Absorption Probability Calculations

Thursday, October 15, 2015

2:00 PM

Homework 2 due Monday, October 19 at 2 PM.

Let's give a mathematical derivation of the recursive formulas for absorption probabilities and expectations of additive functionals for reducible FSDT Markov chains starting from a transient state.

We will only provide the formal mathematical steps; there are some technicalities surrounding the fact that the quantities of interest refer to a random epoch τ. To handle these technicalities, see Resnick Sec. 2.11.

The key concept that we want to illustrate in these derivations is first-step analysis. This shows how to use the Markov property to derive recursive formulas for statistics of Markov chains, and will be useful in more complicated situations later.

Let's first turn to the absorption probability:

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

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

The basic idea of first-step analysis is to use the law of total probability, partitioning on what happens at the first epoch (first step). The argument is somewhat similar to the derivation of the Chapman-Kolmogorov equation, but the difference is that the end epoch is random (τ).

Probabilistic formula to apply: $P(A|C) = \sum_k P(A|B_k, C)P(B_k|C)$

where $\{B_k\}$ is a partition of sample space.

We will take $B_k = \{X_1 = k\}$

$$U_{ij} = P(X_\tau = j \mid X_0 = i)$$

$$= \sum_{k\in S} P(X_\tau = j \mid X_1 = k, X_0 = i)\, \underbrace{P(X_1 = k \mid X_0 = i)}_{P_{ik}}$$

Three possible cases for the first factor:

1) If $k = j$

$$P(X_T = j \mid X_1 = k, X_0 = i) = 1$$

2) If $k \in T^c$, $k \neq j$

$$P(X_T = j \mid X_1 = k, X_0 = i) = 0$$

3) If $k \in T$

$$P(X_T = j \mid X_1 = k, X_0 = i)$$

Know $T \geq 2$

$$= P(X_T = j \mid X_1 = k)$$

by Markov property

strong

$$= U_{kj}$$

(by time shift)

Put everything together:

$$U_{ij} = 1\, P_{ij} + \sum_{\substack{k \in T^c \\ k \neq j}} 0 \cdot P_{ik} + \sum_{k \in T} U_{kj} P_{ik}$$

$$U_{ij} = \underbrace{P_{ij}}_{R_{ij}} + \sum_{k \in T} U_{kj} \underbrace{P_{ik}}_{Q_{ik}}$$

Markov property is what makes this equation a closed-loop recursion.

Apply the same idea for expectations of additive functionals.

$w_i = \mathbb{E}(\overset{Y}{\sum_{n=0}^{\tau-1} f(X_n)} | \overset{C}{X_0 = i})$ for $i \in T$

Let's take f as deterministic for now.

Instead of law of total probability, we will use law of total expectation:

$\mathbb{E}(Y) = \sum_k \mathbb{E}(Y|B_k)P(B_k)$ where $\{B_k\}$ is a partition of sample space.

By applying a common condition C, we can modify this to:

$$\mathbb{E}(Y|C) = \sum_k \mathbb{E}(Y|B_k, C)P(B_k|C)$$

We will apply this to w_i, with $B_k = \{X_1 = k\}$. First step analysis.

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

$$= \sum_{k \in S} \mathbb{E}\left(\sum_{n=0}^{\tau-1} f(X_n) \,\Big|\, X_1 = k, X_0 = i\right) \underbrace{P(X_1 = k | X_0 = i)}_{P_{ik}}$$

Two cases for the first factor:

1) $k \in T^c$ Then $\tau = 1$.

$$\mathbb{E}\left(\sum_{n=0}^{\tau-1} f(X_n) \,\Big|\, X_1 = k, X_0 = i\right)$$

$$= \mathbb{E}\left(f(X_0) \,|\, X_1 = k, X_0 = i\right)$$

$$= \mathbb{E}\left(f(i) \,|\, X_1 = k, X_0 = i\right)$$

$= f(i)$ since f deterministic

2) $k \in T \implies \tau \geq 2$

$$\mathbb{E}\left(\sum_{n=0}^{\tau-1} f(X_n) \,\middle|\, X_1 = k, X_0 = i \right)$$

$$\mathbb{E}\left(f(i) + f(k) + \sum_{n=2}^{\tau-1} f(X_n) \,\middle|\, X_1 = k, X_0 = i \right)$$

$$= \underbrace{f(i) + f(k)}_{\text{deterministic}} + \mathbb{E}\left(\underbrace{\sum_{n=2}^{\tau-1} f(X_n)}_{\text{future}} \,\middle|\, \underbrace{X_1 = k}_{\text{present}}, \underbrace{X_0 = i}_{\text{past}} \right)$$

$$= f(i) + f(k) + \mathbb{E}\left(\sum_{n=2}^{\tau-1} f(X_n) \,\middle|\, X_1 = k \right)$$

Markov property

$$= f(i) + \mathbb{E}\left(f(k) + \sum_{n=2}^{\tau-1} f(X_n) \,\middle|\, X_1 = k \right)$$

$$= f(i) + \mathbb{E}\left(f(X_1) + \sum_{n=2}^{\tau-1} f(X_n) \,\middle|\, X_1 = k \right)$$

$$= f(i) + \mathbb{E}\left(\sum_{n=1}^{\tau-1} f(X_n) \,\middle|\, X_1 = k \right)$$

$$= f(i) + w_k \quad \text{(time shift)}$$

Combining results from two cases:

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

$$w_i = \sum_{k \in T^c} f(i)\, P_{ik} + \sum_{k \in T} \left(f(i) + w_k\right) P_{ik}$$

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

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

$$= f(i) \cdot 1 + \sum_{k \in T} w_k P_{ik}$$

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

If the cost/reward function f was random, but random in a way that was conditionally independent of the Markov chain, i.e., $f = f(X_n, \xi_n)$ where ξ_n are iid random variables which are independent of X_n, then the formula for the expectation of the additive functional would simply change as follows:

$w_i = \mathbb{E} f(i, \xi) + \sum_{k \in T} w_k P_{ik}$ by repeating the same derivation. This is analogous to how we extended the LLN for MC's to compute the long-run quality of the inspection protocol.

We can rewrite the linear recursion equations for the absorption probabilities and additive functionals as a matrix equation if we refer to the canonical decomposition.

Absorption probability recursion formula can be written in matrix form:

$U = R + QU$ where $U = \{U_{ij}\}_{i \in T, j \in T^c}$

Solve for U:

$U = (I - Q)^{-1} R$

One can show that $I - Q$ must be invertible by showing that Q has all eigenvalues less than 1 in magnitude (by Perron-Frobenius argument). (Q is strictly substochastic matrix). I is the identity matrix.

The additive functional formula, in matrix form, is:

$$w = f^{(T)} + Qw$$

where $w = \{w_i\}_{i \in T}$ (column vector)

$f^{(T)} = \{f(i)\}_{i \in T}$ (column vector)

Solve for w:

$$w = (I - Q)^{-1} f^{(T)}$$

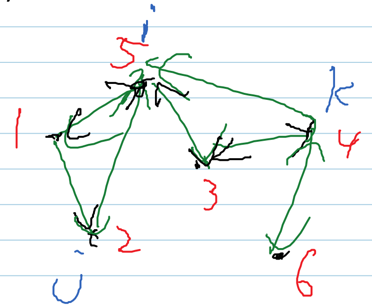
The recursion formulas are typically better to use if one has hope for an analytic solution (as we shall see in an example). If not analytic solution, then the matrix formulations probably more efficient.

The formulas we have derived for absorption probability and expected accumulated cost/reward can be usefully applied for some general questions concerning Markov chain dynamics:

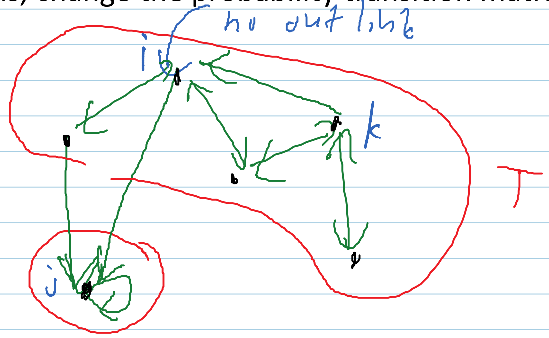
1) Expected number of epochs, starting from a transient state i, until a recurrent state is reached: Use the additive functional formula with $f \equiv 1$. Then $w_i = \mathbb{E}(\tau | X_0 = i)$
2) Expected number of visits to a transient state j, starting from a transient state i: Use additive functional formula with $f(k) \equiv \delta_{jk}$ (Kronecker delta), meaning: $f(j) = 1, f(k) = 0$ if $k \neq j$

   Then $w_i = \mathbb{E}(N_j | X_0 = i)$ where N_j is total number of visits to state j.

Now some further questions we can answer regarding the relationship of a Markov chain to three states i, j, k which are in the same closed (recurrent) communication class.



A) Starting from state i, what is the expected number of epochs until state j is reached? (Mean first passage time calculation). To answer this, modify the Markov chain by making state j absorbing by cutting all of its outlinks. In other words, change the probability transition matrix so that $P_{jm} = \delta_{jm}$.



In this modified Markov chain, j is recurrent, but all other states are transient. Compute the expected time until hit a recurrent state in this modified Markov chain.

B) Starting from state i, what is the expected number of epochs spent in state k, before state j is visited?

Make same modification, but now use cost/reward function which is δ_{km}

C) Starting from state i, what is the probability that state j is visited before state k?

Make both states j and k absorbing, and then do an absorption probability calculation. The probability to be absorbed in state j, starting from state i, in the modified MC will be the same as the probability that the original MC hits state j before state k, starting from state i.

The reason these manipulations are OK is that they only alter the dynamics of the Markov chain, once the uncertainty in the result is settled. The same tricks can also be tried for reducible Markov chains with one closed communication class, so long as the fake absorbing states are the only recurrent states in the modified Markov chain.