Plan Limits and Examples (Countable-state case)

0. Feedback
1. Summary
2. Invariant Distributions and Limit Theorems
3. Extended Examples

Next Time: More Examples and Applications

Homework 4 now on-line, due in two weeks.

Correction 1. "Yanko" in previous handout. The Strong Markov Property reads:

Let T be a stopping time, X be a (discrete-time) Markov chain, and let h be a bounded, measurable function on sample paths. Then,

$$\mathsf{E}\left(h(X_{T+1}, X_{T+2}, \ldots) \mid \mathcal{F}_T\right) 1\{T < \infty\} = \mathsf{E}_{X_T}(h(X_1, X_2, \ldots) 1\{T < \infty\}, \tag{1}$$

where $E_{X_T} h(X_1, X_2, \ldots)$ corresponds to a chain whose initial distribution on $\mathcal{S}$ is the distribution of X_T. (Specifically, if $g(s) = \mathsf{E}_s h(X_1, X_2, \ldots)$, then that random variable is $g(X_T)$.) The indicator says that this equality only applies if $T < \infty$. (Otherwise, for example, X_T does not make sense.)

Note the connection with the standard Markov property for stopping time T and deterministic time n:

$$\mathsf{E}\left(h(X_{T+1}, X_{T+2}, \ldots) \mid \mathcal{F}_T\right) 1\{T < \infty\} = \mathsf{E}_{X_T}(h(X_1, X_2, \ldots) 1\{T < \infty\} \tag{2}$$

$$\mathsf{E}\left(h(X_{n+1}, X_{n+2}, \ldots) \mid \mathcal{F}_n\right) = \mathsf{E}_{X_n}(h(X_1, X_2, \ldots) \quad . \tag{3}$$

Summary 2. Markov Chains: Notation (short form)

We are working in a probability space $(\Omega, \mathcal{F}, \mathsf{E})$.

Concept	Notation
Markov Chain	$X = (X_n)_{n \geq 0}$. X_n is state at time n. X is random function (sample path). "X" is default but not exclusive.
State Space	$\mathcal{S}$ set of possible values of the chain at any given time. $\mathcal{S}$ is the default but not exclusive.
Initial Distribution	μ, a probability distribution on $\mathcal{S}$. In the countable-state case, this is often used as a row vector. μ is the default but not exclusive.
Transition Probabilities	P, either as a transition probability kernel $P(s, A)$ or as a transition probability matrx $P(s, s') \equiv P_{ss'}$. In either case, P^n for $n \geq 0$ is the corresponding n-step transition probabilities. P is typical and default but not exclusive.
Probabilities	Let μ be a distribution on $\mathcal{S}$ and $s \in \mathcal{S}$. E_μ represents the expected value operator using μ for the initial distribution of the chain and E_s represents the expected value operator using a point mass at s for the initial distribution of the chain. Similarly, P_μ and P_s represent the corresponding probability measurs (on Ω).
Return Times etc.	$T_A = \inf \{n \geq 1 \colon X_n \in A\}$ is the *first return time* to A. Note that $T_A = \infty$ if X never returns to A. We write T_s when $A = \{s\}$. for $s \in \mathcal{S}$.

The distribution of T_A for different sets A is informative about the behavior of the chain. Define for $s \in \mathcal{S}$ and $A \subset \mathcal{S}$

$$\begin{aligned} R(s, A) &= \mathsf{P}_s\{T_A < \infty\} \\ M(s, A) &= \mathsf{E}_s T_A, \end{aligned}$$

the probability of return to A (starting at s) and the expected return time. Note that T_A is a stopping time.

$O_A = \sum_{n=1}^{\infty} 1\{X_n \in A\}$ is called the *occupation time* of A. Define for $s \in \mathcal{S}$ and $A \subset \mathcal{S}$

$$O(s, A) = \mathsf{E}_s O_A,$$

the expected occupation time of A starting at s.

We write $R(s, s'), M(s, s'), O(s, s')$, and so forth when $A = \{s'\}$ is a singleton.

Concept	Notation
Taboo Probabilities	$P^n_{!A}(s, B) = \mathsf{P}_s\{X_n \in B, T_A \geq n\}$. The ! stands for "not".
Past History	The history of the process up to time n is the σ-field $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$.

For a stopping time T, the history of the process "up to time T" is the σ-field $\mathcal{F}_T$ given by

$$\mathcal{F}_T = \{A \in \mathcal{F} \colon A \cap \{T = n\} \in \mathcal{F}_n, \text{ for all } n \in \mathbb{Z}_\oplus\}.$$

This contains all events for which, when $T = n$, we have can determine whether they occurred using the history up to time n.

Note that $\mathcal{F}_n \subset \mathcal{F}$ and $\mathcal{F}_T \subset \mathcal{F}$. These are collections of events.

Concept	Notation
Invariant Measures	π is a non-negative mesaure on $\mathcal{S}$ satisfying $\pi = \pi \cdot P$. If it is a probability measure, we call it an *invariant* or *stationary* distribution. π is default but not exclusive.

Summary 3. Countable-State Markov Chains: What We Know So Far (short form)

X is a countable-state Markov chain on $\mathcal{S}$ with initial distribution μ and transition probabilities P.

<u>Idea</u> — <u>Result</u>

Markov Property

Behavior of the chain is determined by initial distribution μ and transition probabilities P.

1. $\mathsf{P}\{X_0 = s_0, \ldots, X_n = s_n\} = \mu(s_0)P(s_0, s_1) \cdots P(s_{n-1}, s_n)$.
2. $\mathsf{P}\{X_n = s_n \mid X_{n-1} = s_{n-1}, \ldots, X_0 = s_0\} = P(s_{n-1}, s_n)$.
3. $\mathsf{E}_\mu(h(X_{n+1}, X_{n+2}, \ldots) \mid X_0, \ldots, X_n = s) = \mathsf{E}_s h(X_1, X_2, \ldots)$.

Strong Markov Property

$\mathsf{E}\left(h(X_{T+1}, X_{T+2}, \ldots) \mid \mathcal{F}_T\right) 1\{T < \infty\} = \mathsf{E}_{X_T}(h(X_1, X_2, \ldots) 1\{T < \infty\}$, for a stopping time T and bounded, measurable function h.

Chapman-Kolmogorov

$P^{n+m} = P^n \cdot P^m$ for any $n, m \geq 0$.

Irreducibility

The chain is irreducible if there is only one communicating class.

Periodicity

For a state $s \in \mathcal{S}$, the period of s is $d(s) = \gcd\{n \geq 1 \colon P^n(s, s) > 0\}$. This funciton is constant on communicating classes.

State Decompositions

The relation $\leftrightarrow$ is an equivalence relation, so $\mathcal{S}$ is a disjoint union of communicating classes. But we can go further.

1. $\mathcal{S} = D \cup \bigcup_i C_i$, where each C_i is an absorbing, communicating class with period d_i and where D is a union of non-absorbing communicating classes.
2. $C_i = \bigcup_{k=0}^{d_i - 1} U_{ik}$ where $P(s, U_{i,k+1 \bmod d_i}) = 1$ for $s \in U_{ik}$, $k = 0, \ldots, d_i - 1$.

This means we can study the behavior of the chain by understanding the irreducible, aperiodic case and the "transient" case.

Recurrence & Transience

A set of state $A \subset \mathcal{S}$ is called *recurrent* if $O(s, A) = \infty$ for all $s \in A$. A set of state $A \subset \mathcal{S}$ is called *transient* if $O(s, A) < \infty$ for all $s \in A$.

We have the following:

1. For any $s \in \mathcal{S}$, $O(s, s) = \infty \iff R(s, s) = 1$.
2. If X is irreducible, either $O(s, s') < \infty$ for all s, s' or $O(s, s') = \infty$ for all s, s'.
3. If X is irreducible, either $R(s, s') = 1$ for all s, s' or $R(s, s) < 1$ for all s.

Thus, an irreducible chain is either *transient* or *recurrent*.

continued...

Invariant Measures

If X is an irreducible, recurrent, countable-state Markov chain, then there exists a measure ρ on $\mathcal{S}$ that satisfies the following:

1. ρ is invariant; that is, $\rho = \rho \cdot P$.
2. ρ is everywhere positive; that is, $\rho(s) > 0$ for all s.
3. ρ is unique up to constant multiples.

The chain is called positive recurrent if $\rho(\mathcal{S}) = \sum_s \rho(s) < \infty$ and null recurrent if $\rho(\mathcal{S}) = \sum_s \rho(s) = \infty$.

Stationary Distribution I

Let X be irreducible.
X is positive recurrent if and only if $M(s,s) < \infty$ for all $s \in \mathcal{S}$.
Then, there exists a (positive) invariant measure ρ on $\mathcal{S}$ with $\rho(\mathcal{S}) < \infty$. Define $\pi(s) = \rho(s)/\rho(\mathcal{S})$ to get a probability distribution on $\mathcal{S}$.

Thus, X has a stationary distribution π if and only if $M(s,s) < \infty$ for all s. The stationary distribution satisfies the following:

1. It is unique.
2. For any $s \in \mathcal{S}$, $\pi(s) = \dfrac{1}{M(s,s)}$.

Stationary Distribution II

Let X be irreducible and aperiodic.
Then for any $s_0, s \in \mathcal{S}$,

$$\lim_{n\to\infty} P^n(s_0, s) = \frac{1}{M(s,s)}. \tag{4}$$

Hence, if X is positive recurrent with stationary distribution π,

$$\lim_{n\to\infty} P^n(s_0, s) = \pi(s), \tag{5}$$

and if X is transient or null-recurrent, $M(s,s) = \infty$, so $\lim_n P^n(s_0, s) = 0$.

If X is positive recurrent, then for any initial distribution μ:

$$\mathsf{P}_\mu\{X_n = s\} \to \pi(s), \tag{6}$$

In particular, if the chain is started with distribution π, it is stationary: $\mathsf{P}_\pi\{X_n = s\} = \pi(s)$.

For the rest of this class, let X be an irreducible, countable-state Markov chain on $\mathcal{S}$ with initial distribution μ and transition probabilities P.

Definition 4. An irreducible, aperiodic, positive recurrent Markov chain is called *ergodic*. An irreducible, aperiodic, null recurrent Markov chain is sometimes called *weakly ergodic*.

Proof of Selected Claims 5.
Claim 1. Either $O(s, s') < \infty$ for all s, s' or $O(s, s') = \infty$ for all s, s'.

Then since for any u, v, $u \to s$ and $s' \to v$, we can find an ℓ, m such that $P^{\ell}(u, s) > 0$ and $P^m(s', v) > 0$. Hence,

$$\sum_n P^{\ell+m+n}(u, v) > P^{\ell}(u, s) \left[\sum_n P^n(s, s') \right] P^m(s', v). \tag{7}$$

It follows that $O(s, s') = \sum_{n=1}^{\infty} P^n(s, s') = \infty$ implies $O(u, v) = \infty$ and $O(u, v) < \infty$ implies $O(s, s') < \infty$. But both pairs of states were arbitrary, so the claim is proved.

Claim 2. $O(s, s) = \infty \iff R(s, s) = 1$.

We can use the same trick we used for random walks early in the class. Notice that for any $s \in \mathcal{S}$,

$$P^n(s, s) = \sum_{k=1}^{n} \mathsf{P}_s\left\{ T_{\{s\}} = k \right\} P^{n-k}(s, s) = \sum_{k=0}^{n} \mathsf{P}_s\left\{ T_{\{s\}} = k \right\} P^{n-k}(s, s), \quad n \geq 1, \tag{8}$$

$$P^0(s, s) = 1, \tag{9}$$

where we've used $\mathsf{P}_s\{T_s = 0\} = 0$. This is called the first-entrance decomposition.

Define generating functions for the sequences $P^n(s, s)$ and $\mathsf{P}_s\{T_s = n\}$.

$$G_s(z) = \sum_n P^n(s, s) z^n \tag{10}$$

$$R_s(z) = \sum_n \mathsf{P}_s\{T_s = n\} z^n. \tag{11}$$

The above recurrence yields

$$G_s(z) = 1 + G_s(z) R_s(z) \tag{12}$$

which implies that

$$G_s(z) = \frac{1}{1 - R_s(z)}. \tag{13}$$

This last is a version of the "renewal equation."

Because $R_s(1) = \mathsf{P}_s\left\{ T_{\{s\}} < \infty \right\}$ (or letting $z \to 1$ to be careful about convergence), we get that $R(s, s) = 1 \iff O(s, s) = \infty$.

Actually, we can get more out of this equation. Suppose for now that $R_s(z)$ has a radius of convergence > 1.

Assume that $R_s(1) = 1$ (that is, $R(s, s) = 1$) and $0 < R_s'(1) < \infty$ (that is, $M(s, s) < \infty$). (The second assumption follows from the stronger assumption on radius of convergence.)

Then $G_s(z)$ has a pole (an isolated, polynomial-like singularity in the complex plane) at $z = 1$. Notice also that $(1 - z)G_s(z) \to 1/R'_s(1) = 1/M(s, s)$ as $z \to 1$. (L'Hopital's rule, for instance.) Take

$$\overline{G}(z) = \frac{R'_s(z)}{(M(s,s))^2(1-z)}, \tag{14}$$

which has $(1 - z)\overline{G}(z) \to -1/M(s, s)$ as $z \to 1$. Then,

$$U(z) = G_s(z) - \overline{G}(z) = \frac{1}{1 - R_s(z)} - \left(\frac{1}{M(s,s)}\right)^2 \frac{R'_s(z)}{1-z}. \tag{15}$$

satisfies $(1 - z)U(z) \to 0$ as $z \to 1$. As a result, we can take $U(z)$ to be analytic in a neighborhood of $z = 1$. (The apparent singularity at $z = 1$ is "removable".) That is,

$$G_s(z) = \overline{G}(z) + U(z). \tag{16}$$

If $R_s(z) = \sum_{n\geq 1} f_n z^n$, then

$$[z^n]\frac{R'_s(z)}{1-z} = \sum_{k=1}^{n} k f_k \to M(s,s) \quad \text{as } n \to \infty. \tag{17}$$

(Why are both these facts true?) Because $U(z)$ is analytic, $[z^n]U(z) \to 0$. Hence,

$$\lim_{n\to\infty} P^n(s,s) = \lim_{n\to\infty}[z^n]G_s(z) = \lim_{n\to\infty}[z^n]\overline{G}(z) = \left(\frac{1}{M(s,s)}\right)^2 M(s,s) = \frac{1}{M(s,s)}. \tag{18}$$

Claim 3. Either $R(s, s') = 1$ for all $s, s' \in \mathcal{S}$ or $R(s, s) < 1$ for all $s \in \mathcal{S}$.

Because of the relation with $O(s, s)$, either $R(s, s) < 1$ for all s or $R(s, s) = 1$ for all s. If the latter is true and $R(s, s') < 1$, then by irreducibility, we have $O(s', s) > 0$ and thus, for some n, $P^n_{!s'}(s', s) > 0$. This implies $R(s', s') < 1$, and the result follows by contradiction.

Claim 4. If X admits a positive stationary distribution π, then every state is recurrent.

Suppose $s \in \mathcal{S}$ is transient. This implies that $O(s, s) < \infty$, so $P^n(s, s) \to 0$ as $n \to \infty$. It follows then that

$$\begin{aligned}
\pi_s &= \sum_{s'} \pi_{s'} P(s', s) && (19)\\
&\leq \sum_{s'\in S_0} \pi_{s'} P(s', s) + \sum_{s'\notin S_0} \pi_{s'} && (20)\\
&\to \sum_{s'\notin S_0} \pi_{s'} && (21)\\
&\to 0, && (22)
\end{aligned}$$

for any finite $S_0 \subset \mathcal{S}$ with the last limit being as $S_0 \uparrow S$. This is a contradiction, so X must be recurrent.

Claim 5. Results on Stationary Distribution from Summary 3.

See the proofs in G&S section 6.4.

To translate between our notation and G&S notation, define $q(s, s') = \sum_n P^n_{!s}(s, s')$ for the taboo probabilies $P^n_{!s}(s, s')$. Notice that $M(s, s) = \sum_{s'} q(s, s')$.

G&S use $\rho_{s'}(s) \equiv q(s, s')$ and $f_{s,s'}(n) \equiv \mathsf{P}_s\{T_{s'} = n\}$.

Example 6. Binomial Runs

Consider a sequence $\Xi_1, \Xi_2, \ldots$ of Bernoulli$\langle p \rangle$ random variables.

Let $Z_0 = 0$ and define Z_n to be the length of the run of 1s looking back from time n. That is, for sequence 00110111010001111, we get for instance $Z_1 = 0$, $Z_2 = 0$, $Z_3 = 1$, $Z_4 = 2$, $Z_5 = 0$, $Z_8 = 3$, $Z_{13} = 0$, and $Z_{17} = 4$. In general, $Z_n = \max\{0 < k \le n : \Xi_n \cdots \Xi_{n-k+1} = 1\}$, with $Z_n = 0$ if the set is empty (i.e., $\Xi_n = 0$).

Is Z a Markov chain?

Yes. The Markov Property follows because the flips are independent as can be calculated directly.

What is the state space of Z?

$\mathcal{S} = \mathbb{Z}_\oplus$

What are the initial distribution μ and transition probabilities P?

By definition, $\mu(j) = \delta_{j0}$. For $j \ge 0$,

$$P(j, j+1) = p \tag{23}$$

$$P(j, 0) = 1 - p \equiv q. \tag{24}$$

What are the communicating classes? Is Z irreducible?

Z is irreducible, meaning there is only one communicating class.

To see that $P^n(i,j) > 0$, it suffices to find an n-step path with non-zero probability. If $j > i$, choose the path is $i \to i+1 \to \cdots \to j$, which has probability $p^{j-i} > 0$. If $j < i$, choose the path $i \to 0 \to 1 \cdots \to j$, which has probability $qp^j > 0$. It follows that $i \leftrightarrow j$ for every $i, j \in \mathbb{Z}_\oplus$.

What are the periods of the communicating classes?

Because Z is irreducible, it suffices to find the period for one state. But $P(0,0) = q$, so the period is 1. Note that for a state $i > 0$, $P^n(i,i) > 0$ for all $n > i$. Since two primes are present in that list, $d(i) = 1$ as well.

Is Z recurrent?

For $i \ge 0$,

$$O(i,0) = \sum_n P^n(i,0) \ge \sum_n P^{n-1}(i,\mathcal{S})q = \sum_n q = \infty. \tag{25}$$

It follows that Z is recurrent.

Is Z positive recurrent?

There are several approaches here. One is to compute $M(i,i)$ for each $i \ge 0$. This will, at the same time, give us the stationary distribution if one exists. Another is to solve for an invariant measure.

Let's try the second one. If $\rho = \rho P$, then we have for $j > 0$,

$$\rho(0) = \sum_{i \ge 0} \rho(i) q = q \tag{26}$$

$$\rho(j) = \sum_{i \ge 0} \rho(i) P(i,j) = p\rho(j-1). \tag{27}$$

It follows that we can take $\rho(j) = p^j(1-p)$. So Z is positive recurrent.

What is it's stationary distribution?

From the last calculation, we have $\pi = \rho$ since ρ is a probability measure. Hence, $\pi(j) = p^j(1-p)$ is the stationary distribution.

For "free," we get that $M(j,j) = p^{-j}(1-p)^{-1}$.

Example 7. Walk on a Finite, Bidirectional Graph

A graph is a collection of vertices joined by edges. Let $\mathcal{V} = \{v_1, \ldots, v_m\}$ be a collection of vertices. Let $\mathcal{E}(v, v')$ be 1 if there is an edge in the graph between v and v' and 0 otherwise. Assume $\mathcal{E}(v, v') = \mathcal{E}(v', v)$.

Let μ be a probability distribution on $\mathcal{V}$. And suppose that Y_0 has distribution μ.

Let P be a transition probability matrix on $\mathcal{V}$ that satisfies $P(v, v') = 0$ if and only if $\mathcal{E}(v, v') = 0$. Given $Y_{n-1} = v$, let Y_n have conditional distribution $P(v, \cdot)$. Then, Y is a Markov chain.

What is the state space of Y?

$\mathcal{S} = \mathcal{V}$

What are the communicating classes? Which classes are absorbing? Under what conditions is Y irreducible?

On each connected component of the graph, there is a path in the graph between v and v'. If it is $v \to u_1 \to \cdots \to u_{k-1} \to v'$, then, $P^k(v, v') \geq P(v, u_1)P(u_1, u_2) \cdots P(u_{k-1}, v') > 0$. Across components there is no such path, so $P(v, v') = 0$ by construction. Hence, each connected component is a communicating class. Since $P(v, \mathcal{V} - C) = 0$ for every class C, every class is absorbing. Y is irreducible if and only if the graph is connected.

Suppose $P(v, v') = \mathcal{E}(v, v') / \sum_{i=1}^m \mathcal{E}(v, v_i)$. *Describe the long-run behavior of the chain.*

The definition of P shows that for vertices in one connected component, all non-zero transition probabilities out of that vertex are equal.

If $C_1, \ldots, C_r$ are the connected components of the graph, then $X_1 \in C_i$ with probability $\mu(C_i)$, and henceforth, it proceeds as an ireducible chain restricted to C_i.

Without knowing more about the edges, we cannot know the periodicity, but define d_i to be the period of C_i.

Now, restrict our attention to a single communicating class, which is equivalent to considering a connected graph.

So for the moment, treat $\mathcal{V}$ as connected. Define

$$\pi(v) = \frac{\sum_{i=1}^m \mathcal{E}(v, v_i)}{\sum_{i,j=1}^m \mathcal{E}(v_j, v_i)}, \tag{28}$$

for all v in the component. Then,

$$\sum_{j=1}^m \pi(v_j) P(v_j, v') = \sum_{j=1}^m \frac{\sum_{i=1}^m \mathcal{E}(v_j, v_i)}{\sum_{i,k=1}^m \mathcal{E}(v_k, v_i)} \frac{\mathcal{E}(v_j, v')}{\sum_{i=1}^m \mathcal{E}(v_j, v_i)} \tag{29}$$

$$= \frac{\sum_{j=1}^m \mathcal{E}(v_j, v')}{\sum_{i,j=1}^m \mathcal{E}(v_j, v_i)} \tag{30}$$

$$= \frac{\sum_{j=1}^m \mathcal{E}(v', v_j)}{\sum_{i,j=1}^m \mathcal{E}(v_j, v_i)} \tag{31}$$

$$= \pi(v'), \tag{32}$$

where the second-to-last step follows from the bi-directionality of the graph. This is a stationary distribution, so the chain is positive recurrent. What's the intuition for this stationary distribution?

This gives us the long-term behavior of the chain for a general, bidirectional graph.

Example 8. Walk on a Finite, Directional Graph

As before, let $\mathcal{V} = \{v_1, \ldots, v_m\}$ be a collection of vertices. Let $\mathcal{E}(v, v')$ 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 $\mathcal{E}(v, v') = \mathcal{E}(v', v)$.

Let μ be an initial distribution on $\mathcal{V}$. Let P be a transition probability matrix on $\mathcal{V}$ that satisfies $P(v, v') = 0$ if and only if $\mathcal{E}(v, v') = 0$. 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 μ_n be the distribution of Y_n.

$$\mu_n(v) = \mu(v)1_{(n=0)} + \sum_{v'} \mu_{n-1}(v')P(v', v)1_{(n>0)}. \tag{33}$$

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

$$\sum_n \mu_n z^n = \mu + \sum_{v'} \sum_n \mu_{n-1}(v') z^n P(v', \cdot) \tag{34}$$

$$\boldsymbol{G}(z) = \mu + z\boldsymbol{G}(z) \cdot P. \tag{35}$$

So,

$$\boldsymbol{G}(z)(I - zP) = \mu \tag{36}$$

$$\boldsymbol{G}(z) = \mu(I - zP)^{-1} = \sum_{n\geq 0} z^n \mu P^n, \tag{37}$$

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 $O(s, \mathcal{S}) = \infty$ implies that the chain is recurrent because the sum $O(s, \mathcal{S}) = \sum_{s'} O(s, s')$ 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 π corresponds to a search for eigenvectors of P with eigenvalue 1, vectors π such that $\pi(I - P) = 0$.

Example 9. Finite, Doubly Stochastic Chain

Suppose $U = (U_n)_{n\geq 0}$ is an irreducible Markov chain on a finite state-space $\mathcal{S}$ 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 π is then just a uniform distribution on $\mathcal{S}$.

Example 10. Renewal Process and Forward Recurrence Time Chain

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

$$Z_n = Z_0 + \sum_{k=1}^{n} \Xi_k \tag{38}$$

for IID random variables Ξ_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 Ξ 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 nth replacement – a "renewal" of the system.

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

Suppose now that Z_0 and Ξ_1 take values in $\mathbb{Z}_+$ and have PMFs $\mu \equiv \mathsf{p}_{Z_0}$ and $\mathsf{p}_\Xi \equiv \mathsf{p}_{\Xi_1}$. Z_n is thus a countable-state Markov chain, but not a terribly interesting one as it marches inexorably toward ∞.

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 $n \geq 0$:

$$V_n^+ = \inf\{Z_m - n \colon Z_m > n\} \tag{39}$$

$$V_n^- = \inf\{n - Z_m \colon Z_m \leq n\}. \tag{40}$$

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 $\mathbb{Z}_\oplus$.

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 ξ. Hence,

$$P(k, k-1) = 1 \quad \text{for } k > 1, \tag{41}$$

$$P(1, k) = \xi(k) \quad \text{for } k \in \mathbb{Z}_+. \tag{42}$$

For V^-, we can reason similarly. Let S_Ξ be the survival function of Ξ_1. Then,

$$P(k, k+1) = \mathsf{P}\{\Xi > k+1 \mid \Xi > k\} = \frac{\mathsf{S}_\Xi(k+1)}{\mathsf{S}_\Xi(k)} \tag{43}$$

$$P(k, 0) = \mathsf{P}\{\Xi = k+1 \mid \Xi > k\} = \frac{\mathsf{p}_\Xi(k+1)}{\mathsf{S}_\Xi(k)}. \tag{44}$$

Is V^+ irreducible? Is it recurrent?

If there exists an $M \in \mathbb{Z}_+$ such that $\mathsf{S}_\Xi(M) = 0$ and $\mathsf{p}_\Xi(M) > 0$, then all states $j > M$ are transient, and all states $\{1, \dots, M\}$ 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 $n > 1$, $P_{!1}^{n-1}(n, 1) = 1$. Hence,

$$R(1,1) = \sum_{n \geq 1} \mathsf{p}_\Xi(n) P_{!1}^{n-1}(n,1) = 1. \tag{45}$$

So the chain is recurrent in this case as well.

What is the long-run behavior of the chain?

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

$$\rho(j) = \sum_{n \geq 1} \mathsf{p}_\Xi(j + n - 1) = \sum_{n \geq j} \mathsf{p}_\Xi(n) = \mathsf{S}_\Xi(j - 1). \tag{46}$$

Notice that

$$\sum_{j \geq 1} \rho(j) P(j,k) = \rho(1)\mathsf{p}_\Xi(k) + \rho(k+1) \tag{47}$$

$$= \mathsf{p}_\Xi(k) + \mathsf{S}_\Xi(k) \tag{48}$$

$$= \rho(k). \tag{49}$$

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

$$\sum_{n \geq 1} \rho(n) = \sum_{n \geq 1} \mathsf{S}_\Xi(n-1) = \sum_{n \geq 1} n \mathsf{p}_\Xi(n) = \mathsf{E}\Xi_1 < \infty. \tag{50}$$

In this case, $\pi(k) = \rho(k)/\mathsf{E}\Xi_1$ is a stationary distribution.

Idea 11. The above argument leads to an interesting idea for the countable case.

Suppose that X is recurrent, Pick a state $s_0 \in \mathcal{S}$ such that it is easy to compute $P_{!s_0}^n(s_0, s)$ for any $s \in \mathcal{S}$. Define

$$\rho(s) = \sum_{n \geq 1} P_{!s_0}^n(s_0, s). \tag{51}$$

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

$$\sum_{s \in \mathcal{S}} \rho(s) P(s, s') = \rho(s_0) P(s_0, s') + \sum_{s \neq s_0} \sum_{n \geq 1} P_{!s_0}^n(s_0, s) P(s, s') \tag{52}$$

$$= P(s_0, s') + \sum_{n \geq 2} P_{!s_0}^n(s_0, s') \tag{53}$$

$$= P_{!s_0}(s_0, s') + \sum_{n \geq 2} P_{!s_0}^n(s_0, s') \tag{54}$$

$$= \rho(s'). \tag{55}$$

This is finite only if

$$\rho(\mathcal{S}) = \sum_{s \in \mathcal{S}} \sum_{n \geq 1} P^n_{!s_0}(s_0, s) \tag{56}$$

$$= \sum_{n \geq 1} \sum_{s \in \mathcal{S}} P^n_{!s_0}(s_0, s) \tag{57}$$

$$= \sum_{n \geq 1} P^n_{!s_0}(s_0, \mathcal{S}) \tag{58}$$

$$= \sum_{n \geq 1} \mathsf{P}_{s_0}\{T_{s_0} \geq n\} \tag{59}$$

$$= \mathsf{E}_{s_0} T_{s_0} \tag{60}$$

$$< \infty, \tag{61}$$

which is just the positive recurrence condition.

Hence,

$$\pi(s) = \frac{\sum_{n \geq 1} P^n_{!s_0}(s_0, s)}{\sum_{n \geq 1} P^n_{!s_0}(s_0, \mathcal{S})} = \frac{1}{\mathsf{E}_{s_0} T_{s_0}}. \tag{62}$$

We already knew the last equality, but this gives a new way of finding a stationary distribution, using equation (51).