{x_{i_r}}{x_{i_1}}\}$. Hence,

$$
QH^{-1}Q^T = \left[\begin{array}{ccccc|c} 1 & -1 & 0 & \cdots & 0 & \frac{x_{i_r}}{x_{i_2}} - \gamma \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & -1 & 0 \\ 0 & 0 & \cdots & 0 & 1 & -\frac{x_{i_r}}{x_{i_{r-1}}} \\ \hline 0 & 0 & 0 & \cdots & 0 & 1 \end{array} \right].
$$

Since $H^{-1}\mathbf{1} \geq 0$, we find that $\gamma \leq \frac{x_{i_r}}{x_{i_2}}$. Since $x_{i_r} < x_{i_2}$, we have $\gamma < 1$, so it must be that case that $\gamma = \frac{x_{i_r}}{x_{i_1}}$. We now deduce that $x_{i_1} = x_{i_2}$. Thus, b) holds.

$$\boxed{\textbf{ELA}}$$

If ii) holds, then b) or c) holds, according as we have $x_{i_{r-1}} \geq x_{i_r}$ or $x_{i_{r-1}} < x_{i_r}$, respectively.

Finally, suppose that iii) holds. Observe that it cannot be the case that $x_{i_{r-1}} > x_{i_r}$, otherwise we have a contradiction to the fact that the induction hypothesis applies to $\tilde{H}$. Hence, it must be the case that $x_{i_{r-1}} \leq x_{i_r}$, so that c) holds. $\square$

Having unearthed some of the structure of the matrix $H(x, D(i_1, \ldots, i_r))$, we next consider its inverse.

LEMMA 3.5. *Suppose that $n \geq 4$ and that $x \in \mathbb{R}^{n-1}$ is a positive vector whose entries are in nondecreasing order. Suppose that $i_1, \ldots, i_{n-1}$ is a permutation of the numbers $1, \ldots, n-1$, and consider $H \equiv H(x, D(i_1, \ldots, i_{n-1}))$. Write H^{-1} as $H^{-1} = I - T$.*

a) *Suppose that $x_{i_1} \leq x_{i_2} \leq \cdots \leq x_{i_{n-1}}$. Then for each $p, q = 1, \ldots, n-1$, we have*

$$t_{i_p, i_q} = \begin{cases} 1, & \text{if } q = p + 1 \\ 0, & \text{if } q \neq p + 1. \end{cases}$$

b) *Suppose that $x_{i_1} \geq x_{i_2} \geq \cdots \geq x_{i_{n-1}}$. Then for each $p, q = 1, \ldots, n-1$, we have*

$$t_{i_p, i_q} = \begin{cases} \frac{x_{i_{p+1}}}{x_{i_p}}, & \text{if } q = p + 1 \\ 0, & \text{if } q \neq p + 1. \end{cases}$$

c) *Suppose that there is an index k with $2 \leq k \leq n-2$ such that $x_{i_1} \geq x_{i_2} \geq \cdots \geq x_{i_k}$ and $x_{i_k} < x_{i_{k+1}} \leq \cdots \leq x_{i_{n-1}}$. Then for each $p, q = 1, \ldots, k$, we have*

$$t_{i_p, i_q} = \begin{cases} \frac{x_{i_{p+1}}}{x_{i_p}}, & \text{if } q = p + 1 \\ 0, & \text{if } q \neq p + 1, \end{cases}$$

and for each $p, q = k, \ldots, n-1$, we have

$$t_{i_p, i_q} = \begin{cases} 1, & \text{if } q = p + 1 \\ 0, & \text{if } q \neq p + 1. \end{cases}$$

Further, for each $p = k + 1, \ldots, n-1$ and $q = 1, \ldots, k$, we have $t_{i_p, i_q} = 0$. Finally, for $l = 1, \ldots, k$, we have

$$t_{i_l, i_{k+1}} = \begin{cases} 1, & \text{if } l = k \\ 0, & \text{if } x_{i_{l+1}} \geq x_{i_{k+1}} \\ \frac{x_{i_l} - x_{i_{l+1}}}{x_{i_l}}, & \text{if } x_{i_{k+1}} \geq x_{i_l} \\ \frac{x_{i_{k+1}} - x_{i_{l+1}}}{x_{i_l}}, & \text{if } x_{i_l} \geq x_{i_{k+1}} \geq x_{i_{l+1}}, \end{cases}$$

$$\boxed{\textbf{ELA}}$$

and for $l = 1, \ldots, k$, $j = k + 2, \ldots, n - 1$, we have

$$t_{i_l,i_j} = \begin{cases} 0, & \text{if } x_{i_{j-1}} \geq x_{i_l} \ \text{or } x_{i_{l+1}} \geq x_{i_j} \\ \frac{x_{i_l} - x_{i_{j-1}}}{x_{i_l}}, & \text{if } x_{i_j} \geq x_{i_l} \geq x_{i_{j-1}} \geq x_{i_{l+1}} \\ \frac{x_{i_l} - x_{i_{l+1}}}{x_{i_l}}, & \text{if } x_{i_j} \geq x_{i_l} \geq x_{i_{l+1}} \geq x_{i_{j-1}} \\ \frac{x_{i_j} - x_{i_{j-1}}}{x_{i_l}}, & \text{if } x_{i_l} \geq x_{i_j} \geq x_{i_{j-1}} \geq x_{i_{l+1}} \\ \frac{x_{i_j} - x_{i_{l+1}}}{x_{i_l}}, & \text{if } x_{i_l} \geq x_{i_j} \geq x_{i_{l+1}} \geq x_{i_{j-1}}. \end{cases}$$

Proof. a) From the hypothesis, we find that $h_{i_l,i_j} > 0$ whenever $1 \leq l \leq j \leq n - 1$. Further, for such l and j, we have $h_{i_l,i_j} = \min\{1, \frac{x_{i_j}}{x_{i_l}}\} = 1$. It now follows readily that $t_{i_l,i_{l+1}} = 1$, $l = 1, \ldots, n - 2$ while the remaining entries of T are zero.

b) The proof in this part follows by an argument analogous to that presented for part a).

c) Let P be the permutation matrix $P = \begin{bmatrix} e_{i_1} & \mid & \cdots & \mid & e_{i_{n-1}} \end{bmatrix}$, and consider $P^T H P$ which we partition as

$$P^T H P = \left[\begin{array}{c|c} H_1 & \widehat{H} \\ \hline 0 & H_2 \end{array} \right],$$

where H_1, H_2 are $k \times k$ and $(n - k - 1) \times (n - k - 1)$, respectively. Since $x_{i_{k+1}} \leq \cdots \leq x_{i_{n-1}}$, we see that H_2 is a matrix of the type described in part a), and since $x_{i_1} \geq \cdots \geq x_{i_k}$, H_1 is a matrix of the type described in part b); consequently, the desired expressions for t_{i_p,i_q} follow readily for the cases $p, q = 1, \ldots, k$ and $p, q = k + 1, \ldots, n - 1$. Evidently, we also have $t_{i_p,i_q} = 0$ for each $p = k + 1, \ldots, n - 1$ and $q = 1, \ldots, k$. It remains only to determine t_{i_l,i_j} when $l = 1, \ldots, k$ and $j = k + 1, \ldots, n - 1$.

Note that for $l = 1, \ldots, k$, $j = k + 1, \ldots, n - 1$, we have

$$t_{i_l,i_j} = e_l^T H_1^{-1} \widehat{H} H_2^{-1} e_{j-k}.$$

Observe also that for each $l = 1, \ldots, k - 1$, $e_l^T H_1^{-1} = e_l^T - \frac{x_{i_{l+1}}}{x_{i_l}} e_{l+1}^T$, while $e_k^T H_1^{-1} = e_k^T$. Further, for each $j = k + 2, \ldots, n - 1$, $H_2^{-1} e_{j-k} = e_{j-k} - e_{j-k-1}$, while $H_2^{-1} e_1 = e_1$. Finally, note that for $l = 1, \ldots, k$, $j = k + 1, \ldots, n - 1$, $\hat{h}_{l,j-k} = \min\{1, \frac{x_{i_j}}{x_{i_l}}\}$.

We now consider the case $j = k + 1$. For each $l = 1, \ldots, k$, we have

$$\begin{aligned} t_{i_l,i_{k+1}} &= e_l^T H_1^{-1} \widehat{H} H_2^{-1} e_1 = e_l^T H_1^{-1} \widehat{H} e_1 \\ &= \begin{cases} e_k^T \widehat{H} e_1, & \text{if } \ l = k \\ (e_l^T - \frac{x_{i_{l+1}}}{x_{i_l}} e_{l+1}^T) \widehat{H} e_1, & \text{if } \ l = 1, \ldots, k - 1 \end{cases} \end{aligned}$$

ELA

$$= \begin{cases} 1, & \text{if } l = k \\ \min\left\{1, \frac{x_{i_{k+1}}}{x_{i_l}}\right\} - \frac{x_{i_{l+1}}}{x_{i_l}}\min\left\{1, \frac{x_{i_{k+1}}}{x_{i_{l+1}}}\right\}, & \text{if } l = 1, \ldots, k-1. \end{cases}$$

Recalling that $x_{i_l} \geq x_{i_{l+1}}$ for $l = 1, \ldots, k-1$, it now follows that for each such l,

$$\min\left\{1, \frac{x_{i_{k+1}}}{x_{i_l}}\right\} - \frac{x_{i_{l+1}}}{x_{i_l}}\min\left\{1, \frac{x_{i_{k+1}}}{x_{i_{l+1}}}\right\} = \begin{cases} 0, & \text{if } x_{i_{l+1}} \geq x_{i_{k+1}} \\ \frac{x_{i_l}-x_{i_{l+1}}}{x_{i_l}}, & \text{if } x_{i_{k+1}} \geq x_{i_l} \\ \frac{x_{i_{k+1}}-x_{i_{l+1}}}{x_{i_l}}, & \text{if } x_{i_l} \geq x_{i_{k+1}} \geq x_{i_{l+1}}. \end{cases}$$

The desired expressions for $t_{i_l,i_{k+1}}, l = 1, \ldots, k$ are now established from the considerations above.

Next we consider the case that $k+2 \leq j \leq n-1$. Note that $t_{i_k,i_j} = e_k^T \widehat{H} e_{j-k} = \hat{h}_{k,j-k} - \hat{h}_{k,j-k-1} = \min\left\{1, \frac{x_{i_{j-k}}}{x_{i_k}}\right\} - \min\left\{1, \frac{x_{i_{j-k-1}}}{x_{i_k}}\right\} = 0$. If $1 \leq l \leq k-1$, then

$$t_{i_l,i_j} = \left(e_l^T - \frac{x_{i_{l+1}}}{x_{i_l}}e_{l+1}^T\right)\widehat{H}(e_{j-k} - e_{j-k-1})$$

(3.3) $$= \min\left\{1, \frac{x_{i_j}}{x_{i_l}}\right\} - \min\left\{1, \frac{x_{i_{j-1}}}{x_{i_l}}\right\} - \frac{x_{i_{l+1}}}{x_{i_l}}\left(\min\left\{1, \frac{x_{i_j}}{x_{i_{l+1}}}\right\} - \min\left\{1, \frac{x_{i_{j-1}}}{x_{i_{l+1}}}\right\}\right).$$

If $x_{i_{j-1}} \geq x_{i_l}$ or $x_{i_{l+1}} \geq x_{i_j}$, we find from (3.3) that $t_{i_l,i_j} = 0$. For the remaining cases, we find that

$$t_{i_l,i_j} = \begin{cases} \frac{x_{i_l}-x_{i_{j-1}}}{x_{i_l}}, & \text{if } x_{i_j} \geq x_{i_l} \geq x_{i_{j-1}} \geq x_{i_{l+1}} \\ \frac{x_{i_l}-x_{i_{l+1}}}{x_{i_l}}, & \text{if } x_{i_j} \geq x_{i_l} \geq x_{i_{l+1}} \geq x_{i_{j-1}} \\ \frac{x_{i_j}-x_{i_{j-1}}}{x_{i_l}}, & \text{if } x_{i_1} \geq x_{i_j} \geq x_{i_{j-1}} \geq x_{i_{l+1}} \\ \frac{x_{i_j}-x_{i_{l+1}}}{x_{i_l}}, & \text{if } x_{i_l} \geq x_{i_j} \geq x_{i_{l+1}} \geq x_{i_{j-1}}. \end{cases} \qquad \square$$

REMARK 3.6. Here we maintain the notation and terminology of Lemma 3.5. It is readily verified that if we are in either case a) or case b) of that lemma, then $(I-T)\mathbf{1} \geq 0$ and $x^T(I-T) \geq 0^T$. Our goal in this remark is to establish that in case c) we also have $(I-T)\mathbf{1} \geq 0$ and $x^T(I-T) \geq 0^T$.

So, suppose that we are in case c) of Lemma 3.5. Letting P be the permutation matrix $\left[\begin{array}{c|c|c} e_{i_1} & \cdots & e_{i_{n-1}} \end{array}\right]$, we find that

$$(I-T) = P\left[\begin{array}{c|c} H_1^{-1} & -H_1^{-1}\widehat{H}H_2^{-1} \\ \hline 0 & H_2^{-1} \end{array}\right]P^T.$$

Hence,

$$(I-T)\mathbf{1} = \left[\begin{array}{c} H_1^{-1}\mathbf{1} - H_1^{-1}\widehat{H}H_2^{-1}\mathbf{1} \\ \hline e_{n-k-1} \end{array}\right],$$

and so we deduce that $(I - T)\mathbf{1} \geq 0$ if and only if $H_1^{-1}\mathbf{1} - H_1^{-1}\widehat{H}H_2^{-1}\mathbf{1} \geq 0$. Since $H_2^{-1}\mathbf{1} = e_{n-k-1}$, we see that $(I - T)\mathbf{1} \geq 0$ if and only if

$$(3.4) \qquad\qquad H_1^{-1}(\mathbf{1} - \widehat{H}e_{n-k-1}) \geq 0.$$

If it happens that $x_{i_{n-1}} \geq x_{i_1}$, then $\widehat{H}e_{n-k-1} = \mathbf{1}$, and certainly (3.4) holds. On the other hand, if there is an index j between 1 and $k-1$ such that $x_{i_{j+1}} \leq x_{i_{n-1}} \leq x_{i_j}$, then

$$\widehat{H}e_{n-k-1} = \begin{bmatrix} \frac{x_{i_{n-1}}}{x_{i_1}} \\ \vdots \\ \frac{x_{i_{n-1}}}{x_{i_j}} \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

In that case, we find that

$$H_1^{-1}(\mathbf{1} - \widehat{H}e_{n-k-1}) = \begin{bmatrix} \frac{x_{i_1} - x_{i_2}}{x_{i_1}} \\ \frac{x_{i_2} - x_{i_3}}{x_{i_2}} \\ \vdots \\ \frac{x_{i_{j-1}} - x_{i_j}}{x_{i_{j-1}}} \\ \frac{x_{i_j} - x_{i_{n-1}}}{x_{i_j}} \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

and again (3.4) holds.

Similarly, we have

$$x^T(I - T) = x^T P \left[\begin{array}{c|c} H_1^{-1} & -H_1^{-1}\widehat{H}H_2^{-1} \\ \hline 0 & H_2^{-1} \end{array} \right] P^T = \left[\begin{array}{c|c} x_{i_1} e_1^T & z^T \end{array} \right] P^T,$$

where $z^T = \left[\begin{array}{ccc} x_{i_1} & \cdots & x_{i_k} \end{array} \right] H_1^{-1}\widehat{H}H_2^{-1} - \left[\begin{array}{ccc} x_{i_{k+1}} & \cdots & x_{i_{n-1}} \end{array} \right] H_2^{-1}$. Consequently, $x^T(I - T) \geq 0^T$ if and only if

$$(3.5) \qquad \left[\begin{array}{ccc} x_{i_1} & \cdots & x_{i_k} \end{array} \right] H_1^{-1}\widehat{H}H_2^{-1} - \left[\begin{array}{ccc} x_{i_{k+1}} & \cdots & x_{i_{n-1}} \end{array} \right] H_2^{-1} \geq 0^T.$$

Proceeding analogously as above, we find that the left hand side of (3.5) is the zero vector if $x_{i_1} \geq x_{i_{n-1}}$. On the other hand, if there is an index p between 0 and $n-k-2$ such that $x_{i_{k+p+1}} \geq x_{i_1} \geq x_{i_{k+p}}$, then the left side of (3.5) is given by

$$\left[\begin{array}{ccccccc} 0 & \cdots & 0 & (x_{i_{k+p+1}} - x_{i_1}) & (x_{i_{k+p+2}} - x_{i_{k+p+1}}) & \cdots & (x_{i_{n-1}} - x_{i_{n-2}}) \end{array} \right].$$

$$\boxed{\textbf{ELA}}$$

Again we see that (3.5) holds.

At last we can present the main result of this section, the characterisation of matrices yielding equality in (2.3).

THEOREM 3.7. *Suppose that A is an irreducible stochastic matrix of order n with stationary distribution vector w. Suppose further that the entries of w are in nondecreasing order, and form $\overline{w}$ from w by deleting its last entry. We have $\mathcal{K}(A) = \sum_{j=1}^{n}(j-1)w_j$ if and only if there is a permutation of the numbers $1,\ldots,n-1$, say $i_1,\ldots,i_{n-1}$, such that one of the following holds:*

a) $w_{i_1} \le w_{i_2} \le \cdots \le w_{i_{n-1}}$ and $A = \left[\begin{array}{c|c} T & (I-T)\mathbf{1} \\ \hline \frac{1}{w_n}\overline{w}^T(I-T) & 1 - \frac{1}{w_n}\overline{w}^T(I-T)\mathbf{1} \end{array} \right]$, where T is as given in Lemma 3.5 a) (taking $\overline{w}$ for x in that lemma);

b) $w_{i_1} \ge w_{i_2} \ge \cdots \ge w_{i_{n-1}}$ and $A = \left[\begin{array}{c|c} T & (I-T)\mathbf{1} \\ \hline \frac{1}{w_n}\overline{w}^T(I-T) & 1 - \frac{1}{w_n}\overline{w}^T(I-T)\mathbf{1} \end{array} \right]$, where T is as given in Lemma 3.5 b) (taking $\overline{w}$ for x in that lemma);

c) $n \ge 4$ and there is an index k with $2 \le k \le n-2$ such that $w_{i_1} \ge w_{i_2} \ge \cdots \ge w_{i_k}$ and $w_{i_k} < w_{i_{k+1}} \le \cdots \le w_{i_{n-1}}$, and

$$A = \left[\begin{array}{c|c} T & (I-T)\mathbf{1} \\ \hline \frac{1}{w_n}\overline{w}^T(I-T) & 1 - \frac{1}{w_n}\overline{w}^T(I-T)\mathbf{1} \end{array} \right],$$

where T is as given in Lemma 3.5 c) (taking $\overline{w}$ for x in that lemma).

Proof. Write A as $A = \left[\begin{array}{c|c} T & (I-T)\mathbf{1} \\ \hline \frac{1}{w_n}\overline{w}^T(I-T) & 1 - \frac{1}{w_n}\overline{w}^T(I-T)\mathbf{1} \end{array} \right]$, and suppose that $\mathcal{K}(A) = \sum_{j=1}^{n}(j-1)w_j$. From Theorem 2.2, we find that necessarily T is nilpotent and $\overline{w}^T(I-T)^{-1}\mathbf{1} = \sum_{j=1}^{n}(n-j)w_j$. Applying Corollary 3.2, we find that there are indices $i_1,\ldots,i_n$ such that $(I-T)^{-1} = H(\overline{w}, D(i_1,\ldots,i_{n-1}))$. Since $(I-T)^{-1}$ satisfies the hypotheses of Proposition 3.4, we find that either i) $w_{i_1} \le w_{i_2} \le \cdots \le w_{i_{n-1}}$, or ii) $w_{i_1} \ge w_{i_2} \ge \cdots \ge w_{i_{n-1}}$, or iii) there is an index k with $2 \le k \le n-2$ such that $w_{i_1} \ge w_{i_2} \ge \cdots \ge w_{i_k}$ and $w_{i_k} < w_{i_{k+1}} \le \cdots \le w_{i_{n-1}}$. Applying Lemma 3.5 now yields the desired expressions for the entries of T in each of the three cases of interest.

Conversely, suppose that the permutation $i_1,\ldots,i_n$ and matrix T are as in the statement, and that $A = \left[\begin{array}{c|c} T & (I-T)\mathbf{1} \\ \hline \frac{1}{w_n}\overline{w}^T(I-T) & 1 - \frac{1}{w_n}\overline{w}^T(I-T)\mathbf{1} \end{array} \right]$. From Remark 3.6 we have that $(I-T)\mathbf{1} \ge 0$ and $\overline{w}^T(I-T) \ge 0^T$, so that A is nonnegative. It now follows that A is stochastic with stationary distribution vector w. Observe that T is nilpotent. Further, from Lemma 3.5, it follows that $(I-T)^{-1} = H(\overline{w}, D(i_1,\ldots,i_{n-1}))$; a straightforward computation shows that $\overline{w}^T(I-T)^{-1}\mathbf{1} =$

ELA

$\overline{w}^T H(\overline{w}, D(i_1, \ldots, i_{n-1}))\mathbf{1} = \sum_{j=1}^{n}(n-j)w_j$, and so from Theorem 2.2 we find that
$\mathcal{K}(A) = \sum_{j=1}^{n}(j-1)w_j$, as desired. $\square$

REMARK 3.8. Let A be an irreducible stochastic matrix of order n. As we saw
in Section 1, necessarily $\mathcal{K}(A) \geq \frac{n-1}{2}$. Suppose now that equality holds in that bound
– i.e., that $\mathcal{K}(A) = \frac{n-1}{2}$. From Remark 2.3, necessarily the stationary distribution
vector of A must be $\frac{1}{n}\mathbf{1}$, and evidently equality must also hold in (2.3). Applying
Theorem 3.7, it now follows that there is a permutation of the integers $1, \ldots, n-1$,
say $i_1, \ldots, i_{n-1}$, such that $a_{i_p, i_{p+1}} = 1$, $p = 1, \ldots, n-2$, while all remaining entries of
the leading $(n-1) \times (n-1)$ principal submatrix of A are zeros. Using the fact that A
is stochastic with stationary distribution vector $\frac{1}{n}\mathbf{1}$, we find readily that $a_{i_{n-1},n} = 1$,
$a_{n,i_1} = 1$, and all remaining entries of the last row and column of A are zeros.
Consequently, our matrix A is in fact the $(0,1)$ adjacency matrix of the directed cycle
of length n given by $n \to i_1 \to i_2 \to \cdots \to i_{n-1} \to n$. This line of reasoning yields an
alternate proof of the characterisation (established in [6]) of the irreducible stochastic
matrices A of order n such that $\mathcal{K}(A) = \frac{n-1}{2}$.

The following example gives a particular instance of the class of matrices appear-
ing in Theorem 3.7 c).

EXAMPLE 3.9. Suppose that $n \geq 4$ and fix an index k with $2 \leq k \leq n-2$.
Suppose that we have a positive vector $w \in \mathbb{R}^n$ whose entries sum to 1, and whose
entries are in nondecreasing order. We consider the permutation of $1, \ldots, n-1$ given
by $i_l = k+1-l, l = 1, \ldots, k, i_j = j, j = k+1, \ldots, n-1$. Below we exhibit the matrix
of Theorem 3.7 c) that arises from the permutation $i_1, \ldots, i_{n-1}$:

$$
\left[
\begin{array}{cccc|cccc|c}
0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 \\
\frac{w_1}{w_2} & 0 & \cdots & 0 & \frac{w_2-w_1}{w_2} & 0 & \cdots & 0 & 0 \\
 & \ddots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & \cdots & \frac{w_{k-1}}{w_k} & 0 & \frac{w_k-w_{k-1}}{w_k} & 0 & \cdots & 0 & 0 \\
\hline
0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & 1 & & 0 \\
\vdots & \vdots & & \vdots & \vdots & & & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 \\
\hline
0 & \cdots & 0 & \frac{w_k}{w_n} & \frac{w_{k+1}-w_k}{w_n} & \frac{w_{k+2}-w_{k+1}}{w_n} & \cdots & \frac{w_{n-1}-w_{n-2}}{w_n} & \frac{w_n-w_{n-1}}{w_n}
\end{array}
\right].
$$

(3.6)

Observe then that for any $n \geq 4$, and any positive vector w as above, there are
at least $n-1$ distinct matrices yielding equality in (2.3), namely the $n-3$ matrices
described in (3.6) (one for each k between 2 and $n-2$) as well as the two matrices
arising from the constructions in Theorem 3.7 a) and b). If it happens that w has

$$\boxed{\textbf{ELA}}$$

distinct entries, it is readily verified that these $n - 1$ matrices are the only ones to yield equality in (2.3).

Our final example illustrates Theorem 3.7 c) for a case in which w has repeated entries.

EXAMPLE 3.10. Suppose that w is a positive vector in $\mathbb{R}^{11}$ such that $w^T \mathbf{1} = 1$. Suppose further that w is in nondecreasing order, with the extra conditions that $w_1 = w_2$, $w_6 = w_7$ and $w_9 = w_{10}$. Next we consider the indices $i_1, \ldots, i_{10}$ given by

$$\begin{bmatrix} i_1 & \cdots & i_{10} \end{bmatrix} = \begin{bmatrix} 9 & 10 & 3 & 1 & 2 & 4 & 5 & 7 & 6 & 8 \end{bmatrix}.$$

The stochastic matrix A (which yields equality in (2.3) for the vector w) that arises from Theorem 3.7 c) with the sequence $i_1, \ldots, i_{10}$ is given by:

$$\begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{w_1}{w_3} & 0 & 0 & \frac{w_3 - w_1}{w_3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & \frac{w_3}{w_{10}} & \frac{w_4 - w_3}{w_{10}} & \frac{w_5 - w_4}{w_{10}} & 0 & \frac{w_7 - w_5}{w_{10}} & \frac{w_8 - w_7}{w_{10}} & 0 & 0 & \frac{w_{10} - w_8}{w_{10}} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{w_9}{w_{11}} & 0 & \frac{w_{11} - w_9}{w_{11}}
\end{bmatrix}.$$

REFERENCES

[1] R. Brualdi and H. Ryser. *Combinatorial Matrix Theory*. Cambridge University Press, Cambridge, 1991.

[2] E. Crisostomi, S. Kirkland, and R. Shorten. A Google–like model of road network dynamics and its application to regulation and control. *International Journal of Control*, 84:633–651, 2011.

[3] E. Crisostomi, S. Kirkland, A. Schlote, and R. Shorten. Markov chain–based emissions models: a precursor for green control. In J. Kim and M. Lee (editors), *Green IT: Technologies and Applications*, Springer–Verlag, New York, 381–400, 2011.

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

[5] J. Hunter. Mixing times with applications to perturbed Markov chains. *Linear Algebra and its Applications*, 417:108–123, 2006.

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

$$\boxed{\textbf{ELA}}$$

[7] S. Kirkland and M. Neumann. *Group Inverses of M–Matrices and their Applications.* CRC
       Press, Boca Raton, 2013.
[8] M. Levene and G. Loizou. Kemeny's constant and the random surfer. *American Mathematical
       Monthly*, 109:741–745, 2002.
[9] C. Meyer. Stochastic complementation, uncoupling Markov Chains, and the theory of nearly
       reducible systems. *SIAM Review*, 31:240–272, 1987.
[10] J. Moon. *Topics on Tournaments.* Holt, Rinehart and Winston, New York, 1968.
[11] E. Seneta. *Non–Negative Matrices and Markov Chains*, second edition. Springer, New York,
       1981.