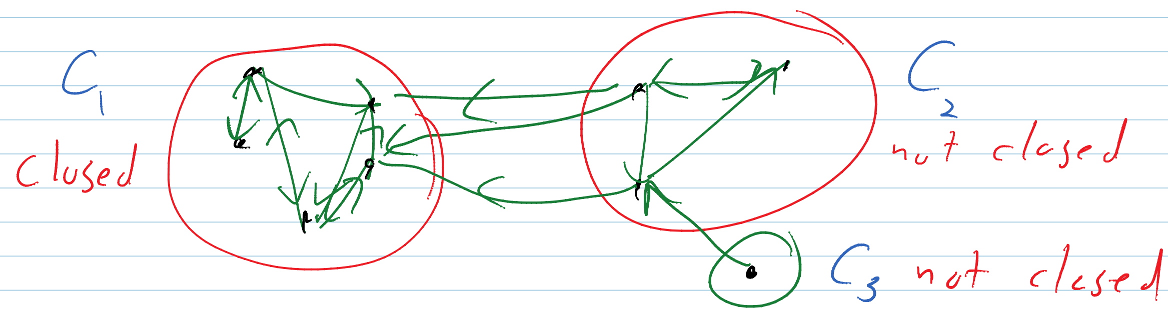
Homework 2 due Monday, October 19 at 2 PM.

The proofs of the assertions about uniqueness of stationary distribution for irreducible Markov chains, and the fact that stationary distribution is a limit distribution for irreducible, aperiodic Markov chains can be found in Resnick Sec. 2.12-2.15.

- The most interesting aspect is the coupling argument for showing the stationary distribution is the limit distribution for FSDT Markov chains which are irreducible and aperiodic. This technique is very useful also in more complex settings.

The stationary distribution, when it serves as a limit distribution for a Markov chain, is adequate to describe most long-time properties of interest. But this requires at least irreducibility for the LLN for MC to apply.

What are the long-term properties of a FSDT MC when it is reducible? If the MC only has one closed communication class, then one can show that it's still true that the MC has only a unique stationary distribution and the LLN of MC applies, and it's a limit distribution if the MC is aperiodic.



A principle that we will often use: If we have a sequence of events $\{A_n\}_{n=0}^{\infty}$ which are independent, and have the property that $\Pr(A_n) \geq \delta > 0$ for all n, then $\Pr(\cap_{n=0}^{\infty} A_n^c) = 0$. In other words, the probability is zero that none of the events will actually occur. This is a special case of something called the second Borel-Cantelli lemma, and can be referred to as the "Infinite Monkey Theorem."

So think of the events A_n being something like: "At epoch $100n$, the Markov chain will enter C_1 at some epoch from epoch $100n$ to epoch $100(n+1)-1$." OK, those events aren't actually independent, but we can still apply a variation of the Infinite Monkey Theorem by using the multiplication principle of conditional probability:

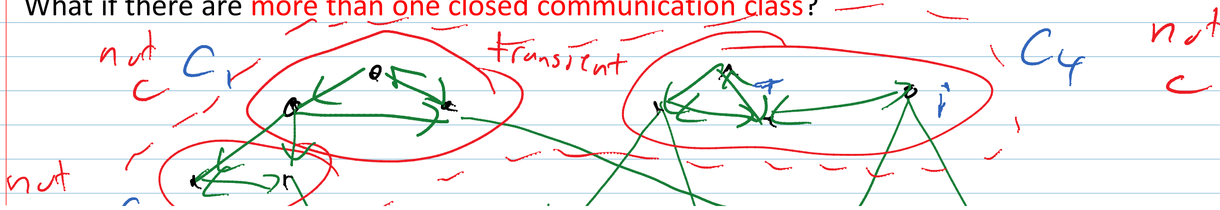
$\Pr(\cap_{n'=0}^{n} A_n^c) = \Pr(A_n^c \mid \cap_{n'=0}^{n-1} A_{n'}^c)\,\Pr(\cap_{n'=0}^{n-1} A_{n'}^c) \leq (1-\delta)\Pr(\cap_{n'=0}^{n-1} A_{n'}^c)$

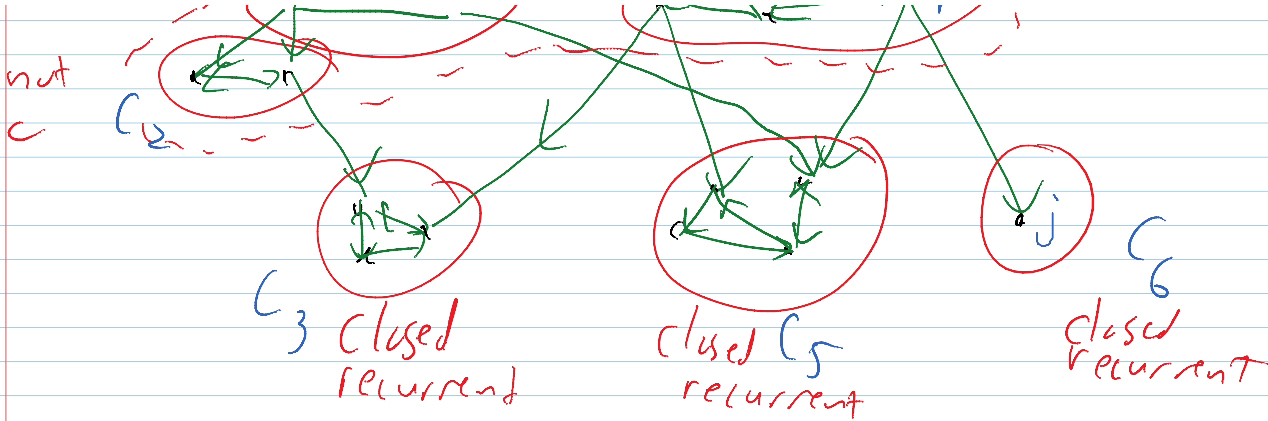
where we are now using: $\Pr(A_n^c \mid \cap_{n'=0}^{n-1} A_{n'}^c) \geq \delta > 0$ for all $n \geq 0$. Therefore

$\Pr(\cap_{n'=0}^{n} A_n^c) \leq (1-\delta)^{n+1}$ by induction, so

$\Pr(\cap_{n=0}^{\infty} A_n^c) = \lim_{n\to\infty} \Pr(\cap_{n'=0}^{n} A_n^c) = 0$.

What if there are more than one closed communication class?





Here we can have nontrivial relationships between the closed and non-closed communication classes, i.e., here.:

$$C_1 \to C_2,\ C_1 \to C_5 \quad C_2 \to C_3,\ C_4 \to C_3,\ C_4 \to C_5,\ C_4 \to C_6$$

The analysis of long-term properties of Markov chain are complicated by trying to determine which closed communication class the Markov chain will eventually reach. Once this is settled, the long-term properties can be deduced by referring to the stationary distribution associated to the Markov chain restricted to a given closed communication class.

This kind of picture, where a Markov chain can be separated into different closed communication classes, or in other words, have long-term fate that is determined by what happens over a finite number of epochs:

- Chemical or biochemical reaction process involving multiple steps
  - Closed communication classes could correspond to the reaction ending in a certain product (or falling back into the reactants)
- Spread of disease in a naïve population
  - Closed communication classes could correspond to a contained spread which goes extinct, or an explosive outbreak that becomes endemic/severe
- Ecosystem and environmental modeling
  - Closed communication classes could correspond to healthy/sustainable state vs. extinction/terrible state.
- Economic models for emerging industries/companies
  - Closed communication classes could correspond to market leader, niche player, or bust.

We will need to introduce some further terminology to classify states and classes of a Markov chain to organize our mathematical formulas and results to follow.

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 again to state j:

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

Recall $T_j(1) = \min_{n>0}(n: X_n = j)$ is the first passage/return time.

A state which is not recurrent is said to be transient.

A few general statements about transience/recurrence:

A recurrent state in fact has probability 1 to return to itself infinitely often.

Recurrence/transience is a class property (all states in a communication class are all transient or all recurrent). Easiest to show that if one state in a communication class is recurrent, then they all are by contradiction. (Infinite monkeys)

Any closed communication class with a finite number of states must be recurrent. (Infinite monkeys)

These statements show that determining transience or recurrence for a finite state Markov chain is a purely topological exercise; we will have to work harder for infinite-state Markov chains to determine transience/recurrence.

- For finite state Markov chains, closed classes are recurrent and non-closed classes are transient.

It can be sometimes useful to decompose a reducible Markov chain using the canonical decomposition:

$S = T \cup_k C_k$ where T are the transient states, and $\{C_k\}$ are the closed communication classes.

In the above example, $S = T \cup C_3 \cup C_5 \cup C_6$

One can write the probability transition matrix in a corresponding canonical form:

$$P = \begin{array}{c} \\ T^c \\ T \end{array} \begin{array}{c} \begin{array}{cc} T^c & T \end{array} \\ \begin{pmatrix} \bar{P} & 0 \\ R & Q \end{pmatrix} \end{array}$$

block matrices

The most useful aspect of this decomposition is the reference to the submatrices P, Q, R which describe the dynamics in/between the recurrent and transient states.

Note that the matrix P is in block diagonal form:

$$\bar{P} = \begin{array}{c} \\ C_3 \\ C_5 \\ C_6 \end{array} \begin{array}{c} \begin{array}{ccc} C_3 & C_5 & C_6 \end{array} \\ \begin{pmatrix} \bar{P}^{(3)} & 0 & 0 \\ 0 & \bar{P}^{(5)} & 0 \\ 0 & 0 & \bar{P}^{(6)} \end{pmatrix} \end{array}$$

block matrixes

The canonical decomposition can be helpful in organizing the mathematical results, but it's not necessarily desirable to actually use the canonical decomposition in practice.

Now we can formulate two key questions to which we'd like computable formulas:

- Starting from a given state, what is the probability to eventually enter a given closed communication class?
- Suppose we associate an additive functional (accumulated cost/reward) to each state of the Markov chain. What is the value of this additive functional at the epoch at which one leaves the transient states and enters into a closed communication class?

Once we answer both of those questions, we can compute long-run properties of the reducible Markov chain by essentially using the stationary distribution which is a weighted average of the stationary distributions associated to each closed communication class, with weight given by the probability to enter that closed communication class. (Law of total probability/expectation combined with LLN for MC and/or stationary distribution as limit distribution)

To prepare to answer these questions, we introduce a random variable
$\tau = \min_{n \geq 0}\{n: X_n \in T^c\}$

which is the first epoch at which the Markov chain enters a recurrent class.

If $X_0 \in T^c$, then everything boils down to the dynamics within the closed communication class to which X_0 belongs, and there is nothing more to compute. So we will assume in what follows that $X_0 \in T$.

Therefore we will have that $\tau \geq 1$ under this assumption, and by infinite monkey theorem, $P(\tau < \infty) = 1$ in a finite-state Markov chain.

Here are the key formulas which we'll derive next time:

1) Define $U_{ij} = P(X_\tau = j | X_0 = i)$ for $i \in T, j \in T^c$

It is the solution of the following recursive equation:

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

2) If we define a deterministic cost/reward function $f(i)$ for $i \in S$, and

Answer is the solution to the recursive equation:

$$w_i = f(i) + \sum_{k \in T} P_{ik} w_k$$