Continuing the discussion of long-time behavior of countable-state Markov chains from last time, we now turn to transient communication classes.

These also have similar properties to the finite state case:

By definition, there is a positive probability that the Markov chain will never return to a transient state. The number of times a Markov chain returns to a transient state is given by a geometric distribution.

\[ \lim_{n \to \infty} p(X_n = j) = 0 \quad \text{for transient } j \]

Analysis of transient states, i.e., to what recurrent communication class do they eventually become absorbed, and how much cost/reward is accumulated by the Markov chain by moving through transient states until it is absorbed into a recurrent communication class -- these questions are addressed with the same formulas as for the finite state case; one just now in principle has to deal with infinite sums (which by similar arguments to last time are guaranteed to converge).

Classification of Countable-State Markov Chains

Topology of the Markov chain by itself does not determine the classification of communication classes with infinitely many states, but it does help, so let's start with it.
Communication classes are still determined by pure topology.

Any communication class that is not closed $C_2$ must be transient because states with one way links in such a communication class to other communication classes have some finite probability of never returning to the original communication class.

Any closed communication class that has finitely many states must be positive recurrent. This allows us to identify $C_1$ as a positive recurrent communication class.

Communication classes that are closed and have infinitely many states ($C_3$) can be of any of the three types and need analysis to determine which type. And for the purposes of this analysis, we treat the closed communication class as its own Markov chain, meaning that we can restrict the following discussion to the case of irreducible (countable-state) Markov chains.

To do this classification, we will combine some fundamental propositions.

Prop 1: Existence of a stationary distribution for an irreducible Markov chain is equivalent to that Markov chain being positive recurrent.

This is justified by first recalling that we showed that any positive recurrent irreducible Markov chain has a unique stationary distribution.

Conversely, how does having a stationary distribution
Imply that the Markov chain is positive recurrent?

Because we know for null recurrent and transient Markov chains:

\[
P(X_n = j) \xrightarrow{\eta \to \infty} 0
\]

\[
\lim_{\eta \to \infty} \begin{pmatrix} \vec{\pi} & \rho^n \end{pmatrix} = \begin{pmatrix} \vec{\pi} & 0 \end{pmatrix}
\]

If we had a stationary distribution for a null recurrent or transient Markov chain, and initialized the Markov chain with this stationary distribution, we would conclude:

\[
\vec{\pi} = \vec{\pi}
\]

Prop. 2: If an irreducible Markov chain does not have an invariant measure (unnormalized stationary distribution), then it must be transient.

Proof: Our argument for deriving the stationary distribution for irreducible positive recurrent Markov chains also showed that every recurrent (null or positive) Markov chain has an invariant measure.

However, this is not an if and only if proposition. Transient Markov chains might have an invariant measure.

Prop. 3: Decisive test for transience.
In an irreducible Markov chain, choose any reference state \( i_x \in S \).

Let \( Q \) be the matrix that is obtained by deleting the row and column from the probability transition matrix corresponding to this reference state. If the only bounded, nonnegative (column vector) solution \( x \) to the equation \( x = Qx \) is the trivial solution \( x = 0 \) (all entries zero), then the Markov chain is recurrent. Otherwise it is transient.

We'll prove this in a minute. Another potentially useful fact (Karlin and Taylor Sec. 3.4) is that under the same setup and definition of \( Q \), if the equation \( Qx = x \) has any unbounded solution

\[
\left\| x \right\|_\infty = \sup_{j \in S \setminus \{i_x\}} x_j = \infty
\]

Then the Markov chain is recurrent. (But not if and only if).

We'll now derive the decisive test for transience.

Choose the reference state \( i_x \).

\[
\beta_j = P(T_{i_x, j} = \infty) \mathbb{1}_{X_0 = j}
\]

This is the probability that starting from state \( j \), I never visit the reference state.

Recurrence is equivalent to the statement:

\[
\beta_j = 0 \quad \text{for all} \quad j \in S \setminus \{i_x\}
\]
We'll derive an equation for \( \beta = (\beta_j)_{j \in S, j \neq i} \)

\[
\beta = \begin{pmatrix}
\rho \\
\frac{1}{2} \\
\frac{1}{2} \\
\frac{1}{2}
\end{pmatrix}
\]

\( \nu_j = 1 - \beta_j \)

\( \nu_j = P(\mathbf{X}_n = i_x \text{ for some } n > 0 | \mathbf{X}_0 = j) \)

To compute this, we can use similar tricks as for finite state Markov chains by setting up an absorption probability calculation on a modified Markov chain. We will modify the Markov chain by making the reference state absorbing.

The canonical decomposition for the probability transition matrix of this modified Markov chain is:

\[
\hat{\rho} = T \begin{pmatrix}
\hat{i}_x & \\ \\
\hat{0} & 0 & \cdots & 0
\end{pmatrix}
\]

One can set up the absorption probability calculation in exactly the same way as for finite state Markov
chains. Note that there is no need for all the absorption probabilities starting from a given state to add to 1. (Any deficit just corresponds to the probability of never being absorbed.) The derivation of this extension is just obtained by repeating the same arguments as we made for the finite state case.

Recall that the equation for the absorption probabilities can be written:

\[ R + \alpha (\mathbf{1} - \beta) = \bar{\mathbf{1}} - \mathbf{0} \]

Normally \( R, Q \) would be matrices but here they are just column vectors because we only have one absorbing state.

\[ Q \beta = \beta + R + \alpha \mathbf{1} - \mathbf{1} \]

\[ R_j + (Q \mathbf{1})_j = \sum_{i \in S} \sum_{k \in S} \left( p_{ij} \right)_{jk} \leq \sum_{i \in S} \left( Q_{ji} \right) + 1 \]

\[ = \sum_{i \in S} \left( p_{ij} \right) + \sum_{i \in S} \left( Q_{ji} \right) \]

\[ = \sum_{i \in S} \left( p_{ij} \right) = 1 \]
If you read the references, this equation is actually derived from first principles using first-step analysis; here we have just carried over absorption probability formulas.

But which solution to this eigenproblem is the one that gives the desired probabilities?

Lemma: \( \vec{\beta} \) is the maximal solution to

\[
\begin{align*}
0 \leq x_j & \leq 1, & 0 \leq x_j & \leq 1 \\
\text{for all } j & \in S \setminus \{i, 3\}
\end{align*}
\]

This means that for any other solution \( \vec{\gamma} \)

Satisfying these conditions, we must have:

\[ Y_j \leq \beta_j \text{ for all } j \in S \setminus \{i, 3\} \]

Proof of Lemma:

We claim that:

\[
\vec{\beta} = \lim_{n \to \infty} Q^n \vec{1}
\]

This is true because:

\[
\beta_j = \lim_{n \to \infty} P( X_n = i, X_0 = j) \text{ under the unif. frod. MC,}
\]
Suppose that \( \mathbf{y} \) is any other solution to:

\[
Q \mathbf{y} = \mathbf{y}, \quad 0 \leq \mathbf{y} \leq \mathbf{1}
\]

(component-wise)

\[
\mathbf{y} \leq Q \mathbf{y} \leq Q^2 \mathbf{1}
\]

\[
\mathbf{y} \leq Q(\mathbf{Q}\mathbf{1}) = Q^2 \mathbf{1}
\]

Continuing by induction, we have:

\[
\mathbf{y} \leq Q^n \mathbf{1}, \quad \forall n \geq 0
\]

\[
\mathbf{y} \leq \lim_{n \to \infty} Q^n \mathbf{1} = \mathbf{v}
\]

So this gives us one way for calculating the probability that a given state in a irreducible Markov chain ever visits another state. But what does this have to do with the
decisive test for transience?

If we found any nonzero, bounded, nonnegative solution to the equation \( x = Qx \), then by rescaling it, we could find a nonzero solution that satisfied:

\[
\begin{align*}
\mathbf{x} &= Q\mathbf{x}, \quad 0 \leq \mathbf{x} \leq 1 \\
\Rightarrow \quad & \mathbf{\beta} \neq \mathbf{0} \quad \text{since} \quad \mathbf{\beta} \geq \mathbf{x} \times \mathbf{0} \\
\Rightarrow \quad & \text{transient.}
\end{align*}
\]

If on the other hand, the only bounded nonnegative solution to \( x = Qx \) is \( 0 \), then we must have \( \mathbf{\beta} = \mathbf{0} \Rightarrow \text{recurrent} \).

Putting the three propositions together with the topological analysis, we obtain a definitive classification procedure for countable-state Markov chains:

1. Decompose the Markov chain into communication classes according to topological considerations
   a. Any communication class that is not closed must be transient
   b. Any closed communication class that is finite must be positive recurrent.

2. Next consider all communication classes that are closed and have infinitely many states. Consider these closed communication classes as their own irreducible Markov chains to do the following analysis:
   a. Look for an invariant measure
      i. If you find an invariant measure that can be normalized into a stationary distribution, then the class is positive recurrent.
      ii. If you can show that no invariant measure exists, then the class must be transient.
iii. If one can find an invariant measure that can't be normalized, then one only knows the class is either transient or null recurrent, and must be decided by the next part.

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 $P$. Look for solutions to $x = Qx$.

i. If you find a nonzero, nonnegative, bounded solution, then the class must be transient.

ii. If you find an unbounded solution, then the Markov chain must be recurrent.

iii. If you prove that the only solution that is nonnegative and bounded is the trivial 0 solution, then the class is recurrent.

You can do parts a or b in either order, perhaps depending on your intuition about what you expect the result to be.

We will now show how to apply this classification procedure to

General birth-death chains

We will assume that

\[ p_0 > 0 \]

\[ p_i, q_i > 0 \quad i \geq 1 \]
So that the Markov chain is irreducible.

Important variation is where we allow $p_0 = 0$, in which case we have just from topology that the absorbing state 0 is positive recurrent and every other state is transient.

But under the stated conditions we have an irreducible Markov chain with infinitely many states. Depending on the parameter values, it could fall into any of the three categories. We will classify what parameter values correspond to each using the analytical component of classification.

We will first attack the classification by looking for an invariant measure.

$$\pi \cdot p = \pi$$

$$\forall_j \geq 0$$

$$\begin{pmatrix}
0 & 1 & 2 & 3 \\
q_0 & p_0 & p_1 & 0 \\
q_2 & r_2 & p_2 & \ddots \\
q_3 & r_3 & \ddots & \ddots \\
0 & \ddots & \ddots & \ddots
\end{pmatrix}$$

$$\forall_0 r_0 + \forall_1 q_1 = \forall_0$$

$$\forall_0 p_0 + \forall_1 r_1 + \forall_2 q_2 = \forall_1$$

$$\forall_{j-1} p_{j-1} + \forall_j r_j + \forall_{j+1} q_{j+1} = \forall_j$$
One gets an infinite system of coupled linear equations. These could be solved by recursion.

But there's a better way. Remember that one trick for solving for stationary distributions is to just try to see if a **detailed balance** solution exists. It's not guaranteed to work, but it's easy to try, and often one is lucky. This detailed balance solution idea also works for invariant measures for the same reason.

We'll just try to see if there's an invariant measure that satisfies detailed balance:

\[ \forall i, j \in \mathcal{S}, \quad \frac{v_i}{p_{ij}} = \frac{v_j}{p_{ji}} \]

The only nontrivial detailed balance equations are those for which \(|i-j|=1|\).

\[
\begin{align*}
  v_j &= v_j \\ v_{j+1} &= v_j \left( \frac{p_{j+1}}{q_{j+1}} \right) \\ v_j &= v_0 \prod_{i=0}^{j-1} \frac{p_i}{q_i} \quad \text{for } j \geq 1
\end{align*}
\]

We have found a detailed balance solution so it is a legitimate invariant measure. Because the Markov chain is irreducible, it is in fact the unique invariant measure (up to multiplication by a constant).