THE ROLE OF KEMENY'S CONSTANT IN PROPERTIES OF MARKOV CHAINS

JEFFREY J HUNTER

SCHOOL OF COMPUTING & MATHEMATICAL SCIENCES

AUT UNIVERSITY
TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

Andrei A Markov (1856 – 1922)

John G Kemeny (1926 – 1992)

Outline

1. Preliminaries
2. Kemeny's constant
3. Expected time to mixing
4. Random surfer
5. Examples
6. Perturbation results
7. Mixing on directed graphs
8. Kirchhoff index
9. Variances of mixing times

Introduction

Let $\{X_n\}$, $(n \geq 0)$ be a finite irreducible (ergodic), discrete time Markov chain (MC).

Let $S = \{1, 2, \ldots, m\}$ be its state space.

Let $p_{ij} = P[X_{n+1} = j \mid X_n = i]$ be the transition probability from state i to state j.

Let $P = [p_{ij}]$ be the transition matrix of the MC.

P stochastic $\Rightarrow \sum_{j=1}^{m} p_{ij} = 1, \; i \in S.$

Let $\{p_j^{(n)}\} = \{P[X_n = j]\}$ be the probability distribution at the n-th trial.

Limiting & Stationary Distribns

When the MC is ***regular*** (finite, aperiodic & irreducible) a limiting distribution exists, that does not depend on the initial distribution and that the limiting distribution is the stationary distribution. ie. $\{X_n\}$ has a unique stationary distribution $\{\pi_j\}, j \in S$ and $\lim_{n \to \infty} p_j^{(n)} = \pi_j$.

When the MC is finite, irreducible and ***periodic*** a limiting distribution does not exist. However there is a unique stationary distribution.

Stationary Distributions

Irreducible or ***ergodic*** MCs $\{X_n\}$ have a unique stationary distribution $\{\pi_j\}, j \in S$.

The stationary probabilities are given as the solution of the stationary equations:

$$\pi_j = \sum_{i=1}^{m} \pi_i p_{ij} \ (j \in S) \text{ with } \sum_{i=1}^{m} \pi_i = 1.$$

The "stationary probability vector" is $\boldsymbol{\pi}^T = (\pi_1, \pi_2, \ldots, \pi_m)$.

Primer on g-inverses of $I - P$

A 'one condition' g-inverse or an 'equation solving' g- inverse of a matrix A is any matrix A^- such that $AA^-A = A$.

Let P be the transition matrix of a finite irreducible MC with stationary probability vector $\boldsymbol{\pi}^T$. Let $\boldsymbol{t}$ and $\boldsymbol{u}$ be any vectors. Let $\boldsymbol{e}^T = (1,1, \ldots, 1)$.

$I - P + \boldsymbol{t}\boldsymbol{u}^T$ is non-singular $\Leftrightarrow$ $\boldsymbol{\pi}^T\boldsymbol{t} \neq 0$ and $\boldsymbol{u}^T\boldsymbol{e} \neq 0$.

$\boldsymbol{\pi}^T\boldsymbol{t} \neq 0$ and $\boldsymbol{u}^T\boldsymbol{e} \neq 0 \Rightarrow$ $[I - P + \boldsymbol{t}\boldsymbol{u}^T]^{-1}$ is a g-inverse of $I - P$.

(Hunter, 1982)

Use of g-inverses

A necessary and sufficient condition for $AXB = C$ to have a solution is that $AA^{-}CB^{-}B = C$.

If this consistency condition is satisfied the general solution is given by $X = A^{-}CB^{-} + W - A^{-}AWBB^{-}$, where W is an arbitrary matrix. (Rao,1966)

$AX = C$ has a solution $X = A^{-}C + (I - A^{-}A)W$, where W is arbitrary, provided $AA^{-}C = C$.

Special g-inverses of $I - P$

If G is any g-inverse of $I - P$ then there exists vectors $\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{t}$ and $\boldsymbol{u}$ with $\boldsymbol{\pi}^T\boldsymbol{t} \neq 0$ and $\boldsymbol{u}^T\boldsymbol{e} \neq 0$ such that

$$G = [I - P + \boldsymbol{t}\boldsymbol{u}^T]^{-1} + \boldsymbol{e}\boldsymbol{f}^T + \boldsymbol{g}\boldsymbol{\pi}^T.$$

$Z = [I - P + \Pi]^{-1}$, $(\Pi \equiv \boldsymbol{e}\boldsymbol{\pi}^T)$ "fundamental matrix" of irreducible (ergodic) Markov chains. (Kemeny & Snell, 1960)

$(I - P)^{\#} = A^{\#} = Z - \Pi$, "group inverse" of $I - P$. (Meyer, 1975)

If G is any generalized inverse of $I - P$, $(I - P)G(I - P)$ is invariant and $= A^{\#}$.
(Meyer, 1975), (Hunter, 1982)

First Passage Times in MCs

Let T_{ij} be the first passage time r.v. from state i to state j,

i.e. $T_{ij} = \min\{n \geq 1$ such that $X_n = j$ given that $X_0 = i\}$,

T_{ii} is the "first return to state i".

The irreducibility of the MC ensures that the T_{ij} are all proper random variables. Under the finite state space restriction, all the moments of T_{ij} are finite.

Let m_{ij} be the mean first passage time from state i to state j.

i.e. $m_{ij} = \mathrm{E}[T_{ij} \mid \mathrm{X}_0 = i]$ for all $(i, j) \in S \times S$.

Mean First Passage Times

For an irreducible finite MC with transition matrix P,

let $M = \left[m_{ij}\right]$ be the matrix of expected first passage

times from state i to state j.

M satisfies the matrix equation

$$(I - P)M = E - PM_d$$

where $E = \mathbf{ee}^T = [1]$, $M_d = [\delta_{ij} m_{ij}] = (\Pi_d)^{-1} \equiv$ D.

Mean first passage times

If G is any g-inverse of $I - P$, then

$$M = [G\Pi - E(G\Pi)_d + I - G + EG_d]D. \qquad \text{(Hunter, 1982)}$$

Under any of the following three equivalent conditions:

(i) $G\boldsymbol{e} = g\boldsymbol{e}$, g a constant,

(ii) $GE - E(G\Pi)_d D = 0$,

(iii) $G\Pi - E(G\Pi)_d = 0$,

$$M = [I - G + EG_d]D. \qquad \text{(Hunter, 2008)}$$

Special cases:

$G = Z$, Kemeny and Snell's fundamental matrix (g = 1)

$G = A^{\#} = Z - \Pi$, Meyer's group inverse of $I - P$, (g = 0)

Mean first passage times

If $G = [\, g_{ij}]$ is any generalized inverse of $I - P$,

then $m_{ij} = \left(\dfrac{g_{jj} - g_{ij} + \delta_{ij}}{\pi_j} \right) + (g_{i\bullet} - g_{j\bullet})$, for all i, j.

$\boldsymbol{Ge} = g\boldsymbol{e} \Rightarrow m_{ij} = \left(\dfrac{g_{jj} - g_{ij} + \delta_{ij}}{\pi_j} \right)$, for all i, j.

$$\text{Thus } m_{ij} = \begin{cases} \dfrac{z_{jj} - z_{ij}}{\pi_j} = \dfrac{a^{\#}_{jj} - a^{\#}_{ij}}{\pi_j}, & i \neq j, \\[2ex] \dfrac{1}{\pi_j} & i = j. \end{cases}$$

where Z = [z_{ij}] (Kemeny & Snell, 1960), A$^{\#} = [a^{\#}_{ij}]$ (Meyer, 1975)

Kemeny's constant

Key Result : For all $i \in S$,

$$\sum_{j=1}^{m} m_{ij}\pi_j = K, \text{ "}\textbf{\textit{Kemeny's constant}}\text{".}$$

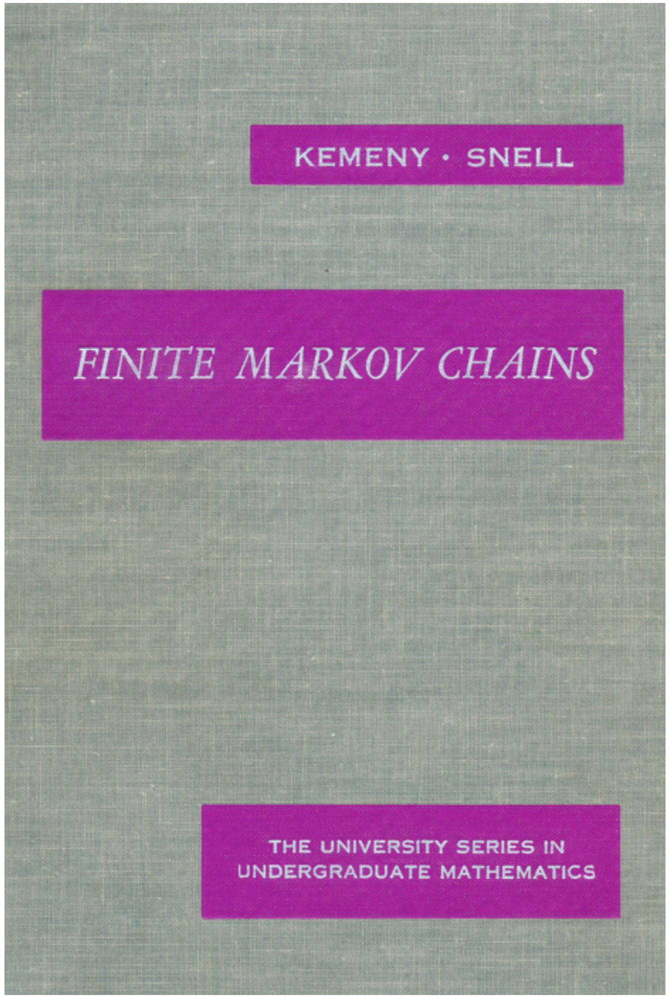
Equivalently, $\quad M\boldsymbol{\pi} = K\mathbf{e}.$

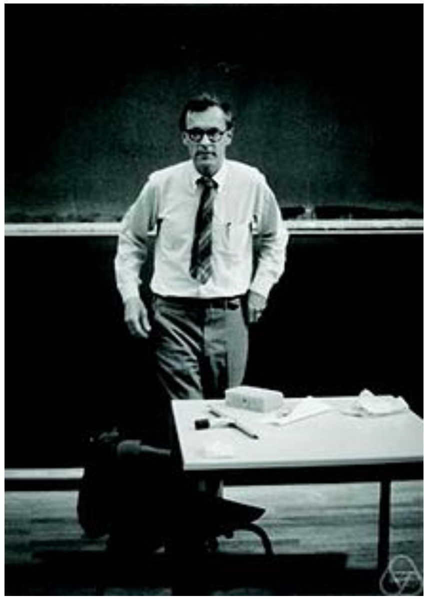
One of the simplest proofs is based upon Z :

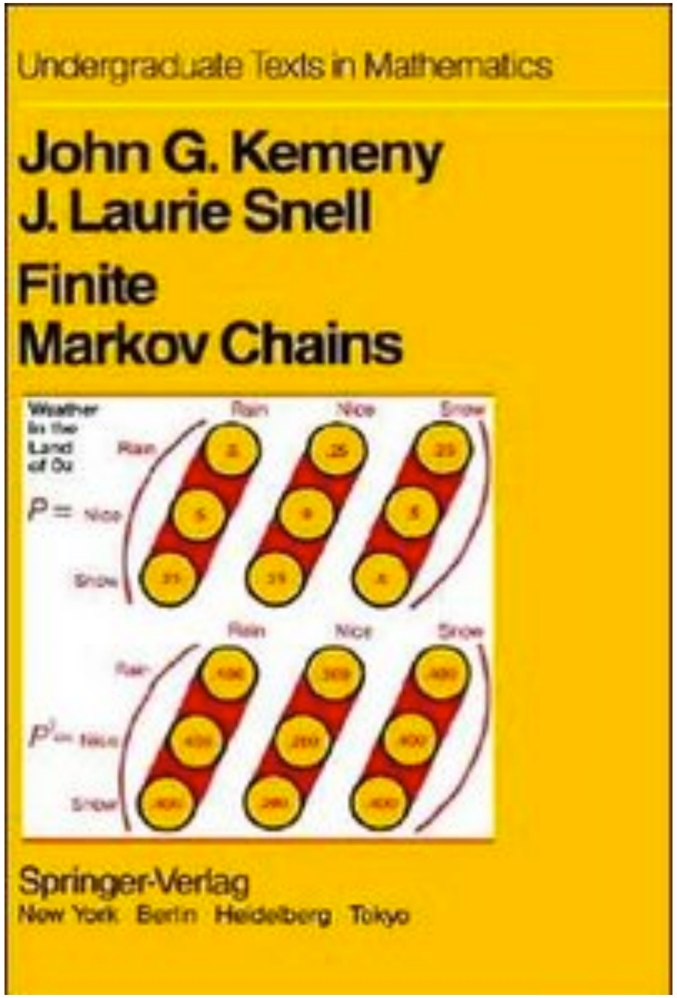
$$\begin{aligned} M\boldsymbol{\pi} &= [I - Z + EZ_d]D\boldsymbol{\pi} \\ &= [I - Z + EZ_d]\mathbf{e} \\ &= \mathbf{e} - Z\mathbf{e} + \mathbf{e}\mathbf{e}^T Z_d \mathbf{e} = K\mathbf{e}, \end{aligned}$$

where $K = \mathbf{e}^T Z_d \mathbf{e} = tr(Z)$.

Initial appearance - 1960







Kemeny & Snell - Initial result

4.4.10 THEOREM. *Let* $c = \sum_i z_{ii}$. *Then* $M\alpha^T = c\xi$.

PROOF.

$$\begin{aligned} M\alpha^T &= (I - Z + EZ_{\mathrm{dg}})D\alpha^T \\ &= (I - Z + EZ_{\mathrm{dg}})\xi \\ &= \xi(\eta Z_{\mathrm{dg}}\xi) = c\xi. \end{aligned}$$

In terms of our notation: $c = tr(Z)$, $\alpha^T = \boldsymbol{\pi}, \eta = \mathbf{e}^T, \xi = \mathbf{e}$ so that

$$M\boldsymbol{\pi} = (tr(Z))\mathbf{e}.$$

(Kemeny & Snell, “Finite Markov Chains”,1960)

Kemeny's constant

Define $\boldsymbol{k} = M\boldsymbol{\pi}$, where $\boldsymbol{k}^T = (K_1, K_2, \ldots., K_m)$.

Since $(I - P)M = E - PM_d$,

$(I - P)\boldsymbol{k} = (I - P)M\boldsymbol{\pi} = E\boldsymbol{\pi} - PM_d\boldsymbol{\pi} = \mathbf{e}\mathbf{e}^T\boldsymbol{\pi} - P\mathbf{e} = \mathbf{e} - \mathbf{e} = \mathbf{0}$.

i.e. $\quad P\boldsymbol{k} = \boldsymbol{k}, \quad$ or $\quad \sum_{j=1}^{m} p_{ij} K_j = K_i$

The irreducubility of the MC implies that $\boldsymbol{k}$ is the right eigenvector of P corresponding to the eigenvalue $\lambda = 1$

$\Rightarrow k = Ke.\ \ i.e\ K_i = K\ for\ all\ i\ =\ 1,\ 2,\ \ldots,\ m.$

$i.e.\ K_i = \sum_{j=1}^{m} m_{ij}\pi_j = K$, "***Kemeny's constant***" for all $i \in S$.

Kemeny's *K* - Clarification

Note that m_{ii} is the "mean recurrence time for state i".

It is well known that $m_{ii} = 1/\pi_i$ and thus $m_{ii}\pi_i = 1$.

Consequently "***Kemeny's constant***"

$$K = \sum_{j=1}^{m} m_{ij}\pi_j = m_{ii}\pi_i + \sum_{j \neq i} m_{ij}\pi_j = 1 + \sum_{j \neq i} m_{ij}\pi_j.$$

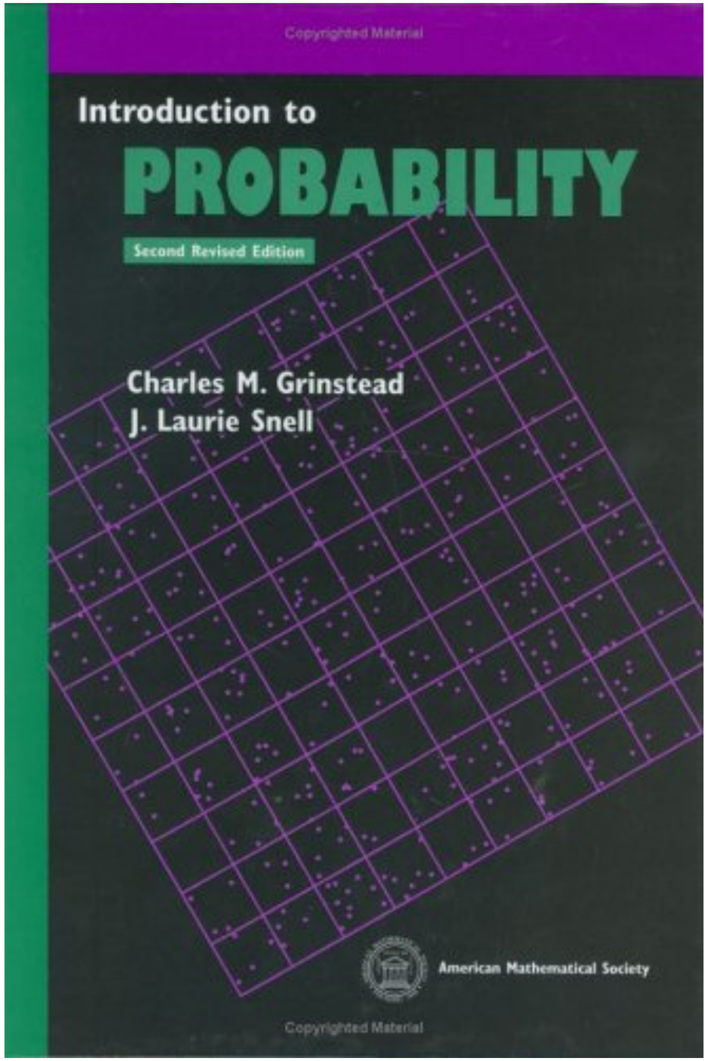
Some authors define, by convention, that $m_{ii} = 0$

so that the expression for the mean first passage times

taken as $m_{ij} = (z_{jj} - z_{ij})/\pi_j$ holds for all i, j.

We will stay with the expression as defined above for K, bearing in mind that in some books and papers K is replaced by $K - 1$.

Grinstead & Snell - 2006 - Update



Grinstead & Snell - Update

19 Show that, for an ergodic Markov chain (see Theorem 11.16),

$$\sum_j m_{ij} w_j = \sum_j z_{jj} - 1 = K \ .$$

By convention $m_{ii} = 0$.

The second expression above shows that the number K is independent of i. The number K is called *Kemeny's constant*. A prize was offered to the first person to give an intuitively plausible reason for the above sum to be independent of i. (See also Exercise 24.)

Grinstead & Snell - Update

24 In the course of a walk with Snell along Minnehaha Avenue in Minneapolis in the fall of 1983, Peter Doyle[25] suggested the following explanation for the constancy of *Kemeny's constant* (see Exercise 19). Choose a target state according to the fixed vector $\mathbf{w}$. Start from state i and wait until the time T that the target state occurs for the first time. Let K_i be the expected value of T. Observe that

$$K_i + w_i \cdot 1/w_i = \sum_j P_{ij} K_j + 1 \ ,$$

and hence

$$K_i = \sum_j P_{ij} K_j \ .$$

By the maximum principle, K_i is a constant. Should Peter have been given the prize?

Peter Doyle – 2009 - Update

The Kemeny constant of a Markov chain

Peter Doyle

Version 1.0 dated 14 September 2009
GNU FDL*

$$M_{iw} = \sum_j P_i^{\ j} M_{jw}.$$

But now by the familiar *maximum principle*, any function f_i satisfying

$$\sum_j P_i^{\ j} f_j = f_i$$

must be constant: Choose i to maximize f_i, and observe that the maximum must be attained also for any j where $P_i^{\ j} > 0$; push the max around until it is attained everywhere. So M_{iw} doesn't depend on i. ∎

Note. The application of the maximum principle we've made here shows that the only column eigenvectors having eigenvalue 1 for the matrix P are the constant vectors—a fact that was stated not quite explicitly above.

This formula provides a computational verification that Kemeny's constant is constant, but doesn't explain *why* it is constant. Kemeny felt this keenly: A prize was offered for a more 'conceptual' proof, and awarded—rightly or wrongly—on the basis of the maximum principle argument outlined above.

Kemeny's constant: G-inverses

If $G = [g_{ij}]$ is any g-inverse of $I - P$, then

$$K = 1 + tr(G) - tr(G\Pi) = 1 + \sum_{j=1}^{m} (g_{jj} - g_{j.}\pi_j).$$

When $G\mathbf{e} = g\mathbf{e}$,

$$K = 1 - g + tr(G) = 1 - g + \sum_{j=1}^{m} g_{jj}.$$

In particular, $K = tr(Z) = \sum_{j=1}^{m} z_{jj}$

and $K = 1 + tr(A^{\#})$.

"Classical result" (Hunter, 2006).

"Random target lemma" (with Z) (Lovasz &Winkler,1998).

Book "Reversible MCs & RWs" (Aldous & Fill, 1999).

Kemeny's constant: Eigenvalues

P irreducible $\Rightarrow$

The eigenvalues of $P, \{\lambda_i\}$ $(i = 1, 2, \ldots, m)$

are such that $\lambda_1 = 1$, with $|\lambda_i| \le 1$ and $\lambda_i \ne 1$ $(i = 2, \ldots, m)$.

$\Rightarrow$ The eigenvalues of $Z = [z_{ij}] = [I - P + \mathbf{e}\pi^T]^{-1}$ are

$$\lambda_i(Z) = 1 \ (i = 1), \ \frac{1}{1 - \lambda_i} \ (i = 2, \ldots, m).$$

Thus $K = tr(Z) = \sum_{i=1}^{m} z_{ii} = \sum_{i=1}^{m} \lambda_i(Z) = 1 + \sum_{i=2}^{m} \frac{1}{1 - \lambda_i}.$

(Levene & Loizou, 2002), (Hunter, 2006), (Doyle, 2009)

Kemeny's constant: Bounds

$K = 1 + \sum_{i=2}^{m} \frac{1}{1-\lambda_i}$ and P is irreducible.

Hence $\lambda_1 = 1$, with $|\lambda_i| \le 1$ and $\lambda_i \ne 1$ $(i = 2, \ldots, m)$.

If any eigenvalue appears on the unit circle $|\lambda| = 1$ must appear as a root of unity and be associated with a periodic chain (whose periodicity cannot exceed m).

Any complex root $\lambda = a + bi$ must be associated with its complex conjugate $\bar{\lambda} = a - bi$, with $a^2 + b^2 \le 1$.

For this pair of conjugate roots

$$\frac{1}{1-\lambda} + \frac{1}{1-\bar{\lambda}} = \frac{2-(\lambda+\bar{\lambda})}{(1-\lambda)(1-\bar{\lambda})} = \frac{2-2a}{1-(\lambda+\bar{\lambda})+\lambda\bar{\lambda}} = \frac{2-2a}{1-2a+a^2+b^2} \ge 1 .$$

Bounds on K

For conjugate pair of roots $\frac{1}{1-\lambda}+\frac{1}{1-\bar{\lambda}} \geq 1$. For any real roots, $-1 \leq \lambda \leq 1 \Rightarrow \frac{1}{1-\lambda} \geq \frac{1}{2}$. The only possible root at $\lambda = -1$ occurs with a periodic MC with even period. Thus taking the real roots individually and complex roots in pairs

$$K = 1 + \sum_{i=2}^{m} \frac{1}{1-\lambda_i} \geq 1 + \frac{m-1}{2} = \frac{m+1}{2}.$$

(Hunter(2006)) Proof based on results of Styan (1964) with λ_i real.

If the **MC is reversible** (all the λ_i real) and regular (aperiodic)

then $\frac{m-1}{2} \leq \sum_{i=2}^{m} \frac{1}{1-\lambda_i} \leq \frac{m-1}{1-\lambda_2}$. (Levene & Loizou, 2002).

Improved Bounds on K

Suppose the the MC is irreducible & reversible so that

$1 = \lambda_1 > \lambda_2 \geq \ldots \geq \lambda_m > -1$. Note $K = 1 + \sum_{i=2}^{m} \frac{1}{1-\lambda_i} = m + \sum_{i=2}^{m} \frac{\lambda_i}{1-\lambda_i}$

Apply the method of Lagrange multipliers to the function

$f(x_2, \ldots, x_m) = \sum_{i=2}^{m} \frac{x_i}{1-x_i},$

subject to $1 + x_2 + \ldots + x_m = 0$ on the domain $1 > x_2 \geq \ldots \geq x_m > -1$

$\Rightarrow$ minimum of $f(x_1, x_2, \ldots, x_m)$ attained at $x_2 = .. = x_m = \frac{-1}{m-1}$.

$\Rightarrow \frac{(m-1)^2}{m} \leq \sum_{i=2}^{m} \frac{1}{1-\lambda_i} \leq \frac{m-1}{1-\lambda_2}$. (Palocois & Remon, 2010).

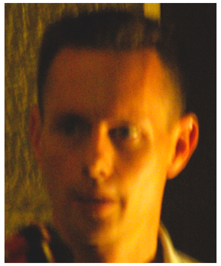
- an improvement on the earlier bounds of Levene & Loizoiu).

Alternative representation of *K*

$$K = tr(A_j^{-1}) - \frac{A_{jj}^{\#}}{\pi_j} + 1,$$

where A_j^{-1} is $(m-1)\times(m-1)$ principal submatrix of $A = I - P$ obtained by deleting j – th row and column.

(Catral, Kirkland, Neumann, Sze, 2010)



The proof is based upon expressing $A^{\#} = [a_{ij}^{\#}]$ in terms of A_n^{-1} and $\boldsymbol{\pi}^T$

Without loss of generality, take $j = m$. Use $m_{ij}\pi_j = a_{jj}^{\#} - a_{ij}^{\#}$

and the result (Meyer, 1973) that if B is the leading $(m-1)\times(m-1)$ principal submatrix of $A^{\#}$, then $B = A^{-1}{}_n + \beta W - A_n^{-1}W - WA_n^{-1}$,

where $\beta = \boldsymbol{u}^T A_n^{-1}\mathbf{e}$, $W = \mathbf{e}\boldsymbol{u}^T$ and $\boldsymbol{\pi}^T = (\boldsymbol{u}^T, \pi_n)$.

Stationarity in Markov chains

For all irreducible MCs (including periodic chains),

if for some $k \geq 0$, $p_j^{(k)} = \mathrm{P}[X_k = j] = \pi_j$ for all $j \in S$,

then $p_j^{(n)} = \mathrm{P}[X_n = j] = \pi_j$ for all $n \geq k$ and all $j \in S$.

How many trials do we need to take so that

$P[X_n = j\ ] = \pi_j$ for all $j \in S$?

Mixing Times in Markov chains

Let Y be a RV whose probability distribution is the stationary distribution $\{\pi_j\}$.

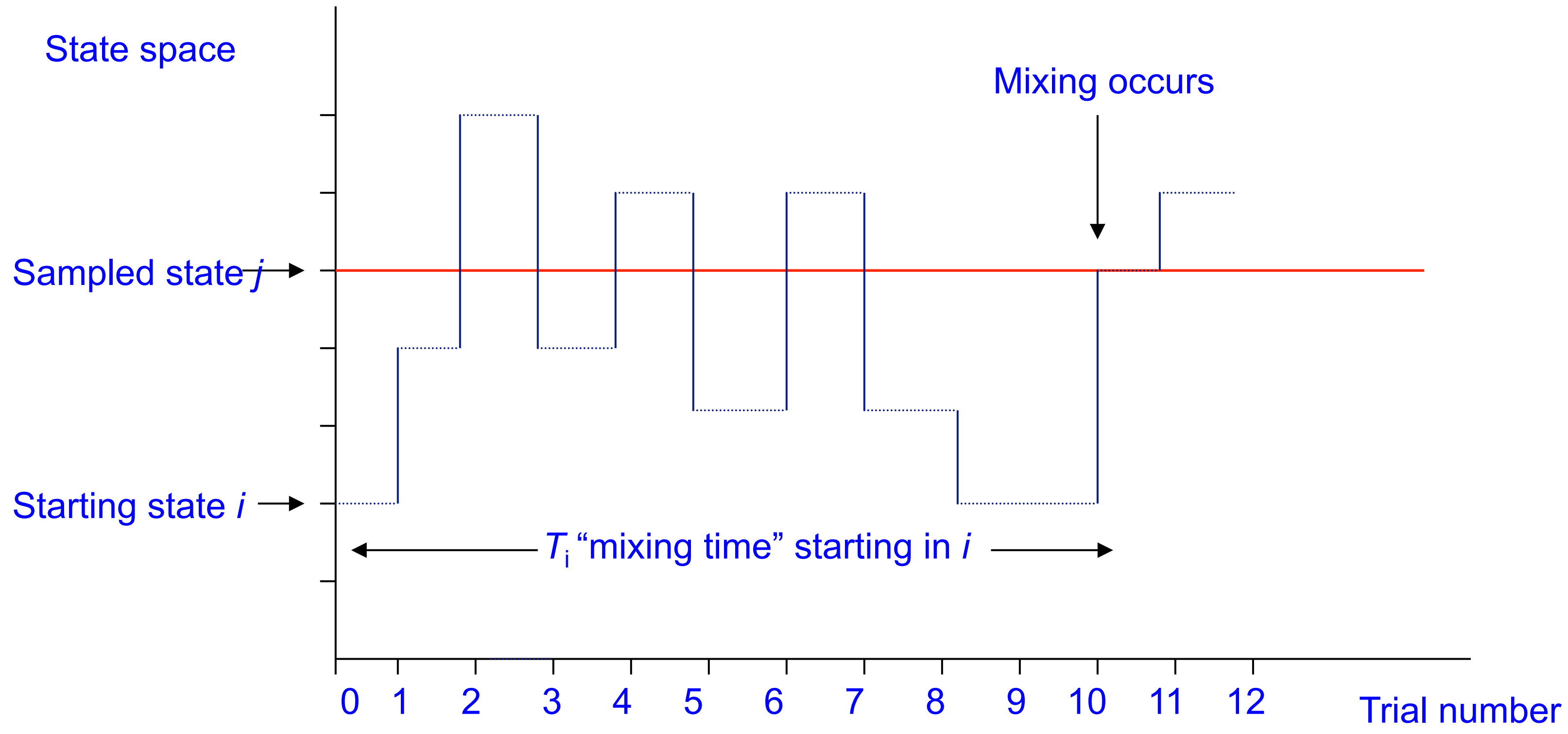
The MC $\{X_n\}$, achieves "mixing", at time $T = k$, when $X_k = Y$ for the smallest such $k \geq 1$.

T is the "time to mixing" in a Markov chain.

Thus, we first sample from the stationary distribution $\{\pi_j\}$ to determine a value of the random variable Y, say $Y = j$.

Now observe the MC, starting at a given state i. We achieve "mixing" at time $T = n$ when $X_n = j$ for the first such $n \geq 1$.

Expected time to Mixing



Expected Time to Mixing

The finite state space & irreducibility of the X_n

$\Rightarrow$ T is finite (a.s), with finite moments.

Let $\tau_{M,i}$ be the "*expected time to mixing*", starting at state i,

(assuming that mixing cannot occur at the first trial).

Conditional upon $X_0 = i$,

$$E[T] = E_Y(E[T \mid Y]) = \sum_{j=1}^{m} E[T \mid Y = j]P[Y = j]$$

$$= \sum_{j=1}^{m} E[T_{ij} \mid X_0 = i]\pi_j = \sum_{j=i}^{m} m_{ij}\pi_j$$

i.e. $\tau_{M,i} = E[T \mid X_0 = i] = \sum_{j=i}^{m} m_{ij}\pi_j = \sum_{j=1}^{m} m_{ij}\pi_j = \tau_M = K.$

i.e. Expected time to mixing, starting in any state, is K.

(Hunter, 2006)

Mixing or Hitting Times

Suppose the sampled stationary state ("mixing state") is j and the initial "starting state" is i.

We have assumed that the MC $\{X_n\}$, achieves "mixing", at time $T = k$, when $X_k = Y$ for the smallest such $k \geq 1$.

Suppose however we allow mixing to be possible when $k = 0$ when i = j. i.e. we permit "mixing" to occur at time $T = 0$, when state i is the "hitting" state (rather than "returned state")

The expected time to mixing in this situation would be

$\sum_{j \neq i} m_{ij}\pi_j = K - 1$, since $m_{ii}\pi_i = 1$.

(Hunter - 2010 preprint - considers the distribution of the time to mixing and time to hitting in each of the above situations.)

Random surfer

Note that $K = \sum_{i=1}^{m} \pi_i \sum_{j=1}^{m} \pi_j m_{ij} = \sum_{i=1}^{m} \pi_i M_i$ where $M_i = \sum_{j=1}^{m} \pi_j m_{ij}$.

M_i can represent the mean first passage time from state i when the destination state is unknown.

$K = \sum_{i=1}^{m} \pi_i M_i$ can be interpreted as the mean first passage time from an unknown starting state to an unknown destination state.

Imagine a random surfer who is "lost" and doesnt know the state he is at and where he is heading.

K can be intrepeted 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 random links and he can expect to arrive at his final destination. (Levene & Loizou, 2002)

Ex: Two state Markov Chains

Let $P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$,

$(0 \le a \le 1,\ 0 \le b \le 1)$. Let $d = 1 - a - b$.

MC irreducible $\Leftrightarrow -1 \le d < 1$.

MC has a unique stationary probability vector

$$\boldsymbol{\pi}^T = (\pi_1, \pi_2) = \left(\frac{b}{a+b}, \frac{a}{a+b} \right) = \left(\frac{b}{1-d}, \frac{a}{1-d} \right).$$

$-1 < d < 1 \Leftrightarrow$ MC is regular and the stationary distribution is the limiting distribution of the MC.

$d = -1 \Leftrightarrow$ MC is irreducible, periodic, period 2.

Ex: Two state Markov Chains

$$K = 1 + \frac{1}{a+b} = 1 + \frac{1}{1-d}.$$

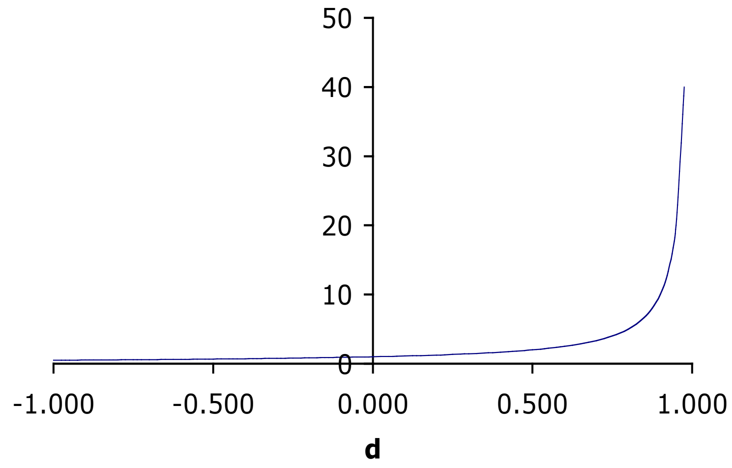
$d = 1 \Leftrightarrow$ Periodic, period 2, MC with $a = 1,\ b = 1.$

$\Leftrightarrow K = 1.5$ (minimum value of K).

$d = 0 \Leftrightarrow$ Independent trials $\Leftrightarrow K = 2.$

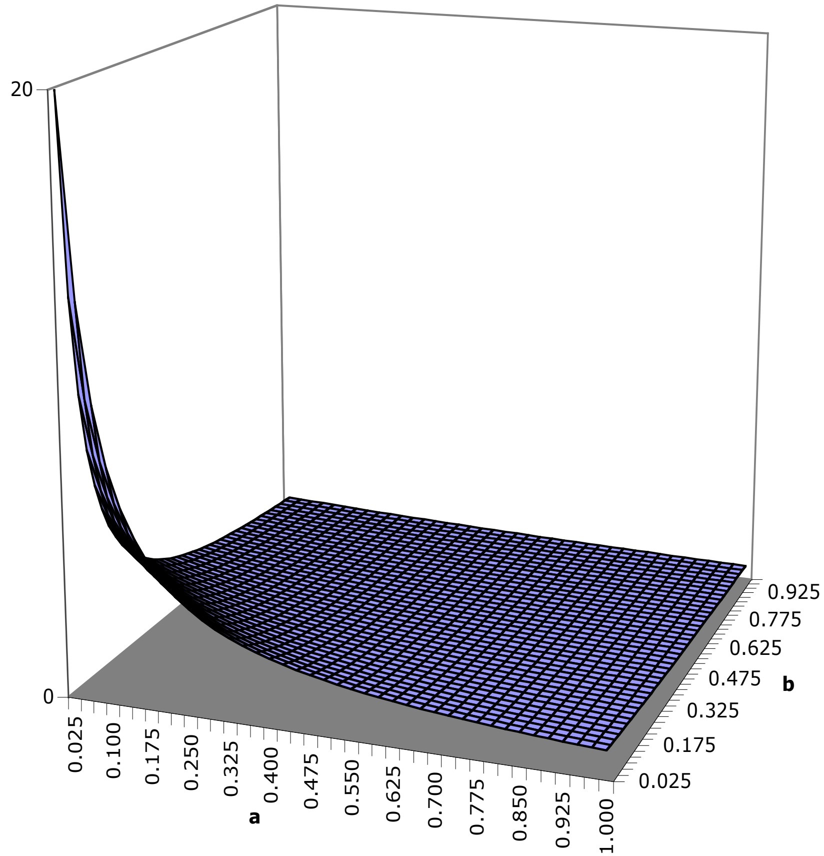
$d \to 1$ (both a $\to 0$ and b $\to 0$) $\Rightarrow$ arbitrarily large K.

For all two state MCs: $1.5 \le K < \infty$



Ex: Two state Markov Chains

Plot of $K = 1 + \dfrac{1}{a+b}$.



Ex: Three state Markov Chains

$$P = \left[p_{ij} \right] = \begin{bmatrix} 1 - p_2 - p_3 & p_2 & p_3 \\ q_1 & 1 - q_1 - q_3 & q_3 \\ r_1 & r_2 & 1 - r_1 - r_2 \end{bmatrix}.$$

Six constrained parameters with

$0 < p_2 + p_3 \leq 1,\ 0 < q_1 + q_3 \leq 1$ and $0 < r_1 + r_2 \leq 1.$

Let $\Delta_1 \equiv q_3 r_1 + q_1 r_2 + q_1 r_1,$

$$\Delta_2 \equiv r_1 p_2 + r_2 p_3 + r_2 p_2,$$

$$\Delta_3 \equiv p_2 q_3 + p_3 q_1 + p_3 q_3,$$

$$\Delta \equiv \Delta_1 + \Delta_2 + \Delta_3.$$

Ex: Three state Markov Chains

MC is irreducible

(and hence a stationary distribution exists)

$\Leftrightarrow \quad \Delta_1 > 0,\ \Delta_2 > 0,\ \Delta_3 > 0.$

Stationary distribution given by

$$(\pi_1, \pi_2,\ \pi_3) = \frac{1}{\Delta}(\Delta_1, \Delta_2, \Delta_3).$$

Ex: Three state Markov Chains

Let $\tau_{12} = p_3 + r_1 + r_2,\ \tau_{13} = p_2 + q_1 + q_3, \tau_{21} = q_3 + r_1 + r_2,$

$\tau_{23} = q_1 + p_2 + p_3,\ \tau_{31} = r_2 + q_1 + q_3,\ \tau_{32} = r_1 + p_2 + p_3,$

Let $\tau = p_2 + p_3 + q_1 + q_3 + r_1 + r_2$

$\Rightarrow \quad \tau = \tau_{12} + \tau_{13} = \tau_{21} + \tau_{23} = \tau_{31} + \tau_{32}.$

$$M = \begin{bmatrix} \Delta/\Delta_1 & \tau_{12}/\Delta_2 & \tau_{13}/\Delta_3 \\ \tau_{21}/\Delta_1 & \Delta/\Delta_2 & \tau_{23}/\Delta_3 \\ \tau_{31}/\Delta_1 & \tau_{32}/\Delta_2 & \Delta/\Delta_3 \end{bmatrix}$$

Ex: Three state Markov Chains

Kemeny's constant: $K = 1 + \dfrac{\tau}{\Delta}$

For all three-state irreducible MCs, $K \geq 2$.

$K = 2$ achieved in "the minimal period 3" case

when $p_2 = q_3 = r_1$, i.e. when $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$.

Ex: Three state Markov Chains

"Period-2 case": Transitions between {1,3} and {2}

$$P = \begin{bmatrix} 0 & 1 & 0 \\ q_1 & 0 & q_3 \\ 0 & 1 & 0 \end{bmatrix}, \ (q_1 + q_3 = 1) \Rightarrow K = 2.5$$

"Constant movement" case:

$$P = \begin{bmatrix} 0 & p_2 & p_3 \\ q_1 & 0 & q_3 \\ r_1 & r_2 & 0 \end{bmatrix}, \ (p_2 + p_3 = q_1 + q_3 = r_1 + r_2 = 1)$$

$$K = 1 + \frac{3}{3 - q_3 r_2 - r_1 p_3 - p_2 q_1} \Rightarrow 2 \le K \le 2.5$$

General *m* - state MCs

Periodic, period-m chain $K = \dfrac{m+1}{2}$.

Independent trials with m possible outcomes: $K = m$.

For all irreducible m - state MCs: $\dfrac{m+1}{2} \leq K < \infty$.

(Hunter, 2006)

Perturbation results

Consider perturbing $P = [p_{ij}]$ (where P associated with an ergodic, m-state MC, to $\overline{P} = [\overline{p_{ij}}] = P + \boldsymbol{E}$ where $\boldsymbol{E} = [\varepsilon_{ij}]$, $(\sum_{j=1}^{m} \varepsilon_{ij} = 0)$.

Let $\pi^T = (\pi_1, \pi_2, \ldots, \pi_m)$ and $\overline{\pi}^T = (\overline{\pi}_1, \overline{\pi}_2, \ldots, \overline{\pi}_m)$ be the associated stationary probability vectors.

For all irreducible m-state MCs undergoing a perturbation $\boldsymbol{E} = [\varepsilon_{ij}]$

$$\| \pi^T - \overline{\pi}^T \|_1 \leq (K-1) \| \boldsymbol{E} \|_\infty$$

i.e. $$\sum_{j=1}^{m} | \pi_j^T - \overline{\pi}_j^T | \leq (K-1) \max_{1 \leq i \leq m} \sum_{k=1}^{m} | \varepsilon_{ki} |.$$

(Hunter, 2006)

Elementary perturbations

Let M and $\overline{M}$ be the mean first passage matrices and K and $\overline{K}$ be the Kemeny constants associated with P and $\overline{P}$

Type 1 perturbation: Let $\overline{P} = P + \boldsymbol{E}$ where $\boldsymbol{E} = \boldsymbol{e}_r \boldsymbol{h}^T$.

Then $\quad \overline{m_{ir}} = m_{ir}$ for all $i \neq r$,

$\overline{m_{ij}} \geq m_{ij} \Leftrightarrow \overline{\pi_j} \leq \pi_j$ for all $i, j \neq r$.

and $\quad K \leq \overline{K} \Leftrightarrow \sum_{i \neq r} (\overline{\pi_i} - \pi_i) m_{ir} \geq 0.$

Type 2 perturbation: Let $\overline{P} = P + \boldsymbol{E}$ where $\boldsymbol{E} = \boldsymbol{e}\boldsymbol{h}^T$.

Then $K = \overline{K}$. (Catral, Kirkland, Neumann, Sze, 2010)

Extended perturbations

Extensions:

1. Let P be a symmetric stochastic, irreducible matrix

$\overline{P} = P - \boldsymbol{E}$ where $\boldsymbol{E}$ is a positive semi definite matrix with

$\overline{P}$ stochastic.

Then $\sum_{j=1}^{m} \overline{m}_{ij} \le \sum_{j=1}^{m} m_{ij}$, and $\overline{K} \le K$.

2. Let P be a stochastic, irreducible matrix and suppose $0 \le \alpha \le 1$.

$\overline{P} = \alpha P + (1-\alpha)\boldsymbol{e}\boldsymbol{v}^T$ where $\boldsymbol{v}^T$ is a positive probability vector,

Then $\overline{K} \le K$.

(Catral, Kirkland, Neumann, Sze, 2010)

Directed Graphs

A directed graph, or digraph, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a collection of vertices (or nodes) $i \in \mathcal{V} = \{1, \ldots, m\}$ and directed edges or arcs $(i \to j) \in \mathcal{E}$.

One can assign weights to each directed edge, making it a weighted digraph.

An unweighted digraph has common edge weight 1.

$\mathcal{G}$ can be represented by its $m \times m$ ***adjacency*** matrix $A = [a_{ij}]$ where $a_{ij} \neq 0$ is the weight on arc $(i \to j)$ and $a_{ij} = 0$ if $(i \to j) \notin \mathcal{E}$.

A digraph $\mathcal{G}$ is strongly connected or a strong digraph if there is a path $i = i_0 \to i_1 \to \ldots \to i_k = j$ for any pair of nodes where each link $i_{r-1} \to i_r \in \mathcal{E}$. We focus on strong digraphs.

Random walks over a graph

A random walk over a graph can be represented as a MC with transition matrix $P = D^{-1}A$ where $D = Diag(A\mathbf{e}) = Diag(\mathbf{d})$.
We assume that every node has at least one out-going edge, which can include self loops. Note that $D_{ii} = d_i$, the degree of node i.

The graph is stongly connected $\Rightarrow$ the associated MC is irreducible with $p_{ij} = 1/d_j$ for those states j such that $i \rightarrow j$, 0 otherwise.

The graph is undirected $\Rightarrow$ the associated MC is reversible, and the stationary probability vector $\boldsymbol{\pi}^T = \mathbf{d}/\mathbf{d}^T\mathbf{e}$.

Mixing on Directed Graphs

For any stochastic matrix P of order m, the *directed graph associated with* P, $D(P)$ is the directed graph on vertices labelled $1, 2, ..., m$ such that for each $i, j = 1, 2, ..., m$, $i \to j$ is an arc on $D(P)$ if and only if $p_{ij} > 0$.

For a strongly connected graph D on m vertices define the class

$$\sum_D = \{P \mid P \text{ is stochastic and } m \times m \text{ and for each } i, j = 1, 2, ..., m, \\ i \to j \text{ is an arc on } D(P) \text{ only if } i \to j \text{ is an arc in } D\}$$

Define Kemeny's constant $K(P)$ with the convention that $m_{ii} = 0$.

Let $\mu(D) = \inf\{K(P) \mid P \in \sum_D \textit{ and } P \text{ has 1 as a simple eigenvalue}\}$

Let $k =$ the length if the longest cycle in D, (i.e. period $m \Rightarrow d = m$)

then $$\mu(D) = \frac{2m - k - 1}{2}.$$ (Kirkland, 2010)

Electric networks and graphs

There is a connection between electric networks and random walks (RWs) and graphs. (Doyle & Snell,1984).

On a connected graph G with vertex set $V = \{1, 2,\ ..., m\}$ assign to the edge (i, j) a resistance r_{ij}. The conductance of an edge (i, j) is $C_{ij} = 1/r_{ij}$. Define a RW on G to be a MC with transition probabilities $p_{ij} = C_{ij} / C_i$ with $C_i = \sum_j C_{ij}$.

The graph is connected $\Rightarrow$ MC is ergodic with a stationary probability vector $\boldsymbol{\pi}^T = (\pi_1, ..., \pi_m)$ where $\pi_j = C_j / C$ with $C = \sum_i C_i$.

The MC is in fact reversible.

On the electric network we define $C_{ij} = \pi_i p_{ij}$.

(If $p_{ii} \neq 0$ the resulting network will need a conductance from i to i.)

Electric networks and graphs

For a network of resistors assigned to the edges of a connected graph choose two points *a* and *b* and put a 1-volt battery across these points establishing a voltage $v_a = 1,\ v_b = 0$.

We are interested in finding the voltages v_i and the currents I_{ij} in the circuit and to give a probabilistic interpretation.

By Ohm's Law $I_{ij} = (v_i - v_j)/r_{ij} = (v_i - v_j)C_{ij}$. Note $I_{ij} = -I_{ji}$.

By Kirchhoff's current law $\sum_j I_{ij} = 0$ for $i \neq a,\ b$.

i.e if $\sum_j (v_i - v_j)C_{ij} = 0 \Rightarrow v_i = \sum_j v_j p_{ij}$ for $i \neq a,\ b$.

Let h_i be the probability of starting at *i*, that state *a* is reached before *b*. Then h_i also satisfies above equations with $v_a = h_a = 1$ and $v_b = h_b = 0$. i.e. interpret the voltage as a "hitting probability".

Electric networks and graphs

Let $E_a T_b$ be the expected value, starting from the vertex *a*, of the hitting time T_b of the vertex *b*.

Let π_i be the stationary probability of the MC at vertex *i*.

When we impose a voltage *v* between points *a* and *b* a voltage $v_a = v$ is established at a and $v_b = 0$ and a current $I_a = \sum_x I_{ax}$ will flow into the circuit from outside the source.

We define the effective resistance between *a* and *b* as $R_{ab} = v_a / i_a$,as calculated using Ohm's Law.

Then

$$E_a T_b = \frac{1}{2} \sum_i C_i \{R_{ab} + R_{bi} - R_{ai}\} \qquad \text{(Palacios \&Tetali,1996)}$$

Kirchhoff index

Let G be a simple connected graph with vertex set

$V = \{1, 2, \ldots, m\}$.

Let R_{ij} be the *effective resistance* between i and j.

The ***Kirchhoff index*** is defined as

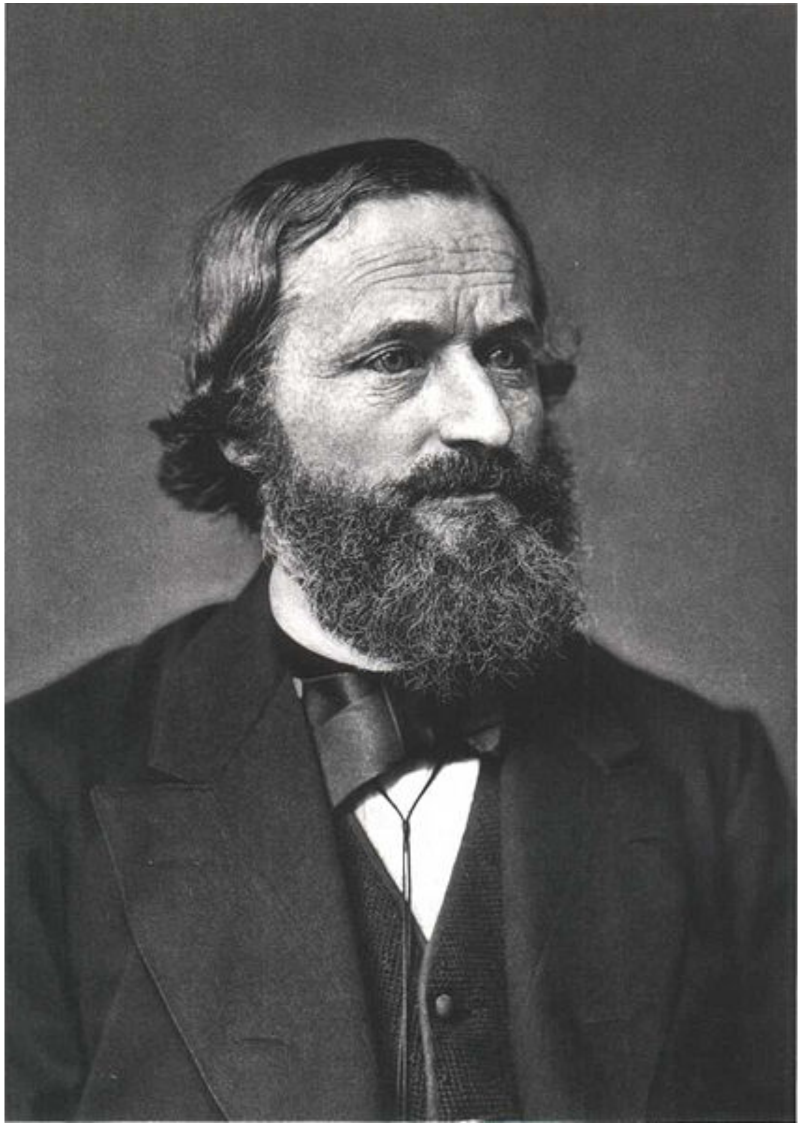
$$Kf(G) = \sum_{i<j} R_{ij}. \qquad \text{(Klein \& Randic,1993)}$$

Since $R_{ij} = R_{ji}$ and $R_{ii} = 0$, $Kf(G) = \frac{1}{2}\sum_{i,j} R_{ij}$.

(Used in Chemistry to discriminate between different molecules with similar shapes and cycle structures.)

A lot of interest in recent years - graph theory, Laplacian and normalised Laplacians, electric networks, hitting times.

Gustav R Kirchhoff (1824 – 1887)



Kirchhoff index

$$Kf(G) = \sum_{i<j} R_{ij}.$$

We use the relations between electric networks and RWs on graphs. For a graph of m vertices computing $Kf(G)$ entails finding $O(m^2)$ values of the R_{ij} which is equivalent to finding $O(m^2)$ values of the $E_i T_j$ for the RW on the graph.

$Kf(G)$ can be characterised as (Palacois & Renom, 2010)

$$Kf(G) = \frac{1}{2|E|} \sum_{i,j} E_i T_j$$

- based on the fact that the "commute times" can be expressed as

$$E_i T_j + E_j T_i = 2|E| R_{ij} \qquad \text{(Aldous \& Fill, 2002)}$$

Kirchhoff index

$Kf(G)$ can also be characterised as $Kf(G) = m\sum_{i=1}^{m-1}\frac{1}{\mu_i}$

(Zhu, Klein, Lukovits, 1996) (Gutman, Mohar, 1996)

where the μ_i's (i = 1, 2, ..,m) with $\mu_m = 0$, are the eigenvalues of the (ordinary or combinatorial) Laplacian matrix L of G, *i.e.* $L = D - A = D(I - P)$.

Using the above characterisation, upper and lower bounds for Kf have been found (Zhou and Trinajstic, 2009). They also found bounds in terms of the eigenvalues of the normalised Laplacian $\mathrm{L} = \mathrm{D}^{-1/2}L\mathrm{D}^{-1/2}$.

Kirchhoff index and Z

In the case of d-regular graphs,(where all vertices have exactly d neighbours) using the characterisation of the Kirchhoff index as

$$Kf(G) = \frac{1}{d} \sum_j E_1 T_j$$

it was shown (Palacois, 2010) that

$$Kf(G) = \frac{m}{d} \, [tr(Z) - 1]$$

where $Z = (I - P + \mathbf{e}\pi^T)^{-1}$, with P the transition matrix of the random walk and $\boldsymbol{\pi}^T$ its stationary probability vector.

Thus we have a connection between the Kirchhof index and Kemeny's constant $K = tr(Z) - 1$.

Variances of mixing times

The expected time to mixing starting in any state is *K*,
a constant independent of the starting state.
What about the variance of the mixing times?
Do these depend on the starting state?
If so, can we choose a desirable starting state?
We explore some results on the second moments of the first passage time variables.
Let $m_{ij}^{(2)}$ be the 2-nd moment of the first passage time from state *i* to state *j*. *i.e.* $m_{ij}^{(2)} = \mathrm{E}[T_{ij}^2 \mid \mathrm{X}_0 = i]$ for all $(i, j) \in S \times S$;
and let $M^{(2)} = \left[m_{ij}^{(2)} \right]$.

Variances of the Mixing Times

Let T be the mixing time variable and let

$\eta_i^{(k)} = E[T^k \mid X_0 = i] = \sum_{j=1}^{m} m^{(k)}{}_{ij}\ \pi_j$ and $\boldsymbol{\eta}^{(k)T} = (\eta_1^{(k)}, \eta_2^{(k)}, \ldots., \eta_m^{(k)})$.

We have seen that $\boldsymbol{\eta}^{(1)T} = K\mathbf{e}$, i.e the expected mixing times, starting at i, is constant.

The variance of the mixing time, starting at i, is given by

$\nu_i = \eta_i^{(2)} - \eta^2$. If $\boldsymbol{\nu}^T = (\nu_1, \nu_2, \ldots, \nu_m)$ then $\boldsymbol{\nu} = \boldsymbol{\eta}^{(2)} - \eta^2\mathbf{e}$.

From (Hunter, 2006), if G is any g-inverse of $I - P$, such that $G\mathbf{e} = \mathbf{e}$

$\boldsymbol{\eta}^{(2)} = [2tr(G^2) - 3tr(G) - (1-2g)(1-g)]\mathbf{e} + 2L\boldsymbol{\alpha},$

$\boldsymbol{\nu} = [2tr(G^2) - (tr(G))^2 - (5-2g)tr(G) - (1-g)(2-3g)]\mathbf{e} + 2L\boldsymbol{\alpha},$

where $L = I - G + EG_d$ and $\boldsymbol{\alpha} = \mathbf{e} - (\Pi G)_d D\mathbf{e} + G_d D\mathbf{e}$.

$\nu_i = \nu$ for all i $\Leftrightarrow L\boldsymbol{\alpha} = l\mathbf{e}$. A sufficient condition is $\boldsymbol{\alpha} = \alpha\mathbf{e}$.

Variances Mixing Times, 2-states

For the 2-state case, $P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$ and $d = 1-a-b$.

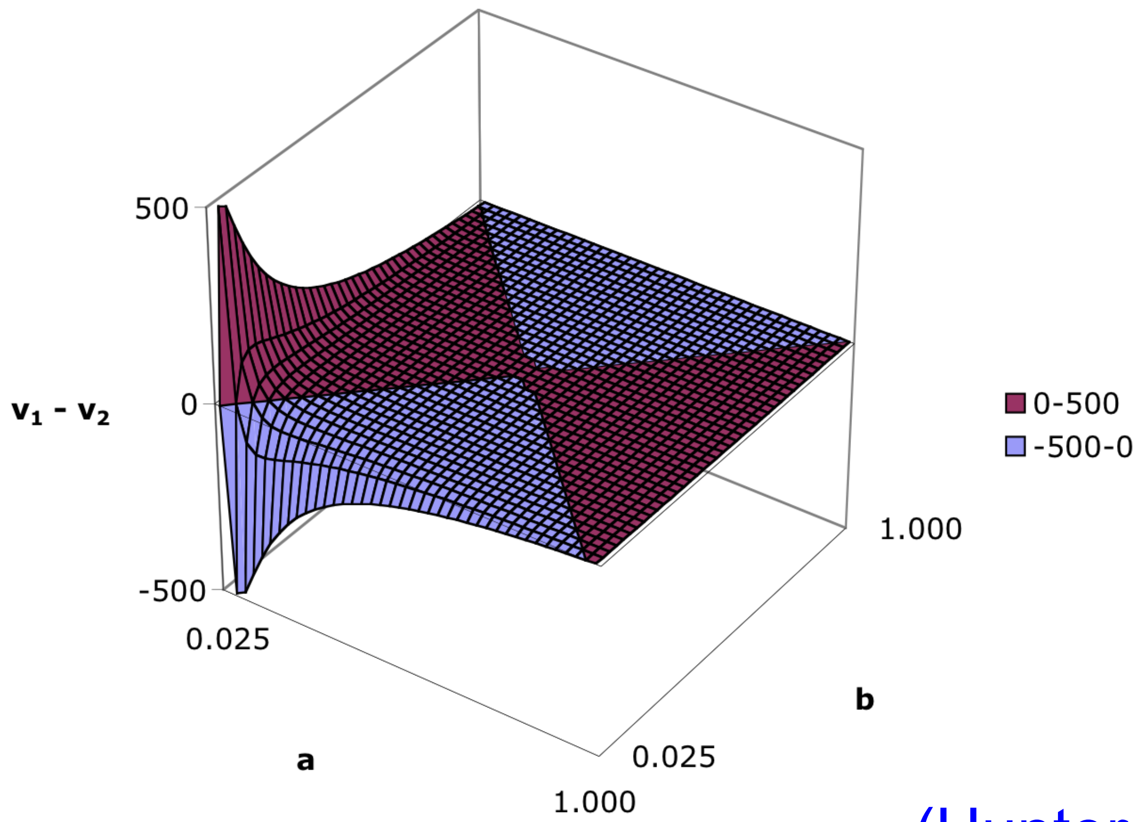
$$\boldsymbol{\nu} = \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} = \frac{1}{ab(1-d)^2} \begin{bmatrix} (2a^2+2b-3ab)(a+b)-ab \\ (2b^2+2a-3ab)(a+b)-ab \end{bmatrix}$$

Lines $a = b$ & $a + b = 1$ partition the parameter space (a,b) to give regions where $\nu_1 = \nu_2, \nu_1 < \nu_2$ and $\nu_1 > \nu_2$.

$\nu_1 < \nu_2$ if $p_{21} < p_{11} < p_{22}$ *or* $p_{22} < p_{11} < p_{21}$.

Variances Mixing Times, 2-states

Graph of $\nu_1 - \nu_2$:



(Hunter, 2008)

References - 1

[1] D. Aldous & J.Fill, Reversible Markov chans and random walks on graphs. http://www.stat.berkeley.edu/users/aldous/RWG/book.html, 2002.

[2] M.Catral, S.J. Kirkland, M. Neumann, N.-S. Sze, The Kemeny constant for finite homogeneous ergodic Markov chains, J. Sci. Computing 45 (2010) 151-166.

[3] P.Doyle, The Kemeny constant of a Markov chain, http://www.math.dartmouth.edu/~doyle/docs/kc/kc.pdf 2009

[4] P.Doyle, J.L.Snell, Random walks and electric networks, Mathematical Association of America, 1984

[5] C.M. Grinstead & J.L.Snell, Introduction to Probability, Second revised Edition, American Mathematical Society, 1997.

[6] I.Gutman, B. Mohar, The quasi-Wiener and the Kirchhoff indices coincide. J. Chem. Inf. Comput. Sci. 36, (1996), 982.

[7] J. J. Hunter, Generalized inverses and their application to applied probability problems, Linear Algebra Appl. 45 (1982) 157-198.

[8] J. J. Hunter, Mixing times with applications to perturbed Markov chains, Linear Algebra Appl. 417 (2006) 108-123.

[9] J. J. Hunter, Variances of First Passage Times in a Markov chain with applications to Mixing Times, Linear Algebra Appl. 429 (2008) 1135-1162.

[10] J. G. Kemeny, J.L. Snell, Finite Markov Chains, Van Nostrand, N.Y., 1960. Republished as Finite Markov Chains : With a New Appendix 'Generalization of a Fundamental Matrix', Springer, 1976.

[11] S.J.Kirkland, Fastest expected time to mixing for a Markov chain on a directed graph, Linear Algebra Applications, 433 (2010), 1988-1996.

[12] D.J.Klein, M. Randic, Resistance distance, J. Math. Chem. 12 (1993) 81-95

References - 2

[13] M.Levene & G. Loizou, Kemeny's constant and the random surfer, American Math Monthly, 109 (2002), 741-745

[14] L. Lovasz, P. Winkler, Mixing times, in: D. Aldous, J. Propp (Eds) Microsurveys in Discrete Probability DIMACS Series in Discrete Math. and Theor. Comput. Sci, AMS. 1998, 85-133

[15] C.D. Meyer, Jr, The role of the group generalized inverse in the theory of finite Markov chains, SIAM Rev 17, (1975), 443-464

[16] C.D. Meyer, Jr, The condition of a finite Markov chain and perturbation bounds for the limiting probabilities SIAM J Algebraic Discrete Methods, 1 (1980), 273-283.

[17] J.L. Palacois, On the Kirchhoff index of regular graphs, Int J Quantum Chem, 110, (2010), 1307- 1309.

[18] J.L. Palacois, J. M. Renom, Bounds for the Kirchhoff index of regular graphs via the spectra of their random walks, Int J of Quantum Chem, 110 (2010), 1637- 1641

[19] J.L. Palacois, P.Tetali, A note on expected hitting times for birth and death chains, Statistics & Probability Letters, 30 (1996), 119-125

[20] C.R. Rao, Generalized inverse for matrices and its applications in mathematical statistics, Research Papers in Statistics, Festschrift for J. Neyman, Wiley, N.Y., 1966

[21] G.P.H. Styan, Selected topics in Markov chain theory with an application to marketing, MA thesis, Columbia University, 1964.

[22] P. Schweitzer, Perturbation theory and finite Markov chains, J. Appl. Probab. 5 (1968), 410-413.

[23] B.Zhou, N Trinajstic, On resistance-distance and Kirchhoff index, J. Math Chem 46 (2009), 283-289

[24] H.-Y Zhu, D.J. Klein, I. Lukovits, Extensions of the Weiner number, J. Chem. Inf. Comput. Sci. 36 (1996), 420-428