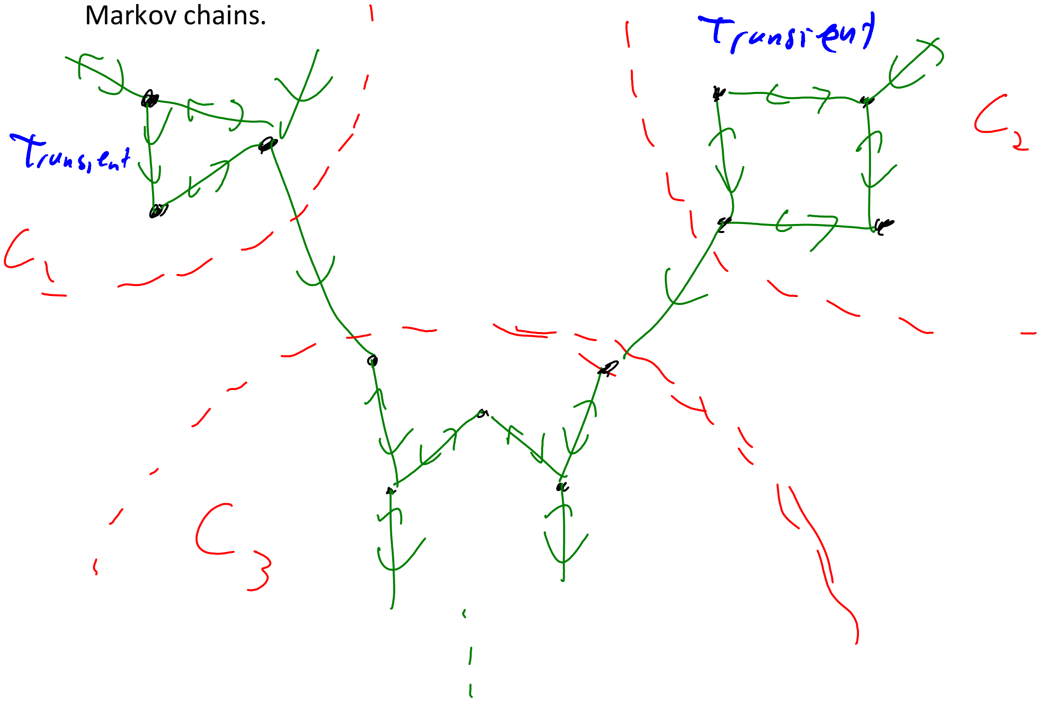
Determining Class Properties in Countable Markov Chains

Tuesday, October 14, 2008
12:03 PM

We will now develop general-purpose techniques for classifying countable-state Markov chains.

The first stage is decomposing the Markov chain into (communication) classes and this proceeds by a topological argument just like finite state Markov chains.



Once you decompose the Markov chain into classes, you can identify some of them as transient directly from the topology, namely those which are not closed.

But the closed communication classes can be transient, null recurrent, or positive recurrent. How do we determine which type it is? We need to do some analysis. For this purpose, we can view the closed communication class as an irreducible Markov chain. So for the following discussion, it suffices to consider irreducible Markov chains because we can extend to reducible Markov chains by the above discussion.

Existence of a stationary distribution for an irreducible Markov chain is equivalent to the Markov chain being positive recurrent.

Stationary distribution: $\vec{\pi}$

$$\pi_j > 0 \qquad j \in S$$

$$\sum_{j \in S} \pi_j = 1$$

$$\vec{\pi} \cdot P = \vec{\pi}$$

Why is this true? We've already discussed that every positive recurrent irreducible Markov chain has a unique stationary distribution. In the other direction, let's prove the contrapositive.

That is, let us show that null recurrent and transient Markov chains do not have a stationary distribution. For such Markov chains, we have

$$\lim_{n \to \infty} (P^n)_{ij} = 0$$

So for any initial probability distribution $\vec{\phi}^{(0)}$

$$\lim_{n\to\infty} \left(\vec{\phi}^{(0)} \cdot P^n\right)_j = \lim_{n\to\infty} \sum_{i \in S} \phi_i^{(0)} (P^n)_{ij} = 0$$

for each $j \in S$

bounded/dominated convergence theorem

In particular then, we can have no solution to

$$\vec{\pi} \cdot P = \vec{\pi} \quad \text{other than } \vec{\pi} = \vec{0}$$

because then $\vec{\pi} \cdot P^n = \vec{\pi} \to \vec{0}$ as $n \to \infty$

This gives a decisive test to determine whether an irreducible Markov chain is positive recurrent -- just see whether it has a stationary distribution!

Every recurrent Markov chain has an invariant measure, so in other words, an irreducible Markov chain with no invariant measure must be transient.

Invariant measure: $\vec{r}$

$$\vec{r} \cdot P = \vec{r}$$

$$r_j > 0 \quad \text{for } j \in S$$

We discussed this previously. (Resnick Ch. 2)

This is not a decisive test for transience, though, because an irreducible Markov chain with an invariant measure can be of any type.

Decisive test for transience (Karlin and Taylor Sec. 3.4)

In an irreducible Markov chain, choose any reference state $i_* \in S$ and let Q be the matrix obtained by deleting the i_*th row and column from the probaiblity tranistion matrix P. If the only bounded, nonnegative solution $\vec{x}$ to $\vec{x} = Q\vec{x}$ is $\vec{x} = 0$,

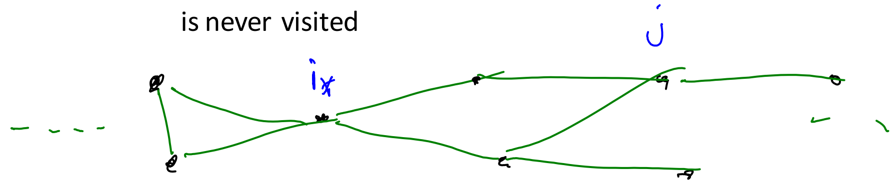
then the Markov chain is recurrent. Otherwise, it is transient.

Why does this work? Consider

$$\beta_j = P(T_{i_*}(1) = \infty \mid X_0 = j)$$

where $T_{i_*}(1) = \min\{n > 0 : X_n = i_*\}$

which is the probability that, starting from state j, reference state i_* is never visited



Recurrence $\Leftrightarrow$ $\beta_j > 0$ for $j \in S$

Transience $\Leftrightarrow$ some $\beta_j > 0$ for $j \in S \setminus \{i_*\}$

(Argue this by first-step analysis).

Let's now derive an equation for $\vec{\beta}$

Observe that $\beta_j = 1 - \tilde{U}_j$ where

$$\tilde{U}_j = P(X_n = i_* \text{ for some } n > 0 \mid X_0 = j)$$

and this remains true if I modify the Markov chain by making state i_* absorbing. This then allows us to invoke the absorption probability formulas because now this modification definitely makes the Markov chain transient.

$\tilde{U}_j$ is the probability to be absorbed at reference state i_* given that the Markov chain starts in state j.

$$P = \begin{array}{c|c} & \begin{array}{cc} i_* & \\ \end{array} \\ i_* & \begin{pmatrix} 1 & 0 \quad 0 \cdots \cdots \quad 0 \\ R & Q \end{pmatrix} \end{array}$$

$$\vec{R} + Q\vec{\tilde{U}} = \vec{\tilde{U}}$$

$$\vec{\tilde{U}} = \{\tilde{U}_j\}_{j \in S \setminus \{i_*\}}$$

$$\vec{R} = \{P_{j i_*}\}_{j \in S \setminus \{i_*\}}$$

$$Q = \{P_{j j'}\}_{\substack{j \neq i_* \\ j' \neq i_*}}$$

$$\vec{\tilde{U}} = \vec{1} - \vec{\beta} \qquad\qquad \vec{1} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$\vec{R} + Q(\vec{1} - \vec{\beta}) = \vec{1} - \vec{\beta}$$

$$(Q\vec{1})_i = \sum_{\substack{j \neq i_* \\ j \in S}} Q_{ij} 1 = \sum_{\substack{j \in S \\ j \neq i_*}} P_{ij}$$

jth component

$$R_i = P_{i i_*}$$

$$R_i + (Q\vec{1})_i = \sum_{\substack{j \in S \\ j \neq i_*}} P_{ij} + P_{i i_*} = \sum_{j \in S} P_{ij} = 1$$

$$1 - (Q\vec{\beta})_i = 1 - \beta_i$$

$$(Q\vec{\beta})_i = \beta_i$$

$$Q\vec{\beta} = \vec{\beta}$$

We derived this expression by using our absorption probability results, but the books derive this from scratch using first-step analysis.

$$\beta_j = \sum_{\substack{k \in S \\ k \neq i_*}} Q_{jk} \, \beta_k$$

k = first state visited at $n > 1$

Actually the equation $Q\vec{\beta} = \vec{\beta}$

does not have a unique solution, so which one do we want?

Lemma: $\vec{\beta}$ is the maximal solution to $Q\vec{x} = \vec{x}$ such that $0 \leq x_j \leq 1$ for all $j \in S \setminus \{i_*\}$.

Proof of Lemma:

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

because
$$\beta_j = \lim_{n \to \infty} (X_n \neq i_* \mid X_0 = j)$$
$$= \lim_{n \to \infty} (Q^n \vec{1})_j = \lim_{n \to \infty} \sum_{k \neq i_*} (Q^n)_{jk}$$
when i_* absorbing.

Let $\vec{x}$ be any solution to

$\vec{x} = Q\vec{x}$ with $0 \leq x_j \leq 1$.

$\vec{x} \leq Q\vec{1}$ (inequality in each component)

nonnegative entries

$\vec{x} = Q^2 \vec{x} \leq Q^2 \vec{1}$

$$\vec{x} = Q^2\vec{x} \le Q^2\vec{1}$$

$$\vdots$$

$$\vec{x} \le Q^n\vec{1} \quad \text{for any} \quad n \ge 0$$

$$\vec{x} \le \lim_{n\to\infty} Q^n\vec{1} = \vec{\beta}$$

Now how can we use this as a test for transience. Consider the equation $\vec{x} = Q\vec{x}$. If the only bounded nonnegative solution is $\vec{x} = \vec{0}$, then $\vec{\beta} = \vec{0}$ so i_* is a recurrent state and since recurrence is a class property, the whole irreducible Markov chain is recurrent.

Conversely, suppose we find some bounded nonnegative solution to $\vec{x} = Q\vec{x}$. By rescaling by a positive number, we can obtain some nontrivial solution such that $0 \le x_j \le 1$ for all $j \in S \setminus \{i_*\}$

$$\Rightarrow \vec{\beta} \ne \vec{0} \Rightarrow \text{transient}$$

A couple of remarks:

1) For transient chains $\sup_{\substack{j \in S \\ j \ne i_*}} \beta_j = 1.$

$\boxed{\text{Resnick Sec. 2.15}}$

2) Another criterion that can sometimes be useful:

With the same setup if the equation $Q\vec{x} = \vec{x}$

has any solution with $\limsup_{j\to\infty} x_j = \infty$

then the Markov chain is recurrent. Karlin and Taylor Section 3.4

Summary of procedure for classifying Markov chains with countably many states

1. Decompose the Markov chain into communication classes based on topological considerations
   - Any communication which is not closed must be transient

2. Consider each closed communication class and analyze it as if it were an irreducible Markov chain with state space identified with that closed communication class
   a. Look for an invariant measure
      i. If you find an invariant measure that can be normalized, then you have a stationary distribution and the class must be positive recurrent.
      ii. If you prove no invariant measure exists, then the class must be transient.
      iii. If your find a non-normalizable invariant measure, then the class is either null recurrent or transient.
   b. Now choose any convenient reference state i_* and let Q be the matrix obtained by deleting the i_* th row and column from the probability transition matrix. Look for solutions to $\vec{x} = Q\vec{x}$.
      i. If you find a nonzero, nonnegative, bounded solution, then the class must be transient.
      ii. If you prove that the only nonnegative bounded solution is the trivial solution $\vec{x} = \vec{0}$ then the class must be recurrent (null or positive).
      iii. If you find an unbounded solution then the Markov chain is recurrent.

Assuming you can do the computations, this procedure will decisively determine the class properties of the Markov chain. Actually you can do parts a and b in either order.