Derivation of limit theorems for irreducible and reducible Markov chains

Monday, September 29, 2008
12:01 PM

Homework 2 will be posted soon (tonight?), due Tuesday, October 14.

Need to show:

$$\text{✱} \quad \sum_{k \in S} r_k P_{k\ell} = r_\ell \qquad \text{for } \ell \in S$$

$$r_k = \sum_{n=0}^{\infty} P\{X_n = k, \tau_j(1) > n \mid X_0 = j\}$$

LHS of ✱:

$$\sum_{k \in S} \sum_{n=0}^{\infty} P\{(\underbrace{X_n = k, \tau_j(1) > n}_{B}) \mid \underbrace{X_0 = j}_{C}\} P(\underbrace{X_{n+1} = \ell}_{A} \mid X_n = k)$$

Suggests: $P(A \text{ and } B \mid C) = P(A \mid B \text{ and } C) P(B \mid C)$ (law cond. prob.)

$A = \{X_{n+1} = \ell\}$

$B = \{X_n = k, \tau_j(1) > n\}$

$C = \{X_0 = j\}$

We will show:

$$P(X_n = k, \tau_j(1) > n \mid X_0 = j) P(X_{n+1} = \ell \mid X_n = k)$$

$$= P(A \mid B \text{ and } C) P(B \mid C) = P(A \text{ and } B \mid C)$$

by showing:

i) $k \neq j \Rightarrow P(X_{n+1} = \ell \mid X_n = k) = P(A \mid B \text{ and } C)$

ii) $k = j \Rightarrow P(B \mid C) = 0$ (so $P(A \text{ and } B \mid C) = 0$ too)

(provided we understand $P(A_1 \mid A_2) \equiv 0$ if $P(A_2) = 0$)

Proof of i): $P(A \mid B \text{ and } C)$

$$= P(X_{n+1} = \ell \mid X_n = k, \underbrace{\tau_j(1) > n}_{\{X_1 \neq j, X_2 \neq j, \ldots, X_n \neq j\}}, X_0 = j)$$

redundant w/ $X_n = k \neq j$

$$= P(X_{n+1} = \ell \mid X_n = k) \checkmark$$

(Markov property)

ii) $P(B \mid C) = P(X_n = k, \tau_j(1) > n \mid X_0 = j)$

$= P(X_n = j,\ T_j(1) > n \mid X_0 = j)$

contradictory/mutually exclusive

$= P(\emptyset \mid X_0 = j)$

$= 0$

LHS of ※

$$= \sum_{n=0}^{\infty} \sum_{k \in S} P(X_{n+1} = \ell,\ X_n = k,\ T_j(1) > n \mid X_0 = j)$$

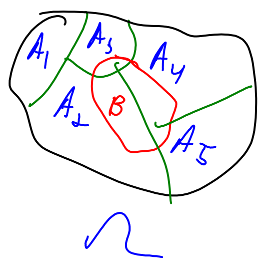
$$= \sum_{n=0}^{\infty} P\Big(X_{n+1} = \ell,\ \underbrace{\coprod_{k \in S} \{X_n = k\}}_{\{X_n \in S\} = \Omega},\ T_j(1) > n \mid X_0 = j\Big)$$

$$= \sum_{n=0}^{\infty} P(X_{n+1} = \ell,\ T_j(1) > n \mid X_0 = j)$$

Partition: $\{A_j\}_{j \in I}$ partition Ω provided:

$A_j \cap A_{j'} = \emptyset$ for $j \neq j' \in I$

$\bigcup_{j \in I} A_j = \Omega$



$\sum_{j \in I} P(A_j) = 1$

$\sum_{j \in I} P(A_j \mid C) = 1$

$\sum_{j \in I} P(A_j \cap B) = P(B)$

$\sum_{j \in I} P(A_j \cap B \mid C) = P(B \mid C)$

Law of Total Probability $\sum_{j \in I} P(B \mid A_j)\, P(A_j) = P(B)$

$\sum_{j \in I} P(B \mid A_j \cap C)\, P(A_j \mid C) = P(B \mid C)$

Example: $A_k = \{X_n = k\}$

$\{A_k\}_{k \in S}$ is partition of Ω

To finish the proof, then we need to show that

$$\text{LHS} = \sum_{n=0}^{\infty} P(X_{n+1} = \ell,\ T_j(1) > n \mid X_0 = j)$$

$$\stackrel{?}{=} \sum_{n=0}^{\infty} P(X_n = \ell,\ T_j(1) > n \mid X_0 = j) = \text{RHS}$$

Show this:

1) For $\ell \neq j$:

$n = 0$ term on RHS $= 0$

$$RHS = \sum_{n=1}^{\infty} P(X_n = \ell,\ \tau_j(1) > n \mid X_0 = j)$$

$n = n'+1$

$$= \sum_{n'=0}^{\infty} P(X_{n'+1} = \ell,\ \tau_j(1) > n'+1 \mid X_0 = j)$$

$\{X_{n'+1} = \ell,\ \tau_j(1) > n'\}$

$= \{X_{n'+1} = \ell,\ \tau_j(1) > n'+1\}$

because $\{X_{n'+1} = \ell,\ \tau_j(1) = n'+1\}$

$= \emptyset$ as $j \neq \ell$

$$= \sum_{n'=0}^{\infty} P(X_{n'+1} = \ell,\ \tau_j(1) > n' \mid X_0 = j)$$

$= LHS$ after $n' \to n$

ii) For $j = \ell$:

$$LHS = \sum_{n=0}^{\infty} P(X_{n+1} = j,\ \tau_j(1) > n \mid X_0 = j)$$

$\quad\quad \shortparallel$

$\{\tau_j(1) = n+1\}$

$$= \sum_{n=0}^{\infty} P(\tau_j(1) = n+1 \mid X_0 = j)$$

(partition, provided

$P(\tau_j(1) < \infty) = 1$)

This condition is guaranteed for irreducible finite-state Markov chains by a technical argument (see texts).

$= 1$

$$RHS = \sum_{n=0}^{\infty} P(X_n = j,\ \tau_j(1) > n \mid X_0 = j)$$

$$= P(X_0 = j,\ \tau_j(1) > 0 \mid X_0 = j) + \sum_{n=1}^{\infty} P(\underbrace{X_n = j,\ \tau_j(1) > n}_{\text{contradict}} \mid X_0 = j)$$

$$= P(\Omega) + \sum_{n=1}^{\infty} 0$$

$= 1 + 0 = 1$ ✓

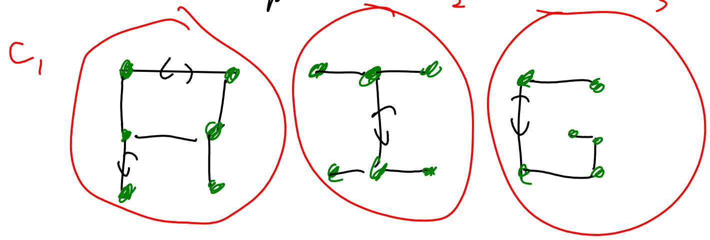
This proves the existence of the stationary distribution as constructed in terms of ratios of numbers of visits to a state and the expected return times.

Proofs of uniqueness and limit theorems concerning stationary distribution: see the references. One noteworthy aspect to these proofs is the coupling argument in Resnick Sec. 2.13 in the proof that the probability distribution for the state of an irreducible, aperiodic Markov chain converges to the limit distribution.

Reducible Markov chains (Lawler Ch. 1)

How do we compute the long-time properties of a Markov chain if it is reducible (not irreducible)?

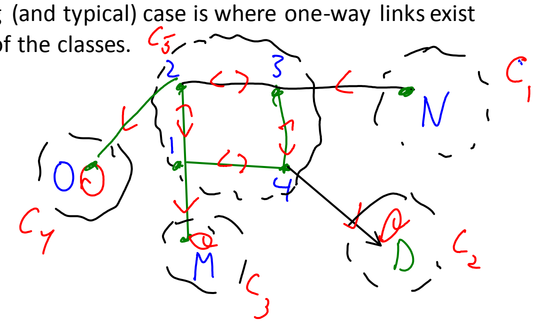
Trivial example:



No communication class is accessible from any other communication class. In this case, the Markov chain can be considered essentially as separate Markov chains corresponding to each class, compute the long-term properties separately within each class as for a irreducible Markov chain. Note that the long-term properties will depend on how much of the initial probability distribution belongs to each class.

More interesting (and typical) case is where one-way links exist between some of the classes.

$T = C_1 \cup C_5$



We see some one-way links between classes.

$C_1 \to C_5$ $\qquad$ $C_5 \to C_3$

$C_5 \to C_4$ $\qquad$ $C_5 \to C_2$

This is a cartoon for more sophisticated applications where this sort of reducibility emerges:

- financial evolution of a company
- (bio)chemical reactant
- ecosystems: organisms moving and getting eaten

  We will also see in the next lecture that consideration of reducible Markov chains also gives us tools for addressing some other questions involving irreducible Markov chains.

To begin, we want to characterize one distinction between classes

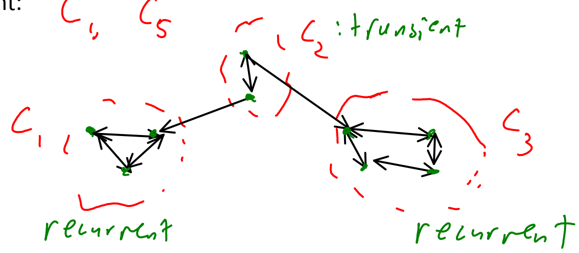
A state j of a Markov chain is said to be recurrent provided that if the Markov chain starts in state j, it will return to state j with probability 1.

$$P(\tau_j(1) < \infty \mid X_0 = j) = 1$$

A state of a Markov chain which is not recurrent is said to be transient.

Facts:
- Recurrence/transience is a class property.
- Recurrent states are actually visited infinitely often (with probability one) if the Markov chain starts in the class containing that state.

In above example:
Recurrent: C_2, C_3, C_4
Transient: C_1, C_5



How do we know the classes we've labelled recurrent are actually recurrent, meaning that I am guaranteed to visit a state I started in?

A closed class is a class in which no other states outside of the class are attainable from within the class. (No way to leave the class).

Fact: A closed class with a finite number of states must be a recurrent class.

Canonical decomposition of a Markov chain state space:

$$S = T \cup C_k$$

transient states

list only recurrent classes

In this decomposition, we don't care about whether the transient states organize themselves into one or more classes.

Now let's consider the long-time properties of a reducible Markov chain under this decomposition. If the Markov chain has initial state in some recurrent class C_k then the Markov chain will always remain in that recurrent class, and one can compute long-time properties by just considering the Markov chain defined with state space equal to the recurrent class C_k. Since this recurrent class is by definition irreducible, one can apply the limit theorems involving stationary distribution already developed (checking aperiodicity if necessary).

But for the cases in which the initial state is transient, new techniques are needed.

Key question: Absorption probability

The Markov chain is guaranteed to eventually visit (and therefore remain) in one of the closed recurrent classes. But what is the probability, starting from a given transient state, to end up in each of the possible recurrent classes?

Secondary question: Accumulated cost/reward during transient phase

In addition to the long-term properties associated with the closed recurrent class in which the state eventually becomes trapped, perhaps quantities like the number of epochs until visiting the recurrent class and associated costs/rewards are desirable to compute.