Countable state DTMC

Thursday, October 22, 2015
2:06 PM

Reading:

- Lawler Ch. 2
- Karlin & Taylor Chs 2 & 3
- Resnick Ch. 1

No class or office hours next week; next class is Monday, November 2.

Homework 3 will be posted by early next week; due mid-November.

Let's now extend our considerations of Markov chains to allow a countably infinite (but still discrete) state space, such as $\mathbb{Z}, \mathbb{Z}^d, \mathbb{Z}_{\geq 0}$, etc. We are not going to cover continuous state spaces in this class.

Why should we care to use countable state spaces?

- queueing theory without a maximum capacity
- population and disease modeling
- spatial dynamics in extended media (atomic lattices)

Even though any such system will really have some kind of upper limit, the upper limit may not be clearly well defined so we leave it off and just allow the stochastic dynamic rules to create statistical limits, and/or interpret a Markov chain whose state runs off to infinity in an appropriate way.

Model formulation is essentially the same.

One can still define stochastic update rules in the same way as before.

And it is still sufficient to characterize the Markov chain by the probability transition matrix(ces)

$P_{ij}^{(n)} = P(X_{n+1} = j | X_n = i)$ for $i, j \in S$.

The only difference is that this matrix now has ∞ many rows and columns. This doesn't wind up creating much difficulty, usually one simply defines the matrix by giving a rule for how to compute $P_{ij}^{(n)}$ for any $i, j \in S$ and $n \geq 0$.

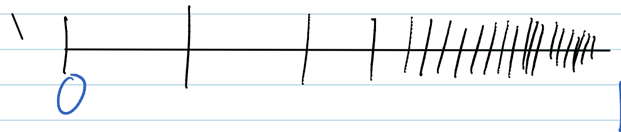
Similarly the initial probability distribution for the Markov chain will be a vector $\phi^{(0)} = \begin{pmatrix} \phi_1^{(0)} \\ \phi_2^{(0)} \\ \vdots \\ \vdots \end{pmatrix} = \{\phi_i^{(0)}\}_{i \in S}$

where $\phi_i^{(0)} = P(X_0 = i)$.

Simulation of countable state Markov chain

If one has a nice stochastic update rule, then simulation proceeds in the same way as before.

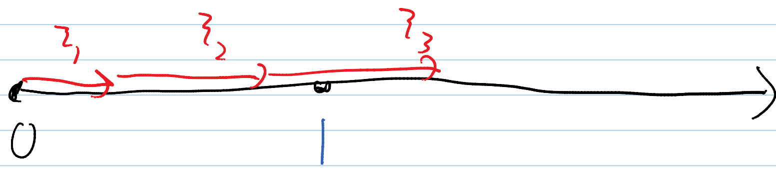
If one is using the probability transition matrix, then in principle (actually this can also happen with the stochastic update rule), one may need to simulate a random variable with an infinite countable state space.



One can extend the simulation procedure for finite random variables to random variables with countable state space, though it can be a bit more awkward/tedious to determine which bin the $U(0,1)$ falls in.

Two things to keep in mind:
- in simulating countable state Markov chains, you often only need to simulate a finite random variable to update the Markov chain
- perhaps the most common random variable with an infinite range that needs to be simulated is the Poisson random variable. Many software packages allow to simulate this directly. Or, one can use the trick:

Suppose we want to simulate a Poisson random variable Y with mean μ.



Generate a stream of exponentially distribution random variables which are iid with mean $1/\mu$: $\{\xi_j\}$. Then

$$Y \sim \max\left(k: \sum_{j=1}^{k} \xi_j < 1\right)$$

How does one simulate exponentially distributed random variables? The inverse transform method tells you that you can simulate ξ by simulating $U \sim U(0,1)$ and mapping it as follows

$$\xi = -\mu \ln U$$

inverse of CDF
of the exponential

Finite horizon statistics

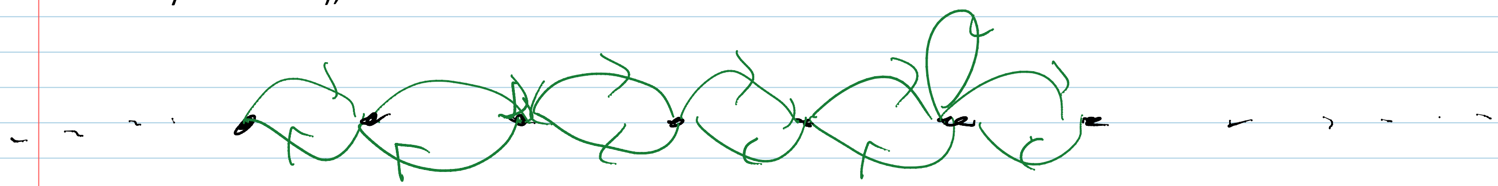
Everything works the same way. The key formulas
for the probability distribution $\phi_j^{(n)} = P(X_n = j)$

was: $\phi^{(n)T} = \phi^{(0)T} P^n$

This matrix-vector multiplication now involves infinite sums, but they are all guaranteed to converge absolutely by elementary Hölder inequalities because the probability distributions must add up to 1, and the row sums add up to 1, and all entries are bounded by 1.

Long Time Properties of Countable State Markov Chains

While any irreducible finite state Markov chain had to be recurrent (by a variation of the Infinite Monkeys Theorem), an irreducible countable state Markov chain need not be.

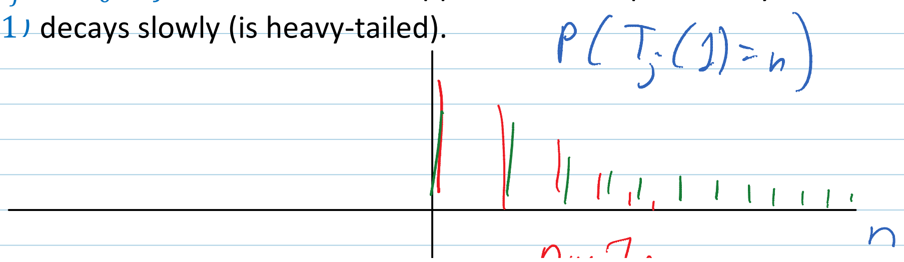


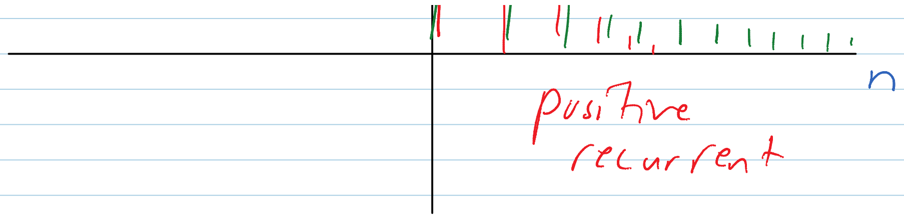
The state can run away to infinity and never come back, particularly if there is some "drift." So we need to revisit the classification procedure.

In fact, the classification procedure needs some refinement.

As before, we say that a state j is transient if and only if $P\,(T_j(1) < \infty | X_0 = j) < 1$, where $T_j(1) = \min\{n > 0: X_n = j\}$, and otherwise we say the state i is recurrent.
In fact, we will now refine the notion of recurrence:

- State j is said to be positive recurrent if $\mathbb{E}(T_j(1)|X_0 = j) < \infty$
- A state which is recurrent but not positive recurrent, is said to be null recurrent. That is, such a state j satisfies: $P\,(T_j(1) < \infty|X_0 = j) = 1$, but $\mathbb{E}(T_j(1)|X_0 = j) = \infty$. This can happen when the probability mass function for $T_j(1)$ decays slowly (is heavy-tailed).





$$\mathbb{E}\, T_j(1) = \sum_{n=1}^{\infty} n\, P(T_j(1)=n)$$

You might think that null recurrence is a pathological situation, but it's not. For standard random walks, one can show by direct calculation (see Karlin and Taylor) that the random walk is null recurrent in 1 and 2 dimensions, and transient in 3 and higher dimensions. This calculation method doesn't generalize well in practice to classify arbitrary Markov chains.

One can show that transience, null recurrence, and positive recurrence are all class properties still. So we can again decompose (in a matter to be described later) reducible Markov chains into transient classes, null recurrent classes, and positive recurrent classes.

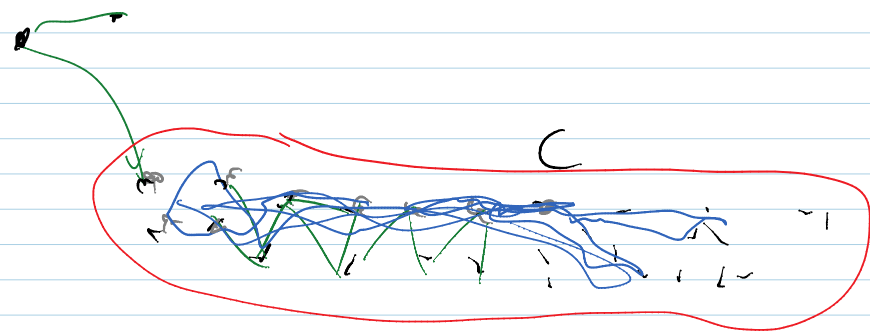
It is still true that recurrent classes must be closed (but not vice versa).

So let's first consider what happens once a Markov chain enters a closed class.
Positive recurrent class

One can show (with the same proof as in Resnick Sec. 2.12) that positive recurrent classes C have a well-defined stationary distribution associated with them:

$\sum_{j\in C}\pi_j = 1, \pi_j \geq 0$, and $\pi^{(T)} = \pi^{(T)}P$ where P is the restriction of the probability transition matrix to the positive recurrent class C. This stationary distribution can be used along with the LLN of MC to characterize long-term behavior within the positive recurrent class, and the stationary distribution is the limit distribution if the positive recurrent class is also aperiodic.
See Resnick Sec. 2.12, 2.13 for proofs.

The basic idea is that positive recurrent classes of countable state Markov chains behave in an effectively "finite" manner.

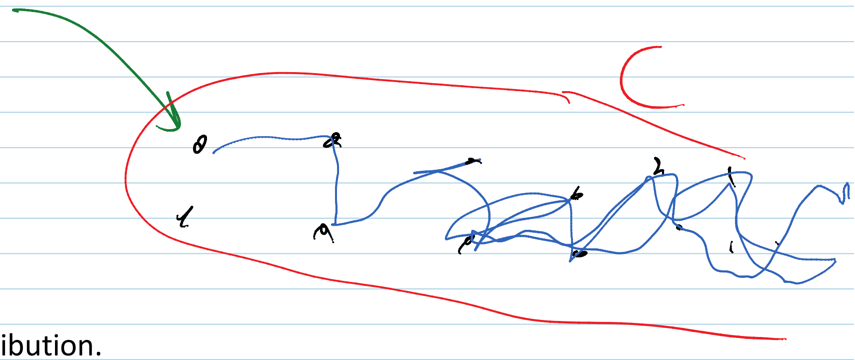


Null recurrent classes

These classes do not have a stationary distribution (it's impossible). If you try to solve:

$\nu^{(T)}P = \nu^{(T)}$

you will be able to find a solution that satisfies $\nu_j \geq 0$ for all $j \in C$, but you can't normalize it to a stationary distribution because $\sum_{j \in C} \nu_j = \infty$. Such non-normalized solutions are called an invariant measure, which have some use in more advanced applications but don't tell us much at an elementary level. (Resnick Sec. 2.12)



Long-time property? No LLN, no limit distribution.

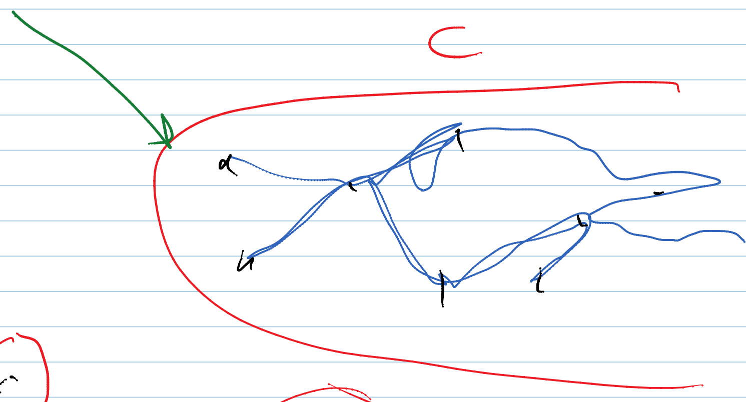
Rather, for any $i, j \in C$: $\lim_{n\to\infty} P(X_n = j | X_0 = i) = 0$.

Transient classes

Now we need to distinguish between closed transient classes and non-closed transient classes.

Just as for null recurrent classes, we have the result:

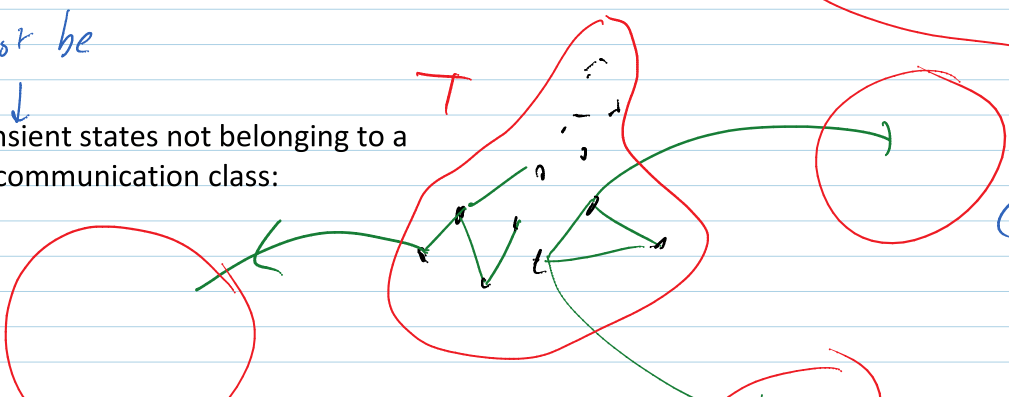
$$\lim_{n\to\infty} P(X_n = j | X_0 = i) = 0$$



Not much more to say or analyze quantitatively bout closed transient classes, up to a point we will return to momentarily.

must be

For transient states not belonging to a closed communication class:



One analyzes these transient states in a similar manner as for finite-state Markov chains, namely:

- compute probabilities for absorption in one of the closed communication classes (even for transient closed communication classes)
  - But the sum of the absorption probabilities may not add up to 1; the defect is the probability to get lost forever in the transient states that don't belong to any closed communication class
- compute expected cost/reward until absorption by a closed communication class
  - One has to be aware that one can get infinite answers since the time of absorption into a closed communication class can be infinite.

One can still do the various modification tricks to Markov chains to study quantitative behavior concerning the behavior of the Markov chain between visits to 2 or 3 states in closed communication classes (whatever type they are).

Classification of Countable State Markov Chains

Here is a systematic procedure:

1. Decompose the Markov chain into communication classes, and decide the properties of whatever communication classes can be determined from topology alone.
   - Any non-closed communication class must be transient.
   - Any closed communication class with finitely many states must be positive recurrent.
2. This only leaves closed communication classes with infinitely many states. Treat each such class as an irreducible MC in its own right for the purpose of the following analysis. Two parallel tracks which can be followed in either order:
   a. Look for an invariant measure ($\nu^{(T)} P = \nu$ with $\nu^{(T)} \geq 0$)
      i. If you can find an invariant measure that can be normalized ($\sum_{j\in S} \nu_j < \infty$) into a stationary distribution, the MC must be positive recurrent.
      ii. If you can show that no invariant measure exists, then the MC must be transient.
      iii. If you find an invariant measure that can't be normalized, then you only know the class is transient or null recurrent. Have to go to test b.
   b. Decisive test for transience. Choose any convenient reference state i_* and form the matrix Q by deleting the corresponding row and column from the probability transition matrix. Look for solutions to the equation $x = Q\,x$.
      i. If you find a nonzero, nonnegative, bounded solution, then the MC must

be transient.

ii. If you find a solution satisfying $x_j \to \infty$ for $|j| \to \infty$, then the MC must be recurrent. Corrected on 11/02/15

iii. If you can prove that the only nonnegative, bounded solution is 0, then the MC must be recurrent.

We will discuss why this classification scheme works in the next lecture.

1) A closed communication class has a well-defined stationary distribution if and only if the class is positive recurrent. This is a decisive test for positive recurrence.