

MAS275 Probability Modelling Chapter 3: Limiting behaviour of Markov chains

Dimitrios Kiagias

School of Mathematics and Statistics, University of Sheffield

Spring Semester, 2020

Returns to start

Suppose a Markov chain (X_n) is started in a particular fixed state i .

If it returns to i at some later (random) time, then, because of the Markov property and time homogeneity, the future behaviour will be as if the process had been started off in state i at this later time.

Hence, returns to the starting state form a renewal process.

Notation

In this renewal process,

$$u_n = p_{ii}^{(n)},$$

where $p_{ii}^{(n)}$ denotes the (i, i) element of the n -step transition matrix $P^{(n)} = P^n$.

The “ f_n ” of the renewal process will then be the probability, starting in i , of the first return to state i being at time n , namely

$$P(X_1 \neq i, X_2 \neq i, \dots, X_{n-1} \neq i, X_n = i | X_0 = i).$$

We will write $f_{ii}^{(n)}$ for this.

Notation II

Also define

$$f_{ii} = \sum_{n=1}^{\infty} f_{ii}^{(n)},$$

i.e. the probability, starting in state i , of ever returning to state i .

If the renewal process is recurrent this will be 1, and the process will keep returning to i .

If it is transient then it will be strictly less than 1, and eventually, with probability 1, the process will visit state i for the last time and never return.

Classification of states

We can refer to any state i of a Markov chain as being **transient** or **recurrent** if the corresponding renewal process is.

Similarly, if state i is recurrent, we can describe it as **positive recurrent** or **null recurrent** if the corresponding renewal process is, . . .

. . . and we can define the **period** of state i as the same as the period of the corresponding renewal process. Again, we describe a state with period 1 as **aperiodic**.

These are called the **recurrence properties** of the states.

Examples

Example

Simple random walk

Example

Mean recurrence time

Delayed renewal processes

If we now take two different states, i and j say, and consider visits to state j having started in state i , then, by a similar argument to the above, we have a delayed renewal process, the delay being the time until the first visit to j .

Notation

We can identify the “ b_n ” and “ v_n ” of this delayed renewal process as

$$b_n = f_{ij}^{(n)} = P(X_1 \neq j, X_2 \neq j, \dots, X_{n-1} \neq j, X_n = j | X_0 = i)$$

(that is, the probability that the first visit to state j happens at time n , starting from state i) and

$$v_n = p_{ij}^{(n)}$$

where $p_{ij}^{(n)}$ is the (i, j) -th element of the n -step transition matrix.

First passage

Note that it may be possible, starting in state i , never to visit state j , in which case the delay distribution is defective.

In fact

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}$$

is the probability, starting in state i , of ever visiting state j .

The probabilities $f_{ij}^{(n)}$ are called **first passage** probabilities.

Finite state space

Finite state spaces tend to be slightly easier to handle, as for example in the following result.

Theorem

If the state space is finite then some of the states must be positive recurrent, and the rest, if there are any, must be transient.

Proof of Theorem 8 – not all transient

In a finite state space the states cannot all be transient.

If they were, after the process had left each of the states for the last time there would be nowhere left for it to go.

So some states are recurrent.

Proof of Theorem 8 – not null recurrent I

If a state i is recurrent, then, starting in i , then consider V_j , the number of visits to another state, j say, before the first return to i .

Let q_{ij} be the probability of visiting j before returning to i , starting in i , and q_{ji} be the probability of visiting i before returning to j .

Then we can write

$$P(V_j = 0) = 1 - q_{ij}$$

and, for $n \geq 1$,

$$P(V_j = n) = q_{ij}(1 - q_{ij})^{n-1}q_{ji}.$$

Proof of Theorem 8 – not null recurrent II

The distribution of V_j has finite mean (by essentially the same calculations as for the geometric distribution).

So, as we are in a finite state space, the total time spent away from i before the first return has finite mean, i.e. the state i is positive recurrent.

Hence there are no null recurrent states in a chain with a finite state space. □

Communication

Two states i and j of a Markov chain are said to **communicate** with each other if it is possible, starting in i , to get to j (not necessarily at the next step) and similarly possible, starting in j , to get to i .

Formally, there exist positive integers m and n such that $p_{ij}^{(m)} > 0$ and $p_{ji}^{(n)} > 0$.

Equivalence relation – symmetry

This relationship is obviously **symmetric**, in the sense that if i and j communicate then so do j and i .

This is implicit in the way we have defined communication.

Equivalence relation – transitivity

If i communicates with j and also j communicates with k , then i communicates with k .

Formally, this follows from the inequality

$$p_{ik}^{(m+r)} \geq p_{ij}^{(m)} p_{jk}^{(r)}$$

which is a straightforward consequence of the Chapman-Kolmogorov equations.

If we can find m and r such that the right hand side is positive, then the left hand side must be positive.

This property of the communication relationship is called

Equivalence relation – reflexivity

If we also, by convention, define a state to communicate with itself, then we have a relationship which is **reflexive**.

A binary relationship between members of a set is called an **equivalence relation** if it is reflexive, symmetric and transitive.

Equivalence classes

If we have all three of these properties then the set in question, in this case the state space of the Markov chain, is partitioned into **equivalence classes**, such that . . .

. . . any two members of the same class communicate with each other, and any two members of two different classes do not communicate with each other.

These are called **communicating classes** or simply classes.

Irreducibility

A Markov chain is called **irreducible** if all its states communicate with each other, i.e. they form a single large class.

Examples

Example

Irreducible Markov chains

Example

Gambler's ruin

Solidarity

The solidarity property can be very simply stated as follows.

Theorem

All states within a communicating class share the same recurrence properties.

Use of theorem

This theorem means, in particular, that if the Markov chain is irreducible then all states share the same recurrence properties, and it is meaningful to refer to the recurrence properties of the whole Markov chain.

In any event, we can refer to the recurrence properties of classes as well as those of individual states.

Corollary

We can immediately conclude the following:

Corollary

An irreducible chain with finite state space has all its states positive recurrent.

Proof By Theorem 8, some of the states must be positive recurrent.

Hence by solidarity (Theorem 9) they all are. □

Closed class

A class C is called **closed** if, once entered, the probability of leaving is zero.

It is also sometimes known as an **absorbing class**.

Formally, C is closed if $p_{ij} = 0$ for all $i \in C$ and $j \notin C$.

In the irreducible case there is a single class $C = S$, which is automatically closed.

Not closed

If a class is not closed, then, once it has been left, there is no possibility of returning to it.

If there were such a possibility then there would be at least one state outside the class which communicates with it, which is a contradiction.

Finite state space

If the state space S is finite, then the concept of a closed class is equivalent to that of a recurrent class, and that of a class which is not closed is equivalent to that of a transient class.

This is because, if a class is recurrent, then it is impossible to leave it, because if it were possible and it happened, then return to the class could not take place, and this contradicts recurrence.

Conversely if a class is transient then each state of the class is only visited finitely many times and so the class must eventually be left.

Infinite state space

If the state space is infinite, there are more possibilities.

For example, in the simple random walk with $p \neq \frac{1}{2}$, all states are transient but they form a single closed class.

Renewal processes and Markov chains

Communication

Solidarity of recurrence properties within classes

Limiting/equilibrium behaviour

Non-irreducible and periodic chains

The renewal theorem

Example

Example

Finding classes and recurrence properties

Limiting/equilibrium behaviour

The aim of this section is to investigate how the distribution of the state a Markov chain evolves with time.

In particular we will see that many chains evolve towards an equilibrium behaviour described by a stationary distribution.

Transient case

We will first of all consider one case where this does not happen: when the chain is transient.

Theorem

Assume we have a Markov chain (X_n) which is irreducible and transient (which necessarily means that its state space S is infinite). Then, for each state $i \in S$, $P(X_n = i) \rightarrow 0$ as $n \rightarrow \infty$.

Furthermore, there is no stationary distribution.

Stationary distribution exists

We now consider the case where the Markov chain is irreducible and aperiodic, and we know that a stationary distribution exists. By Theorem 11 we know that the chain must be recurrent.

Theorem

If a Markov chain (X_n) is irreducible and aperiodic and has a stationary distribution π then the distribution of the state of the Markov chain at time n converges to π .

I.e.

$$\pi^{(n)} \rightarrow \pi$$

as $n \rightarrow \infty$.

Proof of Theorem 12 – Coupling

We can prove this theorem using a “coupling” argument.

The idea of this argument is to consider a second Markov chain, (Y_n) , which has the same transition matrix as the original chain (X_n) but has initial distribution given by the stationary distribution π .

The two Markov chains run independently of each other.

Then $P(X_n = i) = \pi_i^{(n)}$, while $P(Y_n = i) = \pi_i$.

Proof of Theorem 12 – If chains meet

Consider what happens if at some time m , the two chains are in the same state: $X_m = Y_m$.

Then, by the Markov property, what happened before time m does not affect the probabilities of what happens after time n .

So for any $n > m$ the conditional probabilities that $X_n = i$ and that $Y_n = i$ will be the same.

(More formal argument in notes.)

Proof of Theorem 12 – Bivariate Markov chain I

Consider the bivariate Markov chain on $S \times S$ obtained by considering the two chains together, so that its state at time n is (X_n, Y_n) .

By the irreducibility and aperiodicity of our original chain, we know that for any choice of i and j it is the case that for n large enough $p_{ij}^{(n)} > 0$.

Proof of Theorem 12 – Bivariate Markov chain II

So for any pair of states (i_1, i_2) and (j_1, j_2) of the bivariate Markov chain, if n is large enough both $p_{i_1, j_1}^{(n)}$ and $p_{i_2, j_2}^{(n)}$ will be positive, showing that it is possible for the bivariate Markov chain to get from (i_1, i_2) to (j_1, j_2) .

Hence the bivariate Markov chain is irreducible.

(Note that aperiodicity is crucial to this argument.)

Proof of Theorem 12 – Stationary distribution

It also has a stationary distribution, given by

$$P(X_n = i, Y_n = j) = \pi_i \pi_j$$

by independence.

An irreducible Markov chain with a stationary distribution cannot be transient, by Theorem 11.

Proof of Theorem 12 – Conclusion

Hence it is recurrent, and so all states are visited, with probability 1.

This includes states where the two co-ordinates are the same, so this shows that the chains (X_n) and (Y_n) will meet, with probability 1, and thus completes the proof of the theorem. □

Corollary

Corollary

If a Markov chain (X_n) is irreducible and aperiodic, and has a stationary distribution π , then that stationary distribution is unique.

Proof: If another stationary distribution ψ exists, then we can start a chain with the same transition matrix with initial distribution $\pi^{(0)} = \psi$.

By stationarity we have $\pi^{(n)} = \psi$ for all n and hence $\lim_{n \rightarrow \infty} \pi_i^{(n)} = \psi_i$.

But we have shown above that this must be π_i .



Finding a stationary distribution

Now, we show that if we have an irreducible and positive recurrent Markov chain we can construct a stationary distribution.

For each state j , let μ_j be the expected time until the first return to state j , given that the chain starts there.

Write π for the row vector with entry j being $\frac{1}{\mu_j}$, for $j \in S$.

Theorem

The following theorem gives the main results of this section.

Theorem

Assume the Markov chain is irreducible, aperiodic and positive recurrent. Then

- (a) The distribution π is a stationary distribution, and is unique.*
- (b) The distribution of the state of the Markov chain at time n converges to π , i.e.*

$$\pi^{(n)} \rightarrow \pi$$

as $n \rightarrow \infty$.

Using the results

Note that because we know that the that π is a unique stationary distribution, it is usually easiest to find it by solving the equilibrium equations of section 1.7, and then to apply the above results.

Renewal processes and Markov chains

Communication

Solidarity of recurrence properties within classes

Limiting/equilibrium behaviour

Non-irreducible and periodic chains

The renewal theorem

Examples

Example

Modelling the game of Monopoly

Null recurrent

Finally we note that the results of Theorem 11 (that for any state i $P(X_n = i) \rightarrow 0$ as $n \rightarrow \infty$, and that there are no stationary distributions) apply to irreducible null recurrent chains as well as irreducible transient chains.

This is a bit harder to prove, and we omit the proof.

Summary of results

- If a chain has finite state space and is irreducible and all states are aperiodic, then it is positive recurrent by Corollary 10, and by Theorem 14 it has a unique stationary distribution, to which the distribution of the chain at time n converges as $n \rightarrow \infty$.
- If a chain with infinite state space is irreducible, aperiodic and positive recurrent, then again by Theorem 14 it has a unique stationary distribution, to which the distribution of the chain at time n converges as $n \rightarrow \infty$.
- If a chain with infinite state space is irreducible and is either transient or null recurrent, then it does not have a stationary distribution and for any state i $P(X_n = i) \rightarrow 0$ as $n \rightarrow \infty$.

Non-irreducible chains

Consider a Markov chain which has a closed class C which is positive recurrent and not the whole state space.

Then the behaviour of the process if it starts inside C is that of a Markov chain in which the states outside C have been removed from the state space, and the corresponding rows and columns of the transition matrix have been deleted.

This “reduced” Markov chain will be irreducible and will have a stationary distribution.

If this distribution is extended to the original state space by assigning probability zero to each of the states outside C , we

More than one closed class

If there are two or more positive recurrent closed classes, then the stationary distribution is not unique: there will be, at least, one corresponding to each of these classes.

Furthermore, any “mixture” of the form

$$\alpha\pi + (1 - \alpha)\psi$$

where π and ψ are stationary distributions and $0 < \alpha < 1$, will also be a stationary distribution.

So there will be a continuum of them.

Algebraic interpretation

In the case of a finite state space, this situation corresponds to the space of left eigenvectors of the transition matrix corresponding to eigenvalue 1 being more than one-dimensional, and the eigenvalue 1 having multiplicity greater than 1.

Renewal processes and Markov chains

Communication

Solidarity of recurrence properties within classes

Limiting/equilibrium behaviour

Non-irreducible and periodic chains

The renewal theorem

Example

Example

A non-irreducible chain

Periodicity

Periodicity is a nuisance when it comes to looking at limiting behaviour.

The following example illustrates how things can differ from the aperiodic case.

Example

A periodic chain

Back to renewal processes

It is possible to apply the ideas of long-run behaviour of Markov chains to renewal theory.

Given a renewal process, recall the following notation:

f_k is the probability that a given inter-renewal time takes the value k

u_k is the probability that there is a renewal at time k

d is the period of the renewal process

μ is the mean inter-recurrence time, if the process is positive recurrent

Notation

Also define K to be the largest value of k such that f_k is non-zero (i.e. the largest possible inter-renewal time) if this is finite.

Given a renewal process, let X_n be the number of time steps from time n until the next renewal; this is called the **forward recurrence time**.

(Set $X_n = 0$ if there is a renewal at time n .)

Transition probabilities

Then

- $P(X_{n+1} = i - 1 | X_n = i) = 1$ unless $i = 0$
- for $i \geq 0$ we have $P(X_{n+1} = i | X_n = 0) = f_{i+1}$

(To see the latter, note that $X_n = 0$ means there is a renewal at time n and that $X_{n+1} = i$ means that the next renewal after the one at time n occurs at time $n + i + 1$, which has probability f_{i+1} .)

Transition matrix

By the construction of the renewal process, the process (X_n) satisfies the Markov property.

So it is a Markov chain with transition matrix

$$\begin{pmatrix} f_1 & f_2 & f_3 & f_4 & \dots \\ 1 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

and state space $\{0, 1, 2, \dots, K - 1\}$ if K is finite, and the non-negative integers \mathbb{N}_0 otherwise.

Properties of the chain

This chain is irreducible.

If the renewal process is positive recurrent, then the Markov chain is too (we can see this by considering returns to state zero) and it has a stationary distribution given by

$$\pi_i = \frac{1}{\mu}(1 - f_1 - \dots - f_i)$$

for $i = 0, 1, 2, \dots$

Similarly, if the renewal process is transient or null recurrent, then the Markov chain will be too.

Renewal Theorem

We can use the results on Markov chains to deduce the following theorem.

Theorem

(Erdős, Feller, Pollard) (The Discrete Renewal Theorem)
If a renewal process is recurrent and has period d and finite mean inter-renewal time μ , then

$$u_{nd} \rightarrow \frac{d}{\mu} \text{ as } n \rightarrow \infty.$$

In all other cases,

$$u_n \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Proof of renewal theorem

We observe that $u_n = P(X_n = 0)$, as both are the probability that we have a renewal at time n .

So, if the renewal process is aperiodic and positive recurrent, we can immediately deduce $u_n \rightarrow \pi_0 = \frac{1}{\mu}$ from the results on Markov chains.

If the renewal process is transient or null recurrent, then we can similarly deduce $u_n \rightarrow 0$ as $n \rightarrow \infty$.

Proof of renewal theorem – periodic case

We can handle the periodic case by defining a new aperiodic renewal process which has renewals at time n if our periodic renewal process has a renewal at time nd , and applying the theorem to that.

Renewal processes and Markov chains

Communication

Solidarity of recurrence properties within classes

Limiting/equilibrium behaviour

Non-irreducible and periodic chains

The renewal theorem

Example

Example

Bernoulli trials with blocking