

# Directed Forests and the Constancy of Kemeny's Constant

Steve Kirkland\*

Department of Mathematics, University of Manitoba  
Winnipeg, MB, R3T 2N2  
Canada

October 7, 2019

## Abstract

Consider a discrete-time, time-homogeneous Markov chain on states  $1, \dots, n$  whose transition matrix is irreducible. Denote the mean first passage times by  $m_{jk}, j, k = 1, \dots, n$ , and stationary distribution vector entries by  $v_k, k = 1, \dots, n$ . A result of Kemeny reveals that the quantity  $\sum_{k=1}^n m_{jk} v_k$ , which is the expected number of steps needed to arrive at a randomly chosen destination state starting from state  $j$ , is – surprisingly – independent of the initial state  $j$ . In this note, we consider  $\sum_{k=1}^n m_{jk} v_k$  from the perspective of algebraic combinatorics, and provide an intuitive explanation for its independence on the initial state  $j$ . The all minors matrix tree theorem is the key tool employed.

## 1 Introduction

Let  $A$  be an irreducible row stochastic matrix of order  $n$ . Evidently the all ones vector,  $\mathbf{1}$ , is a positive right eigenvector of  $A$  corresponding to the eigenvalue 1, and from Perron–Frobenius theory,  $A$  also has a positive left eigenvector corresponding to the (algebraically simple) eigenvalue 1. If we normalise that left eigenvector, say  $v$ , so that its entries sum to 1, then  $v$  is known as the *stationary distribution vector* for the Markov chain associated with  $A$ . The following argument shows how the entries of  $v$  can be written in terms of determinants. Letting  $\text{adj}(I - A)$  denote the adjoint of  $I - A$ , we have  $\text{adj}(I - A)(I - A) = (I - A)\text{adj}(I - A) = (\det(I - A))I = 0$ . Hence, each column of  $\text{adj}(I - A)$  is a scalar multiple of  $\mathbf{1}$ , while each row of  $\text{adj}(I - A)$  is a scalar multiple of  $v^\top$ . Thus  $\text{adj}(I - A)$  is a scalar multiple of the matrix  $\mathbf{1}v^\top$ . We now find that  $v_j = \frac{\det(I - A_{(j)})}{\sum_{i=1}^n \det(I - A_{(i)})}$ ,  $j = 1, \dots, n$  where for each  $i = 1, \dots, n$ ,  $A_{(i)}$  is the principal submatrix of  $A$  formed by deleting the  $i$ -th row and column.

---

\*email: [stephen.kirkland@umanitoba.ca](mailto:stephen.kirkland@umanitoba.ca) .

From the all minors matrix tree theorem [2],  $\det(I - A_{(j)})$  can be written as  $\sum_{T \in \mathcal{T}_j} w(T)$ , where i)  $\mathcal{T}_j$  is the set of all spanning directed trees in the directed graph of  $A$  (where the loops are ignored) having vertex  $j$  as a *sink*, i.e. all arcs in the tree are oriented towards vertex  $j$ , and ii) for each such directed tree  $T$ , the *weight* of  $T$ ,  $w(T)$ , is the product of the entries in  $A$  that correspond to the arcs in  $T$ . Consequently,  $v_j = \frac{\sum_{T \in \mathcal{T}_j} w(T)}{\sum_{i=1}^n \sum_{T \in \mathcal{T}_i} w(T)}$ ,  $j = 1, \dots, n$ . This result, known as the Markov chain tree theorem [4], establishes a natural connection between the combinatorial structure of  $A$  and its stationary distribution vector.

The stationary distribution vector represents the long-term behaviour of the Markov chain associated with  $A$ . Indeed, when  $A$  is primitive, the iterates of the Markov chain converge to  $v$ , independently of the initial distribution for the chain. A standard measure of short-term behaviour is provided by the *mean first passage times*: given states  $j, k \in \{1, \dots, n\}$ , the mean first passage time from  $j$  to  $k$ ,  $m_{j,k}$ , is equal to the expected number of steps for the chain to enter state  $k$  for the first time, given that it started in state  $j$ . In particular, for each  $j = 1, \dots, n$ ,  $m_{jj}v_j = 1$ ; see [6] for example.

From a remarkable result due to Kemeny [6], it follows that if the eigenvalues of  $A$  are denoted by  $1 \equiv \lambda_1, \lambda_2, \dots, \lambda_n$ , then for any  $j \in \{1, \dots, n\}$ ,

$$\sum_{k=1}^n m_{jk}v_k = 1 + \sum_{\ell=2}^n \frac{1}{1 - \lambda_\ell}. \quad (1)$$

The expression  $\sum_{\ell=2}^n \frac{1}{1 - \lambda_\ell}$  is known as *Kemeny's constant* for the Markov chain with transition matrix  $A$ . Inspecting (1), we see that Kemeny's constant may be interpreted in terms of the expected number of transitions required to reach a randomly chosen destination state, starting from state  $j$ . In particular, (1) yields the surprising observation that  $\sum_{k=1}^n m_{jk}v_k$  does not depend on which initial state  $j$  is chosen, and that fact has given rise to the challenge of providing an intuitive explanation as to why this is the case [5, p. 469]. In [3], the author uses the maximum principle to explain why  $\sum_{k=1}^n m_{jk}v_k$  is independent of  $j$ , while [1] presents an argument based on the number of visits to particular states in order to explain that lack of dependence on  $j$ . In this note, we provide a combinatorial argument to prove that  $\sum_{k=1}^n m_{jk}v_k$  is independent of  $j$ .

## 2 Kemeny's constant via combinatorics

In this section we maintain the notation of section 1. Suppose that we have an index  $k \in \{1, \dots, n\}$ ; since  $A$  is irreducible,  $I - A_{(k)}$  is a nonsingular M-matrix (with positive determinant) from which it follows that  $(I - A_{(k)})^{-1}$  is an entrywise nonnegative matrix. Fix an index  $j \in \{1, \dots, n\}$ . For any index  $k$  with  $k \neq j$ , it is well-known that  $m_{jk} = e_j^\top (I - A_{(k)})^{-1} \mathbf{1}$ . Consequently,

$m_{jk} = \sum_{\ell=1}^{n-1} (I - A_{(k)})_{j\ell}^{-1}$ , and so we have

$$\sum_{k=1}^n m_{jk} v_k = 1 + \sum_{k=1}^n \left( \sum_{\ell=1}^{n-1} (I - A_{(k)})_{j\ell}^{-1} \right) v_k. \quad (2)$$

Next, we consider the adjoint formula for the inverse. For  $j, k, \ell$  as above, it follows that the  $(j, \ell)$  entry of  $(I - A_{(k)})^{-1}$  is equal to  $\frac{|\det((I-A)_{(\ell,k),(j,k)})|}{\det(I-A_{(k)})}$ , where for a square matrix  $X$ , the notation  $X_{(\ell,k),(j,k)}$  denotes the submatrix of  $X$  formed by deleting rows  $\ell$  and  $k$ , and columns  $j$  and  $k$ . We remark here that the absolute value in the numerator of that expression eliminates the need to keep track of the signs of the cofactors in the expression for the adjoint. Recalling the formula for  $v_k$  as a ratio of determinants, we thus find that

$$\left( \sum_{i=1}^n \det(I - A_{(i)}) \right) \left( \sum_{k=1}^n m_{jk} v_k - 1 \right) = \sum_{k=1}^n \left( \sum_{\ell \neq k} |\det((I - A)_{(\ell,k),(j,k)})| \right). \quad (3)$$

In view of (3), it suffices to show that the value of  $\sum_{k=1}^n \sum_{\ell \neq k} |\det((I - A)_{(\ell,k),(j,k)})|$  does not depend on the choice of  $j$ .

Fortunately, the all minors matrix tree theorem [2] provides a useful tool for analysing the expression of interest. Let  $\mathcal{F}$  denote the collection of spanning directed forests each of which is a union of exactly two directed trees, both of which have sinks. For each directed forest  $F \in \mathcal{F}$ , we let  $w(F)$  equal the product of the weights of the two directed trees of which  $F$  is the union (here we take the convention that if a tree has just one vertex, its weight is 1). We next identify some useful subsets of  $\mathcal{F}$ . For each  $k \in \{1, \dots, n\}$  we define  $\mathcal{S}_k^{(j)}$  to be the collection of directed forests in  $\mathcal{F}$  satisfying the following properties: a)  $k$  is the sink of one of the directed trees in  $F$ ; and b)  $j$  and  $k$  are in different directed trees in  $F$ . We note that necessarily  $\mathcal{S}_k^{(k)} = \emptyset$ . It now follows from the all minors matrix tree theorem that for any index  $\ell \neq k$ ,  $|\det((I - A)_{(\ell,k),(j,k)})|$  can be written as  $|\det((I - A)_{(\ell,k),(j,k)})| = \sum_{F \in \mathcal{S}_k^{(j,\ell)}} w(F)$ , where the sum is taken over the set  $\mathcal{S}_k^{(j,\ell)}$  consisting of the directed forests in  $\mathcal{S}_k^{(j)}$  such that  $\ell$  is the sink in the directed tree containing vertex  $j$ . We then find readily that

$$\sum_{\ell \neq k} |\det((I - A)_{(\ell,k),(j,k)})| = \sum_{F \in \mathcal{S}_k^{(j)}} w(F). \quad (4)$$

We claim that the sets of directed forests  $\mathcal{S}_1^{(j)}, \dots, \mathcal{S}_n^{(j)}$  comprise a partition  $\mathcal{F}$ . To see the claim, first fix a forest  $F \in \mathcal{F}$  and write it as a disjoint union of two directed trees,  $T_1$  and  $T_2$ . Since  $F = T_1 \cup T_2$  is spanning, vertex  $j$  is in one of  $T_1, T_2$ , and without loss of generality we take  $j \in T_1$ . Letting  $k_0$  denote the sink vertex of  $T_2$ , we see that necessarily,  $F \in \mathcal{S}_{k_0}^{(j)}$ . Hence  $\mathcal{F} \subseteq \cup_{k=1}^n \mathcal{S}_k^{(j)}$ . Further, for distinct indices  $k_1, k_2$ , we have  $\mathcal{S}_{k_1}^{(j)} \cap \mathcal{S}_{k_2}^{(j)} = \emptyset$ , since forests in  $\mathcal{S}_{k_1}^{(j)}$  have  $k_1$  as the sink in the directed tree not containing  $j$ , while forests in  $\mathcal{S}_{k_2}^{(j)}$  have  $k_2$

as the sink in the directed tree not containing  $j$ . Thus, the sets  $\mathcal{S}_1^{(j)}, \dots, \mathcal{S}_n^{(j)}$  partition  $\mathcal{F}$ , as claimed. Consequently,  $\sum_{k=1}^n (\sum_{\ell \neq k} |\det((I - A)_{(\ell, k), (j, k)})|) = \sum_{k=1}^n \sum_{F \in \mathcal{S}_k^{(j)}} w(F) = \sum_{F \in \mathcal{F}} w(F)$ . This last summation is clearly independent of  $j$ , as desired.

We now summarise the arguments above, maintaining our earlier notation.

**Theorem 1.** *For each  $j = 1, \dots, n$ ,*

$$\begin{aligned} \sum_{k=1}^n m_{jk} v_k &= 1 + \frac{\sum_{k=1}^n \sum_{F \in \mathcal{S}_k^{(j)}} w(F)}{\sum_{k=1}^n \sum_{T \in \mathcal{T}_k} w(T)} \\ &= 1 + \frac{\sum_{F \in \mathcal{F}} w(F)}{\sum_{k=1}^n \sum_{T \in \mathcal{T}_k} w(T)}. \end{aligned} \quad (5)$$

*In particular,  $\sum_{k=1}^n m_{jk} v_k$  does not depend on  $j$ .*

We remark that the equivalence of the first and last members of (5) is known (see [7]). The contribution of the present article is the equivalence of the first two members of (5), as well as the combinatorial argument establishing their independence on  $j$ .

**Acknowledgement** The author's research is supported in part by NSERC Discovery Grant RGPIN-2019-05408. The author is grateful to an anonymous referee, whose comments improved this article.

## References

- [1] D. Bini, J. Hunter, G. Latouche, B. Meini and P. Taylor, Why is Kemeny's constant a constant?, *J. Appl. Probab.* 55: 1025–1036, 2018.
- [2] S. Chaiken, A combinatorial proof of the all minors matrix tree theorem, *SIAM J. Alg. Disc. Meth.* 3: 319–329, 1982.
- [3] P. Doyle, The Kemeny constant of a Markov chain, *ArXiv e-prints*, arXiv:0909.2636v1[math.PR], 2009.
- [4] F. Leighton and R. Rivest, Estimating a probability using finite memory, *IEEE Trans. Information Theory* 32: 733–742, 1986.
- [5] C. Grinstead and J. Snell, *Introduction to Probability*, AMS, Providence, 1997.
- [6] J. Kemeny and J. Snell, *Finite Markov Chains*, Springer-Verlag, New York, 1976.
- [7] S. Kirkland and Z. Zeng, Kemeny's constant and an analogue of Braess' paradox for trees, *Electronic Journal of Linear Algebra* 31 (2016), 444–464.