Kemeny's Constant and the Random Surfer

Mark Levene and George Loizou
School of Computer Science and Information Systems
Birkbeck College, University of London
London WC1E 7HX, U.K.
{mark,george}@dcs.bbk.ac.uk

Abstract

We revisit Kemeny's constant in the context of Web navigation, also known as "surfing". We generalize the constant, derive upper and lower bounds on it, and give it a novel interpretation in terms of the number of links a random surfer will follow to reach his final destination.

1 Introduction

A Markov chain [11, Chapter 4] represents a stochastic process whereby the transition from one state to another depends only on the current state the process is in. The probabilities of a Markov chain are described using a transition matrix, which is nonnegative and row stochastic [5, Section 8.7]. Markov chains are used to model a wide number of physical phenomena [7, Chapter 7] and have been widely studied in the statistical literature since their inception by A.A. Markov in 1907. A Markov chain is *finite ergodic* if its transition matrix is finite, row stochastic, and primitive [5, Definition 8.5.0].

Kemeny's constant [7, Corollary 4.3.6] gives an interesting quantity for finite ergodic Markov chains. Kemeny [4, Section 11.5, Exercise 19] offered a prize for the first person to find an intuitive interpretation for his constant; as far as we know, to date no intuitive and straightforward explanation has yet been found.

The aim of this paper is twofold. First, we derive a general formula for Kemeny's constant involving the eigenvalues of the transition matrix that allows us to derive upper and lower bounds on the constant. Second, we give an intuitive and straightforward interpretation of Kemeny's constant in terms of user navigation through the World-Wide-Web.

We now briefly set the scene for our Markov chain model of the Web. One of the main activities of users interacting with the Web is that of *navigation* (colloquially known as "surfing") whereby users follow links and browse the destination Web pages. During the process of navigation users often experience disorientation and lose track of the context, causing them to be unsure how to proceed to satisfy their original goal; this problem is known as the *navigation problem* [9, 10]. Understanding user navigation patterns and their underlying distribution is important since it can lead to better Web site design [1] with the intention of saving users' effort by directly guiding them to the most relevant Web pages.

We view the Web as a finite irreducible Markov chain, where the probabilities attached to transitions denote the expected utility the user attains from following the corresponding links. This view is naturally realized by considering the user's home page as an artificial starting point for all navigation sessions and assuming that there is a positive probability (however small) of jumping to any other relevant Web page. During a navigation session the user follows links according to the transition probabilities and may eventually return to his home page. Thus a navigation session amounts to a random walk through the Markov chain that is terminated once the user reaches his final destination, i.e., he has no further incentive to continue surfing.

Imagine a random surfer who is following links according to the transition probabilities. At some stage our random surfer is "lost" and does not know the state he is at and where he is heading for. We show that in this context Kemeny's constant can be interpreted as the mean number of links the random surfer needs to follow before reaching his destination.

2 General Bounds on Kemeny's Constant

Let A be an $n \times n$ real or complex matrix. Assume that λ is a nonzero eigenvalue of A with algebraic multiplicity 1, and let $\lambda_2, \lambda_3, \ldots, \lambda_n$ denote the remaining eigenvalues of A including algebraic multiplicities. Let x and y^H, where y^H denotes the conjugate transpose of y, be right and left eigenvectors, respectively, corresponding to λ, so that $Ax = \lambda x \quad \text{and} \quad y^H A = \lambda y^H$. These two eigenvectors, which are unique up to a scalar multiple, are nonorthogonal, and can therefore be normalized so that $y^H x = 1$ [5, Lemma 6.3.10].

We next consider the matrix $A - \lambda x y^H$, which has the same Jordan canonical form as A except that its 1×1 block $[\lambda]$ is replaced by the 1×1 block $[0]$. Moreover, the eigenvalues of $\lambda I - (A - \lambda x y^H)$ are $\lambda, \lambda - \lambda_2, \lambda - \lambda_3, \ldots, \lambda - \lambda_n$, so they are all nonzero. The *fundamental matrix* associated with A and λ is

$$Z \equiv \Big(\lambda I - (A - \lambda x y^H)\Big)^{-1}. \tag{1}$$

The eigenvalues of Z are $\lambda^{-1}, (\lambda - \lambda_2)^{-1}, (\lambda - \lambda_3)^{-1}, \ldots, (\lambda - \lambda_n)^{-1}$, so

$$trZ = \frac{1}{\lambda} + \sum_{i=2}^{n} \frac{1}{\lambda - \lambda_i}. \tag{2}$$

Since x is an eigenvector of Z^{-1}, it is an eigenvector of Z, so

$$Zx = \frac{1}{\lambda} x \quad \text{and} \quad y^H Z x = \frac{1}{\lambda} y^H x = \frac{1}{\lambda}. \tag{3}$$

The number $K(A, \lambda) \equiv trZ - y^H Z x$ is known as *Kemeny's constant* [4, Section 11.5, Exercise 19]. Thus from (2) and (3) it follows that

$$K(A, \lambda) = \sum_{i=2}^{n} \frac{1}{\lambda - \lambda_i}. \tag{4}$$

From (4) it follows that

$$|K(A,\lambda)| \quad \leq \quad \frac{n-1}{|\lambda - \hat{\lambda}|}, \tag{5}$$

where $\hat{\lambda}$ denotes an eigenvalue of A that is closest to λ.

If A and λ are both real, then the right-hand side of (4) is also real; it is *positive* if $\lambda > Re\lambda_j$ for all $j = 2, 3, \ldots, n$. (This is the situation when λ is the Perron eigenvalue of an irreducible nonnegative matrix A.)

If all the eigenvalues of A are real, then from (4)

$$(n-1)\min\left\{\frac{1}{\lambda-\lambda_2}, \frac{1}{\lambda-\lambda_3}, \ldots, \frac{1}{\lambda-\lambda_n}\right\} \quad \leq \quad K(A,\lambda). \tag{6}$$

In order for A to possess a real spectrum, it is sufficient that there be a positive definite matrix Q such that $A^H = QAQ^{-1}$. In this event, we have $\mathcal{A} \equiv Q^{1/2}AQ^{-1/2} = Q^{-1/2}A^HQ^{1/2} = \mathcal{A}^H$, so $\mathcal{A}$ is Hermitian and similar to A. (We make use of this sufficient condition in the next section but only for a positive diagonal Q.)

Suppose that A and λ are both real, and $\lambda \geq |\lambda_j|$ for all $j = 2, 3, \ldots, n$. Then, since non-real eigenvalues occur in conjugate pairs, it follows from (4) that

$$0 < \frac{n-1}{2\lambda} \leq (n-1)\min_{2\leq j\leq n}\frac{\lambda - Re\lambda_j}{|\lambda-\lambda_j|^2} \leq K(A,\lambda). \tag{7}$$

This bound is valid, for example, if A is irreducible and nonnegative, and λ is its Perron eigenvalue. If all the eigenvalues of A are equimodulus, then the lower bound is attained.

If A is real, nonnegative, and irreducible, then its Perron eigenvalue [5, Theorem 8.4.4] λ is positive and algebraically simple, and thus A and λ satisfy (4) and (7). Moreover, there are positive left and right eigenvectors of A associated with λ. If, in addition, A is row stochastic, then (4) and (7) hold with $\lambda = 1$, and an associated eigenvector is h, the vector all of whose entries are 1.

3 Specific Bounds on Kemeny's Constant

Let P be the $n \times n$ transition matrix of a finite irreducible Markov chain. Then $\lambda = 1$ is an algebraically simple eigenvalue of P, $Ph = h$, and P has a positive left eigenvector $\pi = (\pi_1, \pi_2, \ldots, \pi_n)^T$ associated with λ, so that $\pi^TP = \pi^T$. If we normalize π by requiring that $\pi^Th = 1$, we obtain the (unique) *stationary probability vector* of P.

In [6, Theorem 1] the following result regarding a generalisation of the fundamental matrix of a finite irreducible Markov chain was presented.

Theorem 3.1 Let β and g be any two vectors such that β^Th and π^Tg are nonzero. Then $I - P + g\beta^T$ is nonsingular.

Proof: If $x \neq 0$ and $(I-P+g\beta^T)x = 0$, i.e., $I-P+g\beta^T$ is singular, then $(I-P)x = -(\beta^T x)g$, which yields $\pi^T(I-P)x = 0 = -(\beta^T x)\pi^T g$. Hence $\beta^T x = 0$ and so $(I-P)x = 0$. Therefore, $x = \gamma h$ (γ being a nonzero constant) whence $\beta^T h = 0$, since $\beta^T x = 0$. We could also show that $\pi^T g = 0$. $\square$

The matrix $Z \equiv (I - P + g\beta^T)^{-1}$ is an analogue of (1). For a finite irreducible Markov chain, we may take $\beta^T = \pi^T$ and $g = h$, in which case we have

$$Z \equiv (I - P + h\pi^T)^{-1} \tag{8}$$

with $\pi^T h = 1$, $Zh = h$ and $\pi^T Z = \pi^T$.

As a result of Theorem 3.1, (8) is only one of an infinite number of choices for Z, and a form of it that does not involve π can be obtained. For example, let $g = h$ and let β^T be any row vector such that $\beta^T h = 1$. Then

$$Z \equiv (I - P + h\beta^T)^{-1} \tag{9}$$

with $\beta^T h = 1$, $Zh = h$, and $\beta^T Z = \pi^T$.

Since β^T and g can be shown to be left and right eigenvectors of P associated with the algebraically simple eigenvalue $\lambda = 1$, it follows from (2) that all these Z's are associated with the same Kemeny constant, namely

$$K(P) = K(P, 1) = \sum_{i=2}^{n} \frac{1}{1 - \lambda_i}, \tag{10}$$

where λ_i is the ith eigenvalue of P.

A finite irreducible Markov chain is said to be reversible, if $\pi_i P_{ij} = \pi_j P_{ji}$ for all i, j, that is, P and the positive definite diagonal matrix $\Pi = diag(\pi_1, \pi_2, \ldots, \pi_n)$ satisfy $\Pi P = P^T \Pi$, so that $P^T = \Pi P \Pi^{-1}$. Since this condition ensures that all the eigenvalues of P are real, we can order them so that

$$-1 \leq \lambda_n \leq \lambda_{n-1} \leq \cdots \leq \lambda_2 < \lambda = 1. \tag{11}$$

If, additionally, the Markov chain is ergodic then $-1 < \lambda_n$ as well. (See [3], for example, for upper bounds on λ_2.) Thus, when a finite irreducible Markov chain is reversible, (10) and (11) yield

$$\frac{n-1}{2} \leq K(P) \leq \frac{n-1}{1 - \lambda_2},$$

and the lower bound is strict if the chain is ergodic.

4 An Interpretation of Kemeny's Constant

Let P be the transition matrix of a finite irreducible Markov chain. The expected number of steps required to reach state s_j, when starting a Markov chain from state s_i, is called the *mean first passage time* from s_i to s_j, and is denoted by m_{ij}; by convention $m_{ii} = 0$. (Since

P is irreducible $\pi_i \neq 0$ for all i.) Let $M = (m_{ij})$ be the *mean first passage matrix* [4, p. 455] for such a chain. Then, by [4, Theorem 11.16]

$$m_{ij} = \frac{z_{jj} - z_{ij}}{\pi_j}.$$

Since $Zh = h$ and $\pi^T h = 1$,

$$\begin{aligned} K(P) = trZ - \pi^T Zh &= \sum_{j=1}^{n} z_{jj} - 1 \\ &= \sum_{j=1}^{n} \left(\frac{z_{jj} - z_{ij}}{\pi_j} \right) \pi_j \\ &= \sum_{j=1}^{n} \pi_j m_{ij}. \quad \text{(See [4, p. 468].)} \end{aligned}$$

In order to give our interpretation of Kemeny's constant involving a random surfer, we can rewrite $K(P)$ as

$$K(P) = \sum_{i=1}^{n} \pi_i \sum_{j=1}^{n} \pi_j \ m_{ij},$$

due to the fact that $\pi^T h = 1$.

The term $\sum_j \pi_j \ m_{ij}$ gives the mean first passage time from state i when the destination state is unknown; denote this term by M_i. It follows that Kemeny's constant, given by $\sum_i \pi_i \ M_i$, can be interpreted as the mean first passage time from an unknown starting state to an unknown destination state.

Imagine therefore a random surfer who is following links according to the transition probabilities. At some stage our random surfer is "lost" and does not know the state he is at and where he is heading for. In this context Kemeny's constant can be interpreted as the mean number of links the random surfer follows before reaching his destination. Thus the random surfer is not "lost" anymore, he just has to follow $K(P)$ random links and he can *expect* to arrive at his final destination. It is therefore in this context that the upper and lower bounds for $K(P)$ in Sections 2 and 3 should be considered. For the Web, n is very large (greater than 10^9, see [2] or [8]) and therefore a random walk whose length is of the order of the bounds that we present is not feasible. However, if we can restrict the Web graph in some manner to an irreducible (or ergodic) subset (i.e, a Web subgraph that represents an irreducible (or ergodic) Markov chain), for example, by selecting only the pages that satisfy a user query, then the bounds we give may be useful.

References

[1] D.J. Aldous. Reorganizing large web sites. *American Mathematical Monthly*, 108:16–27, 2001.

[2] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, A. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener. Graph structure in the web. *Computer Networks and ISDN Systems*, 30:309–320, 2000.

[3] P. Diaconis and D. Stroock. Geometric bounds for eigenvalues of Markov chains. *Annals of Applied Probability*, 1:36–61, 1991.

[4] C.M. Grinstead and J.L. Snell. *Introduction to Probability.* American Mathematical Society, Providence, RI, 1997.

[5] R.A. Horn and C.R. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge, U.K., 1985.

[6] J.G. Kemeny. Generalization of a fundamental matrix. *Linear Algebra and its Applications*, 38:193–206, 1981.

[7] J.G. Kemeny and J.L. Snell. *Finite Markov Chains.* D. Van Nostrand, Princeton, NJ, 1960.

[8] S. Lawrence and C.L. Giles. Accessibility of information on the web. *Nature*, 400:107–109, 1999.

[9] M. Levene and G. Loizou. Navigation in hypertext is easy only sometimes. *SIAM Journal on Computing*, 29:728–760, 1999.

[10] M. Levene and G. Loizou. A probabilistic approach to navigation in hypertext. *Information Sciences*, 114:165–186, 1999.

[11] S.M. Ross. *Stochastic Processes.* John Wiley & Sons, New York, NY, 2nd edition, 1996.