

Countable-state Markov chains

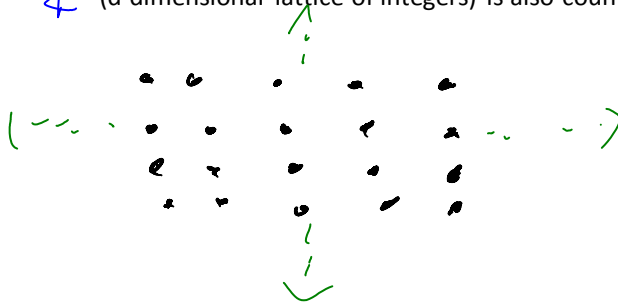
Thursday, October 09, 2008
12:00 PM

- References: [Lawler Ch. 2](#)
[Karlin & Taylor Chs. 2 & 3](#)
[Resnick Ch. 1](#)

Homework 2 due at 12 PM on Tuesday, October 14. Next class is also at 12 PM on October 14.

By a **countable** state space S , we mean that the elements of S can be put into a one-to-one correspondence with some subset of the integers. In other words, one can index all possible states using the integers or some subset thereof.

- Note that \mathbb{Z}^d (d-dimensional lattice of integers) is also countable



In practice, countable-state models exclude state spaces that involve intervals of the real line. (Usually use stochastic differential equations or stationary processes as stochastic models in continuous state space.)

Some typical examples of countable state space Markov chain models:

- Population and epidemic modeling
- Random walks on lattices
- Queuing models (w/o preset upper limit)
- Capital or asset price models (more natural SDE's)

Countable state models are typically used when there is no clear way to set a precise upper limit to the state values of the system. Even if realistically the system cannot take on very large state values, the Markov chain dynamics themselves should make these large state values rare rather than using an upper limit to prevent the state values from becoming large. And if the Markov chain does try to take on arbitrarily large values, that signals information about what to expect in practice (state value will grow in the long run).

What results from our analysis of finite-state Markov chains requires modification for the infinite-state case?

Finite-time formulas

The probability transition matrix now has an infinite number of entries so one has to be a little careful with the matrix manipulations.

$$P_{ij} = P(X_{n+1} = j | X_n = i)$$
$$P^{(n)} = P(X_n = i)$$

$$\begin{aligned} \phi_j^{(n)} &= P(X_n = j) \\ \phi_j^{(0)} &= P(X_0 = j) \\ \vec{\phi}^{(n)} &= \vec{\phi}^{(0)} \cdot P^n \end{aligned}$$

What does mean when the vector and matrix have infinite numbers of entries? Just have infinite sums.

$$\begin{aligned} (P^2)_{ij} &= \sum_{k \in S} P_{ik} P_{kj} \quad \text{etc.} \\ (\vec{\phi}^{(0)} P^n)_j &= \sum_{k \in S} \phi_k^{(0)} P_{kj} \end{aligned}$$

So the main concern is whether these sums converge. In fact they always do in these calculations because:

$$\begin{aligned} \|\vec{\phi}^{(0)}\|_{l^1} &= \sum_{j \in S} \phi_j^{(0)} = 1 \\ \|P\|_{l^\infty} &= \sup_{i,j \in S} |P_{ij}| \leq 1 \\ \|P\|_{l^1} &= \sup_{i \in S} \sum_{j \in S} |P_{ij}| = 1 \end{aligned}$$

Now one can use inequalities from analysis to show:

$$\left| \sum_{k \in S} P_{ik} P_{kj} \right| \leq \|P\|_{l^1} \|P\|_{l^\infty} = 1$$

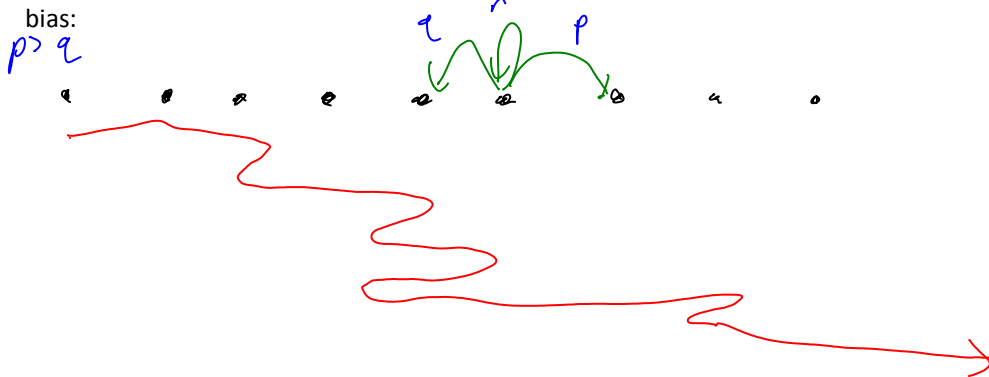
Similarly
$$\left| \sum_{k \in S} \phi_k^{(0)} (P_{kj})^n \right| \leq \|\vec{\phi}^{(0)}\|_{l^1} \|P\|_{l^\infty}^n = 1$$

So the finite-time formulas carry over to countable state case with no problem.

Long-time properties

Here new phenomena arise for the countable state case. In particular the classification of states in terms of recurrent and transient must be revisited.

Transience: For finite state Markov chains, no closed communication class can be transient. But for infinite state Markov chains, it is possible to have closed transient classes and even irreducible Markov chains that are transient. Classical example would be a random walk with a



The random walker drifts off to infinity, eventually doesn't return. So we will need to develop a tool beyond just looking at Markov chain topology to determine whether states are transient. For example for the random walker, it is recurrent without bias ($p=q$) and transient with bias ($p>q$).

Recurrence: Now there are two types of recurrence with qualitatively different behavior:

Positive recurrence: A state j is said to be positive recurrent provided that

$$E(\tau_j(1) | X_0 = j) < \infty$$

where $\tau_j(1) = \min\{n > 0; X_n = j\}$

Null recurrence: A state j is null recurrent provided that:

$$P(\tau_j(1) < \infty | X_0 = j) = 1$$

but

$$E(\tau_j(1) | X_0 = j) = \infty$$

To see why these cases can arise, consider that

$$P(\tau_j(1) < \infty | X_0 = j) = P\left(\bigsqcup_{n=1}^{\infty} \{\tau_j(1) = n\} | X_0 = j\right)$$

$$= \sum_{n=1}^{\infty} P(\tau_j(1) = n | X_0 = j) = 1$$

for recurrent states

$$E(\tau_j(1) | X_0 = j) = \sum_{n=1}^{\infty} n P(\tau_j(1) = n | X_0 = j)$$

$= \infty$ for null recurrent states
because

$P(T_j(1) = n | X_0 = j)$ decays
slowly with n

Null recurrence is not possible for a finite state space.

Positive recurrence and null recurrence and transience are class properties for countable-state Markov chains, meaning that all states in a communication class have the same transience/recurrence property.

Two fundamental questions:

- How do Markov chains with these various properties behave?
- How do you determine the transience/recurrence property of a class?

For this discussion, we will restrict attention to irreducible Markov chains because the results carry over to reducible Markov chains:

- closed classes act like their own irreducible Markov chain
- A class which is not closed is automatically transient.

Positive recurrent Markov chains:

These behave very much like irreducible finite-state Markov chains. In particular, they have a unique stationary distribution

$$\begin{aligned}\sum_{j \in S} \pi_j &= 1 \\ \pi_j &\geq 0 \\ \vec{\pi} \cdot P &= \vec{\pi}\end{aligned}$$

Law of large numbers applies as before.

And if the Markov chain is also aperiodic, then the stationary distribution again serves as a limit distribution

$$\begin{aligned}\lim_{n \rightarrow \infty} P(X_n = j) &= \pi_j \\ \lim_{n \rightarrow \infty} \vec{v} \cdot P^n &= \vec{\pi}\end{aligned}$$

To prove this, see the development in [Resnick Secs. 2.12&2.13](#), some of which we carried out in class. This was based on constructing a hypothetical stationary distribution in terms of first passage times. The most interesting aspect of the proof is the **coupling** argument used to prove that the stationary distribution is a limit distribution when Markov chain is aperiodic. ([Resnick 2.13](#))

Null recurrent Markov chains:

These are a little unusual but do appear in some natural contexts such as unbiased random walks in one or two dimensions.

Null recurrent Markov chains are guaranteed to have an invariant measure \vec{r}

$$\vec{r} \cdot P = \vec{r}$$
$$r_j \geq 0$$

but **not** a stationary distribution. That is, the invariant measure corresponding to a null recurrent Markov chain cannot be normalized

$$\left(\sum_{j \in S} r_j = \infty \right)$$

Invariant measure is unique up to constant multiple. (Proof in Resnick Sec. 2.12).

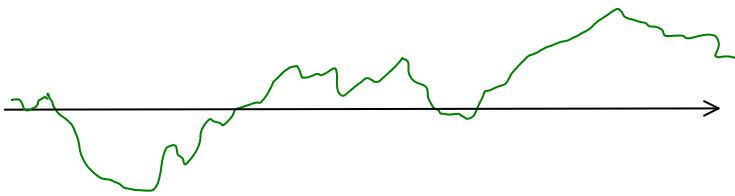
Long-time behavior (Resnick Sec. 2.12, 2.13)

For each $i, j \in S$, $\lim_{n \rightarrow \infty} (P^n)_{ij} = 0$

$$\lim_{n \rightarrow \infty} P(X_n = j | X_0 = i) = 0 \quad \text{for each } i, j \in S$$

Since the Markov chain is recurrent, each state will be visited in the future with probability 1, but if one looks at any given time at a given state, the probability to see the Markov chain in that state becomes small at large times.

What would this mean if one observed a null recurrent chain over a finite time horizon, i.e. with numerical simulation. For example unbiased random walk in one dimension.



See Feller, *An Introduction to Probability Theory and its Applications*, Ch. 3

Transient Markov chain:

$$P(T_j(1) < \infty | X_0 = j) < 1$$

$$T_j = \sum_{n=0}^{\infty} I\{X_n = j\} =$$

number of times state j is visited

has a geometric probability distribution

if $\sum_0=j$)
 and in any case has finite mean.
 And as a consequence, $\lim_{n \rightarrow \infty} P(X_n=j) = 0$.

Transient countable-state Markov chains can be analyzed just the same way as finite-state Markov chains, using absorption probability formulas and cost/reward formulas, simply working now with possibly infinite sums, but these also are guaranteed to converge.

Summary of properties of classes of Markov chains of various types:

- Positive recurrent classes: work with stationary distribution and compute long-time properties using it, very much like finite-state irreducible Markov chains.
- Null recurrent classes: special properties, borderline between transient and positive recurrent, hard to compute much in general
- Transient classes: same type of analysis as for finite-state Markov chain transient classes.

Now, given a communication class in a Markov chain, how do I determine whether it is positive recurrent, null recurrent, or transient?

Topology alone is not enough, as it was for finite-state case.

One method that works well for Markov chains where you can calculate or estimate P^n for large n , such as for random walks

- Karlin and Taylor Sec. 2.5
- Lawler Sec. 2.2
- Resnick Sec. 2.6

By using recursion relations involving first-passage time distributions and their connections to transitions over finite time horizons, one obtains the following results:

- A state i is transient if and only if $\sum_{n=1}^{\infty} (P^n)_{ii} < \infty$
- A state i is recurrent if and only if $\sum_{n=1}^{\infty} (P^n)_{ii} = \infty$
 \Downarrow
 $P(X_n=i | X_0=i)$