Unless otherwise noted, let $X$ be an irreducible, aperiodic Markov chain on countable state-space $S$ with initial distribution $\mu$ and transition probabilities $P$.

**Example 1.** Walk on a Finite, Directional Graph

As before, let $V = \{v_1, \ldots, v_m\}$ be a collection of vertices. Let $E(v, v')$ be 1 if there is an edge in the graph between $v$ and $v'$ and 0 otherwise. Unlike the last case, we do not assume $E(v, v') = E(v', v)$.

Let $\mu$ be an initial distribution on $V$. Let $P$ be a transition probability matrix on $V$ that satisfies $P(v, v') = 0$ if and only if $E(v, v') = 0$. Define $Y = (Y_n)_{n \geq 0}$ as before.

One approach to understanding the chain is to do the conditioning trick that we’ve used before and apply generating functions. Let $\mu_n$ be the distribution of $Y_n$.

$$\mu_n(v) = \mu(v)1_{(n=0)} + \sum_{v'} \mu_{n-1}(v')P(v', v)1_{(n>0)}.$$  \hspace{1cm} (1)

Define $G(z) = \sum_{n \geq 0} \mu_n z^n$ be the vector-valued generating function. Then, from the above recursion

$$\sum_n \mu_n z^n = \mu + \sum_{v'} \sum_n \mu_{n-1}(v')z^n P(v', \cdot)$$  \hspace{1cm} (2)

$$G(z) = \mu + zG(z) \cdot P.$$  \hspace{1cm} (3)

So,

$$G(z)(I - zP) = \mu$$  \hspace{1cm} (4)

$$G(z) = \mu(I - zP)^{-1} = \sum_{n \geq 0} z^n \mu P^n,$$  \hspace{1cm} (5)

with the last assuming the inverse matrix exists. We can use this to understand both the short and long-term behavior of the chain.

If $Y$ is irreducible, then $O(s, S) = \infty$ implies that the chain is recurrent because the sum $O(s, S) = \sum_{s'} O(s, s')$ is finite, implying at least one (and thus all) terms must be infinite. Because any invariant measure on the finite state space will consequently be finite and hence normalizable, $Y$ is positive recurrent as well.

The search for a stationary distribution $\pi$ corresponds to a search for eigenvectors of $P$ with eigenvalue 1, vectors $\pi$ such that $\pi(I - P) = 0$.  

1 28 Feb 2006
Example 2. Finite, Doubly Stochastic Chain

Suppose $U = (U_n)_{n \geq 0}$ is an irreducible Markov chain on a finite state-space $S$ such that $\sum_u P(u, u') = 1$. Because $\sum_{u'} P(u, u') = 1$, this implies that both the rows and columns sum to 1. Such a matrix is called doubly stochastic.

What is the stationary distribution of the chain?

Note that for any constant $c$, $\sum_u cP(u, u') = c$. Hence, this is an invariant measure. The unique invariant distribution $\pi$ is then just a uniform distribution on $S$.

Example 3. Renewal Process and Forward Recurrence Time Chain

Consider a process $Z = (Z_n)_{n \geq 0}$ given by

$$Z_n = Z_0 + \sum_{k=1}^{n} \Xi_k$$

for iid random variables $\Xi_i$ and arbitrary $Z_0$. This is called a renewal process.

As an example, consider a critical part in a system that operates for some time and then fails. When it fails it is replaced. Think of $\Xi$ as the lifetime of the replacement parts and $Z_0$ as the lifetime of the original part. Then, $Z_n$ is the time of the $n$th replacement – a "renewal" of the system.

Given a renewal process, we can define $N_t = \sup\{n \geq 0: Z_n \leq t\}$. Then, $N_t$ is a counting process that counts the renewals up to time $t$.

Suppose now that $Z_0$ and $\Xi_1$ take values in $\mathbb{Z}_+$ and have p.m.f.s $\mu \equiv p_{z_0}$ and $p_{\Xi} \equiv p_{z_1}$. $Z_n$ is thus a countable-state Markov chain, but not a terribly interesting one as it marches inexorably toward $\infty$.

But we can define two related processes that will turn out to be very interesting in general. Define $V^+$ and $V^-$ to be, respectively, the forward and backward recurrence time chains, as follows for $n \geq 0$:

$$V_n^+ = \inf\{Z_m - n: Z_m > n\}$$

$$V_n^- = \inf\{n - Z_m: Z_m \leq n\}.$$  

Then $V_n^+$ represents the time until the next renewal, and $V_n^-$ represents the time since the last renewal. These are sometimes also called the residual lifetime and age processes.

Are these Markov chains? What are the state spaces?

We can check the Markov Property explicitly. But the regeneration of the system at each renewal gives us a simple way of seeing it. When $V_n^+ > 1$, for instance, the next time is determined. When $V_n^+ = 1$, a renewal ensues and the next time is an independent waiting time.

The state space of $V^+$ is $\mathbb{Z}_+$; the state space of $V^-$ is $\mathbb{Z}_{\leq 0}$.
What are the transition probabilities?

If \( V^+_n = k > 1 \), then \( V^+_{n+1} = k - 1 \) by construction. If \( V^+_n = 1 \), then a renewal occurs at time \( n + 1 \), so the time until the following renewal has distribution \( \xi \). Hence,

\[
P(k, k - 1) = 1 \quad \text{for } k > 1, \quad P(1, k) = \xi(k) \quad \text{for } k \in \mathbb{Z}_+.
\] (9) (10)

For \( V^- \), we can reason similarly. Let \( S_\Xi \) be the survival function of \( \Xi_1 \). Then,

\[
P(k, k + 1) = P\{\Xi > k + 1 \mid \Xi > k\} = \frac{S_\Xi(k + 1)}{S_\Xi(k)} \quad \text{(11)}
\]

\[
P(k, 0) = P\{\Xi = k + 1 \mid \Xi > k\} = \frac{p_\Xi(k + 1)}{S_\Xi(k)}. \quad \text{(12)}
\]

Is \( V^+ \) irreducible? Is it recurrent?

If there exists an \( M \in \mathbb{Z}_+ \) such that \( S_\Xi(M) = 0 \) and \( p_\Xi(M) > 0 \), then all states \( j > M \) are transient, and all states \( \{1, \ldots, M\} \) communicate and are recurrent since there is only a finite number. (To see the latter, note that we can find a positive probability path between each pair of states in this set.)

If no such \( M \) exists, then \( V^+ \) is irreducible. Note that for all states \( n > 1 \), \( P_{n-1}^n(n, 1) = 1 \). Hence,

\[
R(1, 1) = \sum_{n \geq 1} p_\Xi(n) P_{n-1}^n(n, 1) = 1.
\] (13)

So the chain is recurrent in this case as well.

What is the long-run behavior of the chain?

Let \( \rho(j) = \sum_{n \geq 1} P_{n}^n(1, j) \). Because \( P_{n}^n(1, j) = p_\Xi(j + n - 1) \) for \( n \geq 1 \), we can write

\[
\rho(j) = \sum_{n \geq 1} p_\Xi(j + n - 1) = \sum_{n \geq j} p_\Xi(n) = S_\Xi(j - 1).
\] (14)

Notice that

\[
\sum_{j \geq 1} \rho(j) P(j, k) = \rho(1)p_\Xi(k) + \rho(k + 1)
\]

\[
= p_\Xi(k) + S_\Xi(k) \quad \text{(15)}
\]

\[
\rho(k).
\] (16) (17)

This invariant measure is positive (on its support) and is finite if and only if

\[
\sum_{n \geq 1} \rho(n) = \sum_{n \geq 1} S_\Xi(n - 1) = \sum_{n \geq 1} np_\Xi(n) = \mathbb{E}\Xi_1 < \infty.
\] (18)

In this case, \( \pi(k) = \rho(k)/\mathbb{E}\Xi_1 \) is a stationary distribution.
Idea 4. The above argument leads to an interesting idea for the countable case.

Suppose that $X$ is recurrent, pick a state $s_0 \in S$ such that it is easy to compute $P^n_{s_0} (s_0, s)$ for any $s \in S$. Define

$$\rho(s) = \sum_{n \geq 1} P^n_{s_0} (s_0, s). \quad (19)$$

Note that $\rho(s_0) = 1$ because the chain is recurrent (i.e., $R(s_0, s_0) = 1$). Then, mimicking the above argument, we get

$$\sum_{s \in S} \rho(s) P(s, s') = \rho(s_0) P(s_0, s') + \sum_{s \neq s_0} \sum_{n \geq 1} P^n_{s_0} (s_0, s) P(s, s') \quad (20)$$
$$= P(s_0, s') + \sum_{n \geq 2} P^n_{s_0} (s_0, s') \quad (21)$$
$$= P_{s_0} (s_0, s') + \sum_{n \geq 2} P^n_{s_0} (s_0, s') \quad (22)$$
$$= \rho(s'). \quad (23)$$

This is finite only if

$$\rho(S) = \sum_{s \in S} \sum_{n \geq 1} P^n_{s_0} (s_0, s) \quad (24)$$
$$= \sum_{n \geq 1} \sum_{s \in S} P^n_{s_0} (s_0, s) \quad (25)$$
$$= \sum_{n \geq 1} P^n_{s_0} (s_0, S) \quad (26)$$
$$= \sum_{n \geq 1} P_{s_0} \{ T_{s_0} \geq n \} \quad (27)$$
$$= E_{s_0} T_{s_0} \quad (28)$$
$$< \infty, \quad (29)$$

which is just the positive recurrence condition.

Hence,

$$\pi(s) = \frac{\sum_{n \geq 1} P^n_{s_0} (s_0, s)}{\sum_{n \geq 1} P^n_{s_0} (s_0, S)} = \frac{1}{E_{s_0} T_{s_0}}. \quad (30)$$

We already knew the last equality, but this gives a new way of finding a stationary distribution, using equation (19).
Example 5. Upper Hessenberg Transition Probabilities

Suppose $S = \mathbb{Z}_0$ and suppose the non-zero transition probabilities are of the form

$$P(0, k) = a_k \quad \text{for } k \geq 0,$$  
(31)

$$P(j, k) = a_{k-j+1} \quad \text{for } j \geq 1, k \geq j - 1,$$  
(32)

where $(a_n)_{n \geq 0}$ is a non-negative sequence satisfying $\sum_n a_n = 1$ and $\sum_n na_n < \infty$.

We want to determine the stability of this chain. Consider the function $\Delta(s)$ on the state space, given by

$$\Delta(s) = \mathbb{E}(X_{n+1} - X_n \mid X_n = s).$$  
(33)

This tells us how much the chain “drifts” on average in one step when starting at $s$. Note that because we are dealing with a time-homogeneous chain,

$$\Delta(s) = \mathbb{E}_s X_1 - s = \sum_{s'} P(s, s') s' - s.$$  
(34)

Let’s compute this function.

$$\Delta(0) = \sum_k P(0, k) k$$

$$= \sum_k ka_k < \infty$$  
(35)

$$\Delta(j) = \sum_k P(j, k) k - j$$

$$= \sum_{k=j-1}^{\infty} a_{k-j+1} k - j$$

$$= \sum_{k=j-1}^{\infty} a_{k-j+1} (k - j + 1) + \sum_{k=j-1}^{\infty} a_{k-j+1} (j - 1) - j$$

$$= \sum_{n=0}^{\infty} na_n - 1 < \infty,$$  
(36)

for $j \geq 1$.

Case (i): $\sum_n na_n > 1$.

In this case, $\Delta(j) > 0$ for every $j$, hence on average in any state, we tend toward higher states. This “positive drift” seems to suggest (though does not prove) transience.

Case (ii): $\sum_n na_n < 1$.

In this case, $\Delta(0) > 0$ and $\Delta(j) < -\epsilon \equiv \sum_n na_n - 1$ for all $j \geq 1$. Hence, on average, whenever the chain is away from zero, it tends to move back toward zero. This suggests recurrence.

What we can we make of this?
**Definition 6. Drift**

Let $V$ be a non-negative function on the state space (that is, $V: S \to \mathbb{R}_+$). Define the drift operator $\Delta_X$ by

$$\Delta_X V = PV - V.$$  (41)

That is,

$$(\Delta_X V)(s) = \sum_{s' \in S} P(s, s')V(s') - V(s) = \mathbb{E}(V(X_{n+1}) - V(X_n) \mid X_n = s) = \mathbb{E}_s(V(X_1)) - V(s).$$  (42)

Note that, in general, $\Delta_X V$ takes values in $[-\infty, \infty]$.

**Theorem 7. Foster’s Drift Criterion**

Suppose there exists a non-negative function $V: S \to \mathbb{R}_+$, an $\epsilon > 0$, and a finite set $S_0 \subset S$ such that

$$|\Delta_X V(s)| < \infty \quad \text{for } s \in S_0$$  (43)
$$\Delta_X V(s) \leq -\epsilon \quad \text{for } s \notin S_0.$$  (44)

Then, $X$ is positive recurrent.

**Proof**

For $s \in S_0$, $|\Delta_X V(s)| < \infty$ implies that $|PV| < \infty$ on $S_0$. Define

$$u^{[n]}(s) = \sum_{s'} P^n(s, s')V(s'),$$  (45)

for $n \geq 0$.

Notice that for $m \geq 0$,

$$u^{[m+1]}(s) = \sum_{s'} P^{m+1}(s, s')V(s')$$  (46)

$$= \sum_{s'} \sum_{t \in S} P^m(s, t)P(t, s')V(s')$$  (47)

$$= \sum_{t \in S} P^m(s, t) \sum_{s'} P(t, s')V(s')$$  (48)

$$= \sum_{t \in S} P^m(s, t)(PV)(t)$$  (49)

$$\leq \sum_{t \in S_0} P^m(s, t)(PV)(t) + \sum_{t \in S_0} P^m(s, t)(PV)(t)$$  (50)

$$\leq \sum_{t \in S_0} P^m(s, t)((PV)(t) + \epsilon) + \sum_{t \in S_0} P^m(s, t)(V(t) - \epsilon)$$  (51)

$$= \sum_{t \in S_0} P^m(s, t)((PV)(t) + \epsilon) + u^{[m]}(s) - \epsilon.$$  (52)

This gives us an upper bound for $u^{[m+1]} - u^{[m]}$. Summing these together by telescoping gives

$$0 \leq u^{[n+1]}(s) \leq u^{[0]}(s) + \sum_{t \in S_0} \sum_{m=0}^{n} P^m(s, t)(PV(t) + \epsilon) - (n + 1)\epsilon.$$  (54)

6  28 Feb 2006
Rearranging and dividing by \( n + 1 \) gives
\[
\frac{u[0](s)}{n + 1} + \sum_{t \in S_0} \left( \frac{1}{n + 1} \sum_{m=0}^{n} P^m(s, t) \right) (PV(t) + \epsilon) \geq \epsilon. \tag{55}
\]

Taking limits on both sides as \( n \to \infty \) and using the fact that \( S_0 \) is finite and \( 0 \leq PV(t) < \infty \) on \( S_0 \), gives us
\[
\lim_{n \to \infty} \frac{1}{n + 1} \sum_{m=0}^{n} P^m(s, t) > 0, \tag{56}
\]
for sum \( t \in S_0 \). But it follows that (Césaro summation), that \( \sum_n P^n(s, t) \) cannot converge. Hence, \( X \) is recurrent. Positive recurrence follows as well, though somewhat more delicately.

**Example 5 cont’d Upper Hessenberg Transition Probabilities**

We know from the above that \( \sum_n na_n < 1 \) implies that this chain is positive recurrence. What about when \( \sum_n na_n \geq 1 \)? Can we deduce transience from this?

It turns out that Foster’s criterion cannot simply be reversed. We can get the following.

**Theorem 8.** Suppose there exists a bounded, non-negative function \( V \) on \( S \) and \( r \geq 0 \) such that \( \{ s \in S : V(s) > r \} \) and \( \{ s \in S : V(s) \leq r \} \) are both nonempty and
\[
\Delta_X V(s) > 0 \quad \text{if } V(s) > r. \tag{57}
\]

Then, \( X \) is transient. The converse is also true.

The proof of this relies on a rather cute result:

**Lemma** Let \( C \subset S \). Let \( h_*(s) = H(s, C) \) be the hitting probability of \( S_0 \) from \( S \). (Recall, the hitting time \( S_C \) is zero if the chain starts in \( C \) and is otherwise equal to \( T_C \).) Then, if \( h:S \to \mathbb{R}_{\geq 0} \) is a solution to
\[
\Delta_X h(s) \leq 0 \quad \text{if } s \in C^c \tag{58}
\]
\[
\Delta_X h(s) \geq 1 \quad \text{if } s \in C \tag{59}
\]

Then, \( h^* \leq h \).

Now, to the theorem, suppose that \( |V| \leq M \). We must have, by the conditions, that \( M > r \). (Why?) Define
\[
h_V(s) = \begin{cases} 
1 & \text{if } V(s) \leq r \\
\frac{M-V(s)}{M-r} & \text{if } V(s) > r.
\end{cases} \tag{60}
\]

Then, we can show that \( h_V \) solves (58) and (59) with \( C = \{ s : V(s) \leq r \} \). So \( h_* \leq h_V \). But then \( h_*(s) \leq h_V(s) < 1 \) if \( s \notin C \), which shows that \( R(s, s') < 1 \) for \( s \in C^c \) and \( s' \in C \). Transience follows.

**Example 5 cont’d Upper Hessenberg Transition Probabilities**

Can we show that \( \sum_n a_n > 1 \) implies transience, using the above?
Example 9. Storage Model

Consider a storage system (dam, warehouse, insurance policy) that receives inputs at random times but otherwise drains at a regular rate.

Let $T_0 = 0$ and let the rest of the $T_i$s be iid $\mathbb{Z}_{\geq 0}$-valued with CDF $G$. These are inter-arrival times for the inputs to our storage system. Let the $S_n$s be iid $\mathbb{Z}_{\geq 0}$ with CDF $H$. These are the amounts input at the time $Z_n = T_0 + \cdots + T_n$. Assume that the $S_n$s and $T_n$s are independent of each other as well. Suppose also that the storage system “drains” or outputs at rate $r$ between inputs.

Define a process $(V_n)_{n \geq 0}$ by

$$V_{n+1} = (V_n + S_n - rT_{n+1})_+.$$  \hfill (61)

Here, $V_n$ represents the contents of the storage system just before the $n$th input (that is, at time $Z_n$).

*Is $V$ a Markov chain?*

*What is the structure of the transition probabilities?*

*What can we say about the long-run behavior of the chain?*

*What is special about the state $\{0\}$?*

*How might we generalize this model to make it more realistic?*