Mixing Time of Markov Chains

Steven Finch

August 4, 2015

We will concentrate on two specific examples, leaving general theory aside. Consider the cycle Z_n (integers modulo n) as our state space. A **lazy random walk** is a particle that moves left or right, each with probability 1/4, or remains motionless with probability 1/2. Let us assume that the starting point is at 0. After how many time steps is the distribution of the particle close to uniform?

The transition matrix Q, whose ij^{th} element conveys the odds that the particle is at site j given it was at site i one step earlier, is

$$Q = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

when $n = 6$. If we wish information on the odds over a separation of t (positive integer) steps, then the matrix product Q^t is required.

Let μ_t denote the first row of Q^t and ν denote the vector $(1/n, 1/n, \ldots, 1/n)$. Define

$$d(t) = \frac{1}{2} \left\| \mu_t - \nu \right\|_1,$$

one-half the L_1 norm of the vector difference (a sum of absolute values). This is called the **total variation distance**. Now define

$$t_{\text{mix}}(\varepsilon) = \min \left\{ t \geq 1 : d(t) \leq \varepsilon \right\},$$

$$t_{\text{mix}} = t_{\text{mix}}(1/4)$$

the **mixing time**. For the case $n = 6$, we compute

$$\mu_3 = \left(\begin{array}{cccccc} \frac{5}{16} & \frac{15}{64} & \frac{3}{32} & \frac{1}{32} & \frac{3}{32} & \frac{15}{64} \end{array} \right),$$

$$\mu_4 = \left(\begin{array}{cccccc} \frac{35}{128} & \frac{7}{32} & \frac{29}{256} & \frac{1}{16} & \frac{29}{256} & \frac{7}{32} \end{array} \right)$$

and $d(3) = 9/32 > 0.28$, $d(4) = 27/128 < 0.22$, therefore $t_{\text{mix}} = 4$. Our interest is in the growth of t_{mix} as $n \to \infty$. It is known that [1]

$$c\,n^2 < t_{\text{mix}} \le n^2$$

for some $c > 0$; simulation suggests that t_{mix}/n^2 approaches a constant ≈ 0.0949.

A **(non-lazy) random walk** is a particle that moves left or right, each with probability 1/2. The transition matrix P is

$$P = \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \end{pmatrix}$$

when $n = 7$. For technical reaons, we must restrict the cycle length n to be odd (to ensure aperiodicity). Let μ_t denote the first row of P^t and everything else be as before. For the case $n = 7$, we compute

$$\mu_8 = \left(\tfrac{35}{128} \quad \tfrac{9}{256} \quad \tfrac{7}{32} \quad \tfrac{7}{64} \quad \tfrac{7}{64} \quad \tfrac{7}{32} \quad \tfrac{9}{256} \right),$$
$$\mu_9 = \left(\tfrac{9}{256} \quad \tfrac{63}{256} \quad \tfrac{37}{512} \quad \tfrac{21}{128} \quad \tfrac{21}{128} \quad \tfrac{37}{512} \quad \tfrac{63}{256} \right)$$

and $d(8) = 253/896 > 0.28$, $d(9) = 223/896 < 0.24$, therefore $t_{\text{mix}} = 9$. Again, the growth rate of t_{mix} is quadratic in n; simulation suggests that t_{mix}/n^2 approaches a constant ≈ 0.1898. We also mention rigorous bounds [2, 3, 4]

$$\left(\frac{2n^2}{\pi^2} - 1\right)\ln(2) \le t_{\text{mix}} \le \frac{4n^2}{\pi^2}\ln(2)$$

which imply that the ratio falls between 0.14 and 0.28. Similar bounds could be determined for the lazy case. The non-lazy mixing time is at most twice the lazy mixing time, but may be less.

A remarkable equation for the lazy constant $C \approx 0.0949$ was announced in [5]:

$$\frac{1}{2}\int_0^1 \left| -1 + \sum_{k=-\infty}^{\infty} \frac{1}{\sqrt{C}\,\pi} \exp\left(-\frac{(x-k)^2}{C}\right) \right| dx = \frac{1}{4}$$

which gives a more accurate estimate $C = 0.0948705678\ldots$. The justification involved passage from discrete (n-cycle) to continuous (circle), Fourier analysis, and reinterpretation of random walks as heat flow. Unfortunately the authors of [5] never completed their proof – their draft preprint is no longer available online – and we are left wondering if/how the challenging details can be brought together. It appears likely that $2C$ is the corresponding non-lazy constant, but verification remains open as well.

Setting $\varepsilon = 1/4$ is, of course, arbitrary. For many Markov chains (not our two examples), there is a more natural choice of threshold. In such scenarios, the variation distance $d(t)$ is fairly large and essentially flat for small t, then abruptly changes character and decays exponentially to zero as t increases beyond a certain point. It is believed that such *cut-off phenomena* are widespread, although they have been rigorously ascertained only sporadically (for example, riffle shuffles of 52 cards [6, 7, 8, 9]). How are the group theoretic properties of the state space related to the existence or non-existence of a cut-off? This is a difficult question; we must often settle for the order of magnitude (as a function of n) of a possible threshold. Only rarely are these results so accurate as to yield tight bounds on the level of a constant.

0.1. Addendum. On the one hand, given any $\varepsilon > 0$, the equation for $t_{\mathrm{mix}}(\varepsilon)/n^2$ in the limit as $n \to \infty$ is the same as that for C except $1/4$ on the right-hand side is replaced by ε. For example, if $\varepsilon = 1/10$, then the limit is 0.1875465011....

On the other hand, consider a random walk in which a particle moves left or right, each with probability 1/3, or remains motionless with probability 1/3. What does the heuristic in [5] predict for the value of $t_{\mathrm{mix}}(\varepsilon)/n^2$? Intuition suggests that the variance of the walk generator is key. The walk with probabilities $\{1/4, 1/2, 1/4\}$ has variance 1/2; the walk with probabilities $\{1/3, 1/3, 1/3\}$ has variance 2/3; dividing 1/2 by 2/3 yields 3/4. For example, if $\varepsilon = 1/4$, then the limit is ≈ 0.0712 via simulation; if $\varepsilon = 1/10$, then the limit is ≈ 0.1406. These compare well with multiplying 0.0948705678... and 0.1875465011... respectively by 3/4.

0.2. Acknowledgements. I am grateful to Peter Winkler, Aaron Smith, Stefan Steinerberger, Natesh Pillai and Ravi Montenegro for helpful discussions.

REFERENCES

[1] D. Levin, Y. Peres and E. Wilmer, *Markov Chains and Mixing Times*, Amer. Math. Soc., 2009, pp. 3–5, 9, 47–48, 53–55, 65, 96; http://pages.uoregon.edu/dlevin/MARKOV/; MR2466937 (2010c:60209).

[2] P. Diaconis, *Group Representations in Probability and Statistics*, Inst. Math. Stat., 1988, pp. 17–30; http://jdc.math.uwo.ca/M9140a-2012-summer/Diaconis.pdf; MR0964069 (90a:60001).

[3] R. Montenegro, The mixing time of simple random walk on a cycle, http://faculty.uml.edu/rmontenegro/8843/notes/lecture6.pdf.

[4] R. Montenegro, Exponential decay and lower bounding mixing time, http://faculty.uml.edu/rmontenegro/8843/notes/lecture7.pdf.

[5] A. Bloemendal and K. Yang, Mixing time of lazy simple random walk on the n-cycle, unpublished note (2010); http://www.individual.utoronto.ca/09yangka/research.html.

[6] D. Aldous and P. Diaconis, Strong uniform times and finite random walks, *Adv. Appl. Math.* 8 (1987) 69–97; https://statistics.stanford.edu/sites/default/files/EFS NSF 249.pdf; MR0876954 (88d:60175).

[7] P. Diaconis, The cutoff phenomenon in finite Markov chains, *Proc. Nat. Acad. Sci. U.S.A.* 93 (1996) 1659–1664; http://statweb.stanford.edu/~cgates/PERSI/papers/cutoff.pdf; MR1374011 (97b:60112).

[8] P. Diaconis, The mathematics of mixing things up, *J. Stat. Phys.* 144 (2011) 445–458; http://statweb.stanford.edu/~cgates/PERSI/papers/mixing.pdf; MR2826629 (2012j:60207).

[9] L. Saloff-Coste, Random walks on finite groups, *Probability on Discrete Structures*, ed. H. Kesten, Springer-Verlag, 2004, pp. 263–346; http://statweb.stanford.edu/~cgates/PERSI/papers/rwfg.pdf; MR2023654 (2004k:60133).