Homework 2 due Friday, March 7.

Last time we showed how to prove existence of stationary distributions for irreducible finite state Markov chains.

The proofs of the statements about uniqueness and limit distributions can be found in Resnick Sec. 2.13. These proofs are fairly uninteresting with the exception of the coupling argument.

We've discussed how stationary distributions can be used to calculate long-term properties of irreducible FSDT Markov chains, and also long-term properties of FSDT Markov chains that aren't irreducible but do have a single closed communication class.

The ideas of stationary distributions can also be extended simply to Markov chains that are reducible (not irreducible; some states don’t communicate) if the Markov chain can be expressed as a union of closed communication classes (i.e., the letters S O C in our previous example). In this case, one simply needs to consider how much of the initial probability distribution falls within each of the closed communication classes. Those probabilities to be in a communication class remain constant because those communication classes are closed. Conditioned on being in one of these communication classes, the probability distribution of the state is given by the stationary distribution associated to that closed communication class.

But stationary distributions do not fully describe the long-term properties of Markov chains which have more than one closed communication class and have some communication classes which are connected to each other only in one direction (so one class is accessible from another, but not vice versa).

We can depict the relationship between the communication classes as follows:

\[ C_1 \rightarrow C_2, \quad C_1 \rightarrow C_3, \quad C_1 \rightarrow C_4 \]

\[ C_4 \rightarrow C_3, \quad C_4 \rightarrow C_5 \]
One could construct (and this is often useful) stationary distributions within the closed communication classes, which will be informative about the long-run properties of realizations that eventually enter those closed communication classes. But if one starts in one of the non-closed communication class, then which closed communication class is eventually entered? Need concepts beyond stationary distribution to determine this.

Practical applications that have topology like this:
- Spread of a disease in a naïve population
  - closed communication classes could correspond to quick control of disease, or a widespread epidemic
- Ecosystem and environmental modeling
  - closed communication classes might represent sustainable vitality or permanent disturbance or damage
- Economic modeling for emerging economies or businesses
  - closed communication classes could be remaining a small fringe business, collapsing, grabbing a permanent large market share
- Material response with possibility of "catastrophic" result
  - closed communication classes could be failure or strained but held configuration

To present the results for how to approach such systems, we introduce some definitions.

A state \( j \) of a Markov chain is said to be **recurrent** if the Markov chain, starting from state \( j \), has probability 1 to return to state \( j \).

\[ P(T_j(1) < \infty | X_0 = j) = 1. \]

Any state that is not recurrent is called **transient**.

**Recurrence and transience are class properties** (same for all states in a communication class).

Any communication class that is not closed is transient.

Any finite state communication class that is closed is recurrent by the coin flipping argument.

Therefore, determining transience and recurrence for FSDT MCs is a purely topological exercise. This is not the case for countable state MCs.

One can also show that any **recurrent state** has probability 1 to be
visited infinitely many times, if the Markov chain visits that state at all. On the other hand transient states will, with probability 1, only be visited finitely many times (or not at all), and governed by a geometric distribution.

In presenting the results concerning the long-time properties of reducible Markov chains, it is useful to refer to a canonical decomposition of the state space.

\[ S = T \cup C_k \ (\text{disjoint}) \]

\[ \uparrow \quad \uparrow \]

transient \hspace{1cm} closed recurrent classes

Associated to this canonical decomposition is a canonical form of the probability transition matrix, which simply involves ordering the states of the MC according to the canonical decomposition. List all the closed communication classes first, then the transient states.

Then the probability transition matrix will have the form:

\[
P = \begin{pmatrix}
T^c & T \\
0 & 0 \\
T & R & Q
\end{pmatrix}
\]

\[ T^c = \text{recurrent states} \]

Assuming we've grouped the recurrent states together by communication class, \( P \) will have a block diagonal form:

\[
\bar{P} = \begin{pmatrix}
C_1 & C_2 & \cdots & C_n \\
0 & \bar{P}^{(2)} & \cdots & 0 \\
0 & 0 & \ddots & \vdots \\
0 & \cdots & 0 & \bar{P}^{(n)}
\end{pmatrix}
\]

This canonical decomposition will be helpful in presenting the main
results, but it’s not actually always necessary or helpful in practice to conduct a canonical decomposition.

Two key questions concerning long-time properties of reducible Markov chains that require calculations beyond the stationary distribution:

- Starting from a given state, which recurrent communication class does the Markov chain eventually enter, and with what probability?
- If one is associating an additive functional (cost/reward) to each state of the Markov chain, then how much cost/reward is accumulated while the Markov chain is in the transient states?

If these two questions are answered, then one can combine those answers with the stationary distributions associated to each closed communication class in order to answer properties about the long-time probabilities to be in a state (use the stationary distribution conditioned on the recurrent class that is eventually entered) and the long-run averages of additive functionals (law of large numbers for MCs, again conditioned on which recurrent communication class is eventually entered.)

To prepare, we introduce the key random variable $\tau$ which is the first epoch in which a recurrent class is entered.

$$\tau = \min\{n \geq 0: X_n \in T^c\}$$

In everything that follows, we will assume that $X_0 \in T$ else there’s no need for these calculations; just use the stationary distribution associated to the closed communication class of the initial state.

Then given this condition, we have that $\tau \geq 1$ but also, for finite state Markov chains, we are guaranteed by a coin flipping argument that $P(\tau < \infty) = 1$.

**Absorption probability**

To answer the first question, we want to compute:

$$U_{ij} = P(\tau = j | X_0 = i) \text{ for } i \in T \text{ and } j \in T^c$$

This is the probability that the first recurrent state visited, starting from a transient state $i$, is state $j$. A possibly more relevant quantity for application is:

$$\sum_{j \in C_k} U_{ij}, \text{ which is the probability that, starting from transient state } i, \text{ the Markov chain eventually lives in recurrent class } C_k.$$
But it's easier to derive a formula for $U_{ij}$ and this is done using an important concept in Markov chain analysis called **first step analysis**. The idea here is to use the idea of the Chapman-Kolmogorov equation, which we used to derive formulas for finite-horizon calculations, by setting up a recursion based on the state of the Markov chain at an intermediate step. But the twist here is that the end epoch is random. Still, what we'll do is we'll introduce a partition over an intermediate epoch, namely the first epoch, and that will set up a recursion we can solve for the absorption probability.

We will use the law of total probability, with an added side condition:

$$P(A|C) = \sum_k P(A|B_k, C)P(B_k|C) \quad \text{provided that } \{B_k\} \text{ is a partition of } S.$$  

$$U_{ij} = P(X_T = j | X_0 = i)$$

$$= \sum_{k \in S} P(X_T = j, X_1 = k, X_0 = i)P(X_1 = k | X_0 = i)$$

Then it must be the case that $\tau \geq 2$

$$P(X_T = j | X_1 = k) \rightarrow \text{Markov property after summing up all possible finite values of } T$$

(See Resnick Sec. 2.11 for fully rigorous argument of these two steps.)

$$= U_{kj}$$

i) $k \in T$

ii) $k = j \in T^C$
That means that $\tau = 1$ and $X_\tau = j$

$$P(X_{\tau} = j \mid X_i = j, X_{\tau - 1} = i) = P(X_i = j)X_{\tau - 1} = i$$

$$= 1$$

iii) $k \in T^c, k \neq j$

Then $\tau = 1, X_\tau = k \neq j$

$$P(X_{\tau} = j \mid X_i = k, X_{\tau - 1} = i) = P(X_i = j)X_{\tau - 1} = k, X_{\tau} = i$$

$$= 0$$

Putting together these three cases:

$$U_{ij} = \sum_{k \in T} U_{ik} P_{ik} + \sum_{k \in T^c} P_{ik} U_{kj}$$

This yields the recursive formula:

$$U_{ij} = P_{ij} + \sum_{k \in T} P_{ik} U_{kj}$$

for $i \in T, j \in T^c$

Note this recursion formula does not require that one set up the canonical form of the transition matrix.

But if one does set up the canonical form of the transition matrix, then the above recursion formula can be written in matrix form using the canonical blocks.

$$U = R + Q U$$

$$(I - Q) U = R$$
The matrix being inverted can be shown to be guaranteed to be invertible for finite state Markov chains by the Perron-Frobenius theorem.

Expectations of Additive Functionals of Markov Chains while Transient

An additive functional of a Markov chain can be thought of as accumulated cost/reward, which simply sums up a function of the Markov chain at each epoch.

$$\sum_{n} f(X_n)$$

Because we know how to evaluate such sums using the LLN MC inside recurrent classes, let's focus on how to evaluate this additive functional over the epochs while the Markov chain remains transient.

$$\sum_{n=0}^{T-1} f(X_n), \quad T = \min \{n \geq 0 : U \cap X_n \neq +\}$$

We will develop a formula for the expectation of this additive functional, conditioned on the initial state.

$$w_i = \mathbb{E} \left( \sum_{n=0}^{\tau-1} f(X_n) | X_0 = i \right)$$

for $i \in T$.

One can also find and develop formulas for higher moments (variance) of this additive functional, but this requires considerably more work because then correlations between the Markov chain at different epochs matter.