Continuous time Markov Chains

Thomas Mountford

EPFL

May 10, 2021

Recall

In this course we will restrict attention to the minimal Markov chains. These will be what we consider to be continuous time Markov chains. The following gives two equivalent ways of looking a these Markov chains.

Theorem

Given a general Q-matrix Q and the associated minimal semigroup $P(t)_{t\geq 0}$ the following are equivalent for a process $(X_t)_{t\geq 0}$ that is cadlag on I

a) Conditional upon $X_0 = i$, the jump chain of X, $(Y_n)_{n\geq 0}$ is a discrete time (δ_i, Π) Markov chain (where Π is derived from Q in the usual way) and conditional upon $(Y_n)_{n\geq}$, the holding times $(S_n)_{n\geq 1}$ are independent $\mathcal{E}xp(q_{Y_{n-1}})$ random variables.

b) Given integer $n \geq 0$ and times $0 \leq t_0 < t_1 < \cdots < t_{n+1}$ and $i_0, i_1 \cdots i_{n+1} \in I$,

$$P(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, X_{t_1} = i_1 \cdots X_{t_n} = i_n) = P_{i_n i_{n+1}}(t_{n+1} - t_n)$$

In this part we use a) rather than b) to analyze the chains

Class Structure

Definition

Given Q-matrix Q, we say i leads to j if

$$\mathbb{P}_i(\exists t \geq 0 : X_t = j) > 0$$

We write $i \to j$ if this is the case.
Definition We say i and j communicate if $i \to j$ and $j \to i$

Theorem

The following are equivalent to $i \to j$ for i and j distinct

(i) $i \to j$ *for the jump chain* $(Y_n)_{n \geq 0}$

(ii) $\exists n > 0$ *and* $i = i_0, i_1 \cdots i_n = j$ *so that* $\forall 1 \leq j \leq n, q_{i_{j-1} i_j} > 0$

(iii) $\forall t > 0, P_{ij}(t) > 0$

(iv) *for some* $t > 0, P_{ij}(t) > 0$

Proof

(i) is obviously equivalent to $i \to j$ given the jump chain representation. Furthermore (i) is equivalent to the existence of integer $n > 0$ and distinct $i = i_0, i_1 \cdots i_n = j$ so that $\forall 1 \leq j \leq n, \pi_{i_{j-1} i_j} > 0$. But this is equivalent to the exisitence of n and distinct $i = i_0, i_1 \cdots i_n = j$ so that $\forall 1 \leq j \leq n, q_{i_{j-1} i-j} > 0$, that is (ii). If (ii) holds with n and $i = i_0, i_1 \cdots i_n = j$, then for any $t > 0$

$$\mathbb{P}_i(X_t = j) \geq \mathbb{P}_i(Y_j = i_j \forall j \leq n, S_j < t/n \forall j < n, S_n > t) > 0.$$

It is immediate that (iii) implies (iv). But, again from the jump chain representation, it is clear that (iv) implies (i)

Given our definition for i and j communicating, we can easily see that this relation partitions I into communicating classes, as in Chapter 1. We say the chain is irreducible if there is a single communicating class (and it is easily seen that this is the same as the jump chain being irreducible). We similarly speak of *closed* classes and *absorbing* sites (i is absorbing if and only if $q_i = 0$).

Hitting probabilities

Some questions for continuous time Markov chains can be rewritten as questions fore the discrete time jump chain. An important example is calculating $\mathbb{P}_i(D_A < \infty)$ for $i \in I$, $A \subset I$ and $D_A = \inf\{t \geq 0 : X_t \in A\}$. The event $D_A < \infty$ is exactly the event that $\exists n \geq 0 : Y_n \in A$ for Y the discrete time jump chain for X.

So if we write h_i^A for $\mathbb{P}_i(D_A < \infty)$ which equals $\mathbb{P}_i^Y(T_A < \infty)$, we have

Theorem

The function h^A is the smallest positive function satisfying

- $h_i^A = 1$ *for* $i \in A$;
- $\sum_j q_{ij} h_j^A = 0$ *for* $i \notin A$;

Proof: We know that h_i^A is the smallest positive function satisfying $h_i^A = 1$ for $i \in A$ and $\sum_j \pi_{ij} h_j^A - h_i^A = 0$ for $i \notin A$. We note that for i not in A if it is absorbing then $h_i^A = 0$ and $\sum_j q_{ij} h_j^A = 0$; if it is not then $q_i(\sum_j \pi_{ij} h_j^A - h_i^A) = 0$ which is precisely $\sum_j q_{ij} h_j^A = 0$. The reverse holds in the same way.

Just as in chapter one, we have

Corollary

For A and B disjoint subsets of I. Let $h_i^{A,B} = \mathbb{P}_i(D_A < D_B)$ The function $h^{A,B}$ is the smallest positive function satisfying

- $h_i^{A,B} = 1$ *for* $i \in A$;
- $h_i^{A,B} = 0$ *for* $i \in B$;
- $\sum_j q_{ij} h_j^{A,B} = 0$ *for* $i \notin A \cup B$;

Again, as in Chapter 1 we can consider function

$$k_i^A = \mathcal{E}_i(D_A)$$

Theorem

Suppose that $q_i > 0$ for each $i \in I$ The function k^A is the smallest positive function satisfying

- $k_i^A = 0$ *for* $i \in A$;
- $\sum_j q_{ij} k_j^A = -1$ *for* $i \notin A$;

Remark: We need the hypothesis that $q_i > 0$ as otherwise if $i \notin A$, the second equation cannot be true.

Proof

Unlike the previous result which was just a reqriting of the corresponding result for the jump chain, this statement requires a reworking of the proof of the analagous result for discrte time processes. Again we use the jump chain. We first show k^A satisfies the claimed equations. Obviously $k_i^A = 0$ on A so we suppose that $i \notin A$. In this case we condition on the first jump time S_1 for the chain starting at i:

$$k_i^A = \mathbb{E}_i(E(D_A|S_1)) = \mathbb{E}_i\left(S_1 + \sum_j \pi_{ij} k_j^A\right) =$$

$$\frac{1}{q_i} + \sum_{j \neq i} \pi_{ij} k_i^A = \frac{1}{q_i}(1 + \sum_{j \neq i} q_{ij} k_j^A)$$

So multiplying both sides by $q_i = -q_{ii}$ we get $-q_{ii} k_i^A = 1 + \sum_{j \neq i} q_{ij} k_j^A$ which is the desired equation for i not in A.

Proof continued

It remains to prove minimality. Again we use the same argument as for Chapter 1: we regenerate whenever possible. We need only consider i not in A. Then for $\tilde{k}^A$ another positive solution to the above equations

$$\tilde{k}_i^A = \frac{1}{q_i} + \sum_{j \neq i} \pi_{ij} \tilde{k}_j^A) = \frac{1}{q_i} + \sum_{j \notin A} \pi_{ij} \tilde{k}_j^A)$$

$$= \frac{1}{q_i} + \sum_{j \neq i, \notin A} \pi_{ij} (\frac{1}{q_j} + \sum_{l \notin A} \pi_{jl} \tilde{k}_l^A)$$

We note that $\sum_{j \neq i} \pi_{ij} (\frac{1}{q_j}$ is $\mathbb{E}_i(S_2 I_{D_A > S_1})$. Continuing with our expansion we have for each n, $\tilde{k}_i^A$

$$= \frac{1}{q_i} + \mathbb{E}_i(S_2 I_{D_A > S_1}) + \cdots \mathbb{E}_i(S_n I_{D_A > S_{n-1}}) + \text{positive term.}$$

to finish we simply observe that as n tends to infinity $\frac{1}{q_i} + \mathbb{E}_i(S_2 I_{D_A > S_1}) + \cdots \mathbb{E}_i(S_n I_{D_A > S_{n-1}})$ converges to k_i^A.

Recurrance and Transience

Definitions:

Given Q a site $i \in I$ is *recurrent* if

$$\mathbb{P}_i(\{t : X_t = i\} \text{ is unbounded}) = 1$$

and is *transient* if

$$\mathbb{P}_i(\{t : X_t = i\} \text{ is unbounded}) = 0.$$

It is immediately clear that i is recurrent for Q if and only if i is recurrent for the jump chain with transition matrix Π. In fact we have

Theorem

(i) *i is recurrent for $(X_t)_{t\geq 0}$ if it is recurrent for $(Y_n)_{n\geq 0}$,*
(ii) *i is transient for $(X_t)_{t\geq 0}$ if it is transient for $(Y_n)_{n\geq 0}$,*
(iii) *Each site is either transient or recurrent*
(iv) *Transience or Recurrence are class properties.*

The theorem is immediate from our Jump chain representation. We will, as before, speak of transient or recurrent chains if all sites i are so. In particular we will speak of transient/recurrent chains if Q is irreducible.

Expected time at i

Definition

Given Q Markov chain $(X_t)_{t\geq 0}$, we write T_i for $\inf\{t \geq J_1 : X_t = i\}$. It is immediate that

$$i \text{ is recurrent if and only if } i \text{ is absorbing or } \mathbb{P}_i(T_i < \infty) = 1.$$

We have the following analogies of the Chapter 1 criterion for recurrence/transience

Theorem

If i is recurrent if and only if $\int_0^\infty P_{ii}(t)dt = \infty$

Proof:

By Fubini's Theorem $\int_0^\infty P_{ii}(t)dt = \mathbb{E}_i(\sum_{k\geq 0} I_{Y_k=i} S_{k+1}) = \sum_{k\geq 0} \mathbb{E}_i(I_{Y_k=i} S_{k+1})$. But given $(Y_k)_{k\geq 0}$ the S_k are exponential random variables of appropriate parameter. In particular if $Y_k = i$, then $S_{k+1})$ is an exponential q_i random variable of expectation $1/q_i$, so independence yields $\int_0^\infty P_{ii}(t)dt = \frac{1}{q_i} \sum_{k\geq 0} \mathbb{P}_i(Y_k = i)$ which is finite or infinite according to whether i is transient or recurrent by Chapter 1.

The h Skeleton

For a Markov chain $(X_t)_{t\geq 0}$, the discrete time process $(Z_n)_{n\geq 0} \equiv (X_{nh})_{n\geq 0}$ is a Markov chain by the semigroup characterization of our Markov chain X with transition matrix given by $P_{ij}(h)$ (Strictly speaking (when explosions are possible, this Matrix is a sub probability matrix but if this bothers you simply adjoin a value ∞ to I.

Theorem

for any $h > 0$ and $i \in I$, i is recurrent for X if and only if i is recurrent for Z.

Proof:
If i is transient for X then $\mathbb{P}_i$ a.s. the times t for which $X_t = i$ form a bounded set. This certainly implies that the n such that $Z_n = i$ form a bounded (which is to say finite) set under $\mathbb{P}_i$. That is i is transient. If i is recurrent for X then $\int_0^\infty P_{ii}(t)dt = \sum_{n\geq 0} \int_{nh}^{(n+1)h} P_{ii}(t)dt = \infty$. But by the semigroup property

$$\forall t \in [nh, (n+1)h] P_{ii}((n+1)h) \geq P_{ii}(t)e^{-q^i((n+1)h-t)} \geq P_{ii}(t)e^{-q^i h}$$

Thus $h\sum_{n\geq 1} P(Z_n = i) \geq e^{-q_i h} \int_{nh}^{(n+1)h} P_{ii}(t)dt = \infty$. Again, we conclude by Chapter 1 that i is recurrent for Z.