Notes on finite-state Markov chains
6.436J/15.085J, Fall 2005

Notation:
$p_{ij} = \mathbf{P}(X_{n+1} = j \mid X_n = i)$.
P: matrix with entries p_{ij}
$r_{ij}(n) = \mathbf{P}(X_n = j \mid X_0 = i)$
$t_s^* = \mathbf{E}[\#$ of transitions in a cycle that starts and ends at $s]$
See [BT, Problem 7] for two equivalent characterizations of aperiodicity.

Fact: $\mathbf{P}(X_n$ is transient$) \to 0$ as $n \to \infty$; furthermore, X_n eventually enters the set of recurrent states with probability 1 [BT, Problem 5].

Absorption probability equations
Let R be a recurrent class. For each i, let a_i be the probability that the process eventually enters R, given that it starts at i. Then, the a_i are the unique solution to the system of equations

$$a_i = \begin{cases} 1, & i \in R, \\ 0, & i \text{ recurrent}, \ i \notin R, \\ \sum_j p_{ij} a_j, & i \text{ transient.} \end{cases}$$

See [BT, problem 29] for a proof of uniqueness.

Expected time to absorption
Let μ_i be the expected time until the process enters a recurrent state, starting from state i. Then, the μ_i are the unique solution to the system of equations

$$\mu_i = \begin{cases} 0, & i \text{ recurrent}, \\ 1 + \sum_j p_{ij} \mu_j, & i \text{ transient} \end{cases}.$$

First passage times
Suppose there is a single recurrent class and that s is a recurrent state. Let t_i be the expected time until the state first becomes s, starting from state i. Then, the t_i are the unique solution to the system of equations

$$t_i = \begin{cases} 0, & \text{if } i = s, \\ 1 + \sum_j p_{ij} t_j, & i \neq s. \end{cases}$$

Mean recurrence time
Let s be a recurrent state. Let t_s^* be the mean time between successive visits to the state s. Then,

$$t_s^* = 1 + \sum_j p_{sj} t_j.$$

Steady-state behavior

Theorem:

(a) There exists a nonnegative vector π such that

$$\pi_j = \sum_k \pi_k p_{kj}, \quad \forall j, \qquad \sum_i \pi_i = 1. \tag{1}$$

(We call such a vector π, an *invariant distribution.*)

(b) One particular invariant distribution can be constructed as follows. Let R be a recurrent class and fix some state $s \in R$. Let $\pi_i = 0$ for $i \notin R$. For $i \in R$, let

$$\pi_i = \frac{\mathbf{E}[\#\text{ of transitions into } i \text{ in a cycle that starts and ends at } s]}{\mathbf{E}[\#\text{ of transitions in a cycle that starts and ends at } s]}.$$

(In particular, $\pi_s = 1/t_s^*$, where t_s^* is the mean recurrence time of state s.)

(c) If π is an invariant distribution and i is a transient state, then $\pi_i = 0$.

(d) If there is a single recurrent class, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} r_{ij}(t) = \pi_j, \qquad \forall\ i, j.$$

(e) If there is a single recurrent class, which is aperiodic, then

$$\lim_{n\to\infty} r_{ij}(n) = \pi_j, \qquad \forall i, j.$$

(f) There exists only one invariant distribution if and only if there is a single recurrent class.

(g) There exists only one solution to the system (1) if and only if there is a single recurrent class.

Proof pointers.

(a) Use Brouwer's fixed point theorem: if $S \subseteq \Re^n$ is closed, bounded, and convex, and $f : S \to S$ is continuous, there exists some $x \in S$ such that $f(x) = x$. Apply this to the map $\pi^T \mapsto \pi^T P$, with S being the set of all nonnegative vectors whose components sum to 1.

(b) See [BT, Problem 32].

(c) If $\mathbf{P}(X_0 = i) = \pi_i$ for all i, then $\mathbf{P}(X_n = i) = \pi_i$ for all i. But for transient states i, the latter probability goes to zero.

(d) Omitted.

(e) The coupling argument in [BT, Problem 17] shows that $r_{ij}(n) - r_{kj}(n) \to 0$. Use this to show that $\max_{n \geq m}(r_{ij}(n) - r_{ij}(m)) \to 0$ as $m \to \infty$. Thus $r_{ij}(n)$ is a "Cauchy sequence", which guarantees that it converges to something. This limit must be the same for all i because of the previous fact.

(f) See [BT, Problem 18].

(g) Proof of the "if" part. Let π be a nonnegative solution (exists by part (a) and $\pi_i > 0$ when i is recurrent, by part (c)). Let $\overline{\pi}$ be another solution. Check that $\overline{\pi}_i = 0$ for i transient. If α is a large enough positive number, then $\alpha\pi + \overline{\pi}$ (normalized so that its components sum to 1) is a second nonnegative solution, contradicting part (f).

The "only if" part is an immediate consequence of part (f).