Continuous time Markov Chains

Thomas Mountford

EPFL

May 17, 2021

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.

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 $q_i = 0$).

Invariant probabilities, positive recurrence

Definition: Given Q-matrix Q, a measure λ on I is invariant if $\lambda Q = \overline{0}$.

Note that IF I is finite and λ is invariant, then we have the backward equation and

$$(\lambda P(t))' = \lambda P(t)' = \lambda Q P(t) = \overline{0}.$$

This implies that for all $t > 0, \lambda P(t)$ is constant and so for $t > 0, \lambda P(t) = \lambda P(0) = \lambda$. That is to say if X_0 has law λ, then so does X_t for every $t > 0$. This justifies the description that λ is *invariant* .

BUT

Things become less clear when I is countably infinite. We introduce the example 3.5.4. We consider a Birth and Death chain on $\mathbb{N}$. We note that if for every $i \geq 0, \nu_i q_{ii+1} = \nu_{i+1} q_{i+1i}$, then for every i

$$(\nu Q)_i = \nu_{i+1} q_{i+1i} + \nu_{i-1} q_{i-1i} - \nu_i q_{ii}$$

$$= \nu_i q_{ii+1} + \nu_i q_{ii-1} - \nu_i q_{ii}$$

$$= \nu_i (q_{ii+1} + q_{ii-1} - q_{ii}) = 0$$

But we can always (given $q_{ii-1}, q_{ii+1} > 0$ for possible values) find ν satisfying the detailed balence equations. For our example we take $\forall i > 0,\ q_{ii-1} = \mu q_i$ and $q_{ii+1} = \ \lambda q_i$ for arbitrary q_i. We also take $q_{01} = q_0 \lambda$

continued.

In this case the solution is a multiple of $\nu_i = \frac{1}{q_i}(\frac{\lambda}{\mu})^i$ We now get concrete.

(i) $1 < \frac{\lambda}{\mu} < 2$

(ii) $q_i = 2^i$

The condition that $1 < \frac{\lambda}{\mu}$ ensures that our jump chain is transient and so all sites for our Markov chain (continuous time or discete) are transient. Condition $q_i = 2^i$ implies that $\sum_i \frac{1}{q_i} < \infty$ which implies (as is easily seen) that the continuous time chain is explosive with probability one. Finally $\frac{\lambda}{\mu} < 2$ and the choice of q_i (ii) implies that $\sum_i \nu_i < \infty$ and so for some positive c, $c\lambda$ is an invariant distribution.

So we have an invariant distribution for which the chain is (highly) transient. Indeed, for which $\nu P(t)$ converges to $\overline{0}$ very rapidly as t becomes large.

Theorem

Let Π be the jump chain for associated Q-matrix Q. Then measure λ is invariant for Q if and only if $\mu : \mu_i = \lambda_i q_i$ is invariant for Π

Proof: we use the fact that

$$\forall i, j \quad q_i(\pi_{ij} - \delta_{ij}) = q_{ij}$$

(Check the special cases $q_i = 0$ and $i = j$ separately). So if $Q\lambda = 0$, then for every i $\sum_j \lambda_j q_{ji} = 0 = \sum_j \lambda_j q_j(\pi_{ji} - \delta_{ji}) =$
$= \sum_j \mu_j(\pi_{ji} - \delta_{ji}) = \sum_j \mu_j \pi_{ji} - \mu_i$
We remark that this should not be unexpected in that if μ is a stationary distribution for (suppose) irreducible jump matrix Π, then the long run ratio of times spent (by the continuous time chain) in sites i and j

should be $\frac{\mu_i/q_i}{\mu_i j q_j}$ which "ought" to be invariant for Q.
A second remark is that if $\sum_i \lambda_i$ is finite, there is no reason why $\sum_i \lambda_i q_i$ should not be infinite and vice versa.

Recall

Theorem

For an irreducible recurrent Markov chain fix $k \in I$ and define

$$\gamma(j) = \mathbb{E}_k\left(\sum_{r<T_k} I_{X_r=j}\right)$$

Then

- $\gamma(k) = 1$
- $0 < \gamma(j) < \infty$
- γ *is invariant.*

and in this case γ is unique up to multiplication by positive constants.If we do not have recurrence then if ν is an invariant measure with $\nu(k) = 1$, then $\nu \geq \gamma$.

From this we obtain immediately

Corollary

If Q is irreducible and recurrent then there exists an invariant measure λ and it is unique up to positive multiples.

Positive Recurrence and invariance

We have for $i \in I$, the return time $\inf\{t \geq S_1 : X_t = i\}$ is denoted by T_i.
Definition: A site i is *positive recurrent* if either $q_i = 0$ or $\mathbb{E}_i(T_i) < \infty$.
Obviously if i is positive recurrent, then it is recurrent. Just as in Chapter 1, we have that i is positive recurrent if and only if

$$\liminf_{t\to\infty} \frac{1}{t}\int_0^t P_{ii}(s)ds > 0$$

if the chain is irreducible this is if and only if for any j

$$\liminf_{t\to\infty} \frac{1}{t}\int_0^t P_{ji}(s)ds > 0$$

If i leads to j

$$\frac{1}{t}\int_0^t P_{jj}(s)ds \geq \frac{t-2}{t}\frac{1}{t-2}\int_0^{t-2} P_{ji}(1)P_{ii}(s)P_{ij}(1)ds > 0$$

This gives

Theorem

For Q-matrix Q positive recurrence is a class property.

We now show that positive recurrence is related to invariant distributions

Theorem

For irreducible Q-matrix Q the following are equivalent

(i) *The chain is positive recurrent*

(ii) *The chain is nonexplosvie and has an invariant distribution.*

If either condition holds then $\mathbb{E}_i(T_i) = 1/(\lambda_i q_i)$ for λ the invariant distribution.

Remark the Birth and Death chain example shows that we need the condition that Q be nonexplosive.

Proof (i) implies (ii)

Let i be in I. It is positive recurrent under assumption (i). So it is recurrent and the chain is nonexplosive. We define measure

$$\mu^i_j = \mathbb{E}_i(\int_0^{T_i} I_{X_s=j} ds)$$

we easily have that $\mu^i_j = \mathbb{E}_i(\sum_{k\geq 0} I_{k<N_i} I_{Y_k=j} S_{k+1}$ which by description a) of Markov chains is equal to $\frac{1}{q_j}\mathbb{E}_i(\sum_{k\geq 0} I_{k<N_i} I_{Y_k=j})$. This is equal to γ^i_j/q_j where $\gamma^i_j = \mathbb{E}_i(\sum_{k\geq 0} I_{k<N_i} I_{Y_k=j}$ is Π invariant. Thus by the preceding Theorem , μ^i_j is an invariant measure for Q. But by definition $\sum_j \mu^i_j = \mathbb{E}_i(T_i) < \infty$ by assumption (i). Thus we have an invariant finite measure, and so after dividing by the total mass, an invariant distribution. (ii) is shown.

Proof (ii) implies (i)

We suppose (ii). Let λ be an invariant distribution. Then by the preceding Theorem, $\lambda_j q_j$ is an invariant measure for Π. Thus for i a fixed site in I, $\nu_j = \lambda_j q_j/(q_i\lambda_i)$ is invariant and $\nu_i = 1$. So by Chapter 1 $\nu_j \geq \gamma_j^i(\gamma^i$ as above). We cannot yet claim equality as we do not yet know that the chain is recurrent! But

$$m_i = \sum_j \mu_j^i = \sum_j \gamma_j^i/q_j \leq \sum_j \nu_j/q_j = \sum_j \lambda_j/(q_i\lambda_i)$$

$= 1/(q_i\lambda_i) < \infty$. So the Markov chain is positive recurrent. That is (i). But since it is recurrent we can go back and have equality between ν and γ_j^i which gives equality between m_i and $1/(q_i\lambda_i)$.

Invariant distributions and stationarity

To finish this section it remains to show

Theorem

For irreducible recurrent Q-matrix Q, λ a measure and $s > 0$, the following are equivalent

(i) $\lambda Q = \overline{0}$

(ii) $\lambda P(s) = \lambda$

The chain is recurrent and thus nonexplosive. From the preceding slides

$$\mu_j^i = \mathbb{E}_i\left(\int_0^{T_i} I_{X_t = j} dt\right)$$

is Q invariant. Thus it is a positive multiple of λ. So it is enough to show that (i) and (ii) are equivalent for measure μ_j^i. The trick is to apply the Strong markov property at T_i to see that

$$\mathbb{E}_i\left(\int_0^{s} I_{X_t = j} dt\right) = \mathbb{E}_i\left(\int_{T_i}^{T_i + s} I_{X_t = j} dt\right).$$

Accordingly

$$\mu_j^i = \mathbb{E}_i\left(\int_0^{T_i} I_{X_t=j}dt\right) =$$

$$\mathbb{E}_i\left(\int_0^{T_i} I_{X_t=j}dt\right) - \mathbb{E}_i\left(\int_0^{s} I_{X_t=j}dt\right) + \mathbb{E}_i\left(\int_{T_i}^{T_i+s} I_{X_t=j}dt\right)$$

which equals $\mathbb{E}_i(\int_s^{T_i+s} I_{X_t=j}dt)$. But we can rewrite this as

$$\mathbb{E}_i\left(\int_s^{\infty} I_{t<T_i+s} I_{X_t=j}dt\right)$$

changing variables to $u = t - s$ gives

$$= \mathbb{E}_i\left(\int_0^{\infty} I_{u<T_i} I_{X_{u+s}=j}du\right) = \mathbb{E}_i\left(\int_0^{\infty} I_{u<T_i} \sum_k I_{X_u=k} P_{kj}(s)du\right) = \sum_k \mu_k^i P_{kj}(s)$$